Genetic analysis of breeding-related traits in *Brassica rapa*

Hedayat Bagheri

Thesis committee

Thesis supervisor
Prof. dr. ir. Maarten Koornneef
Professor of Genetics
Wageningen University

Thesis co-supervisor
Dr. ir. Mark G.M. Aarts
Associate Professor, Laboratory of Genetics
Wageningen University

Other members
Prof. dr. ir. Evert Jacobsen, Wageningen University
Prof. dr. ir. Harro J. Bouwmeester, Wageningen University
Prof. dr. Christian Jung, Christian-Albrechts-University of Kiel, Germany
Dr. ir. Guusje Bonnema, Wageningen University

This research has been conducted under the auspices of the Graduate school of Experimental Plant Sciences.

Genetic analysis of breeding-related traits in *Brassica rapa*

Hedayat Bagheri

Thesis

Submitted in partial fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Doctorate Board
to be defended in public
on Tuesday 15 December 2009
at 1:30 PM in the Aula

Hedayat Bagheri
Genetic analysis of breeding-related traits in *Brassica rapa,*
130 pages.

Thesis, Wageningen University, Wageningen, NL (2009)
With references, with summaries in Dutch and English.

ISBN 978-90-8585-549-1

Contents

Chapter 1	General introduction	1
Chapter 2	*Agrobacterium*-mediated transformation of a transposon construct into *Brassica rapa*	17
Chapter 3	Identification of seed related QTLs in a new genetically segregating *Brassica rapa* F2 population	43
Chapter 4	Genetic mapping and QTL analysis of plant morphology and seed related traits in new *Brassica rapa* recombinant inbred line population	67
Chapter 5	Genetic analysis of the concentrations of health-related compounds in *Brassica rapa*	89
Chapter 6	General discussion	111
	Summary	121
	Samenvatting	123
	Acknowledgements	125
	About the author	127
	Education certificate	129

Dedicated to my beloved wife, Taeibah, and my lovely daughter Mozhdeh/11

Chapter 1

General introduction

Brassica rapa

Brassica rapa L. is a diploid crop in the family of *Brassicaceae.* This family contains about 3500 species and 350 genera and is one of the economically most important plant families (Rich, 1991, Quijada et al., 2007). The family of *Brassicaceae* is an important source of edible roots, stems, leaves, buds and inflorescences, as well as of edible or industrial oils, condiments and forage. *B. rapa* and *B. campestris* were first described as two species by Linnaeus, with *B. rapa* being the turnip form and *B. campestris* the wild weedy form. Later on it was shown that these were the same species so the taxa were combined under the name *B. rapa* (Toxopeus et al., 1984).

There are six *Brassica* species which have the highest agricultural importance and are referred to as 'crop Brassicas' (Gómez-Campo 1999). The triangle of U (U, 1935) shows the relationship between these 'crop Brassicas'. Initially three ancestral diploid species: *B. rapa* (AA, n=10), *B. nigra* (black mustard) (BB, n=8) and *B. oleracea* (CC, n=9) existed. Through spontaneous hybridization followed by chromosome doubling, three amphidiploid species emerged: *B. napus* (AACC, n=19), *B. carinata* (BBCC, n=17) and *B. juncea* (AABB, n=18). Artificial resynthesis of *B. napus* from *B. rapa* and *B. oleracea* showed the same agreement with the triangle of U (Olsson, 1960), which has been confirmed by molecular analysis (Warwick and Black, 1991).

B. rapa, which is the putative ancestor of many oriental *Brassica* vegetables, originates from the high plateau regions in today's Iran–Iraq–Turkey (Dixon, 2007) and seems to have spread naturally to the Western Mediterranean region and to Central Asia. *B. rapa* is the first domesticated *Brassica* crop and it has the widest distribution, with secondary centers of diversity in Europe, Western Russia, Central Asia, and the Near East (Quijada et al., 2007). Various *Brassica* leafy vegetables have been differentiated in East Asia.

Wild *B. rapa* subssp. *oleifera* is regarded as the subspecies from which var. *rapa* L. (cultivated turnip) and *var. silvestris* (Lam.) Briggs (turnip-rape) originated. Prakash and Hinata (1980) suggest that oleiferous *B. rapa* developed in two places giving rise to two different races, one European and the other Asian. The Asian or Indian forms: ssp. *trilocularis*, known as "yellow sarson" and ssp. *dichotoma* known both as "toria" and "brown sarson"; and the west European and north American form: *B. rapa* ssp. o*leifera.* The earliest reference to yellow sarson was in the Sanskrit book by Upanisadas and Brahamanas (c. 1500 BC) where it was referred to as 'Siddhartha' (Prakash, 1961) and used as an oil plant. Yellow sarson is characterized by yellow colored seeds and self-compatibility. Because many of the cultivars have three to four carpels, it was previously named *B. trilocularis* (Gómez-Campo, 1999). In Europe, the first oil types of *B. rapa* are believed to have been planted in the early middle ages and commercial plantings of rapeseed were recorded in the Netherlands as early as the 16th century. Rapeseed oil was used primarily as oil for lamps. Later it became used as

a lubricant for steam engines. Although widely used as edible oil in Asia, through breeding for improved oil quality and development of improved processing techniques, *B. rapa* has become important in western nations. Since the Second World War, rapeseed production (now predominantly *B. napus*) in Europe and Canada has increased dramatically as a result of improved oil and meal quality.

In addition to the oil and turnip types, there are many leafy and inflorescence vegetable types of *B. rapa* such as, var. *campestris*, var. *pekinensis*, var. *chinensis*, var. *parachinensis*, var. *narinosa*, and var. *japonica*, predominantly originating from Asia. Zhao et al. (2005) used AFLP fingerprints on 161 *B. rapa* accessions collected from different parts of the world, and divided them in two main groups: one of Asian and the second of mainly European origin. There are different morphotypes (morphological variation within a species population) in both groups. He concluded that either there is an independent origin for both groups and/or a long and separate domestication and breeding history in both regions (Zhao et al., 2005).

Breeding

B. rapa is an important oil and vegetable crop in many parts of the world, with seed used for oil, and leaves, flowers, stems and roots used as vegetables. It is also an important source of healthy food components, such as vitamins, minerals and anti-carcinogenic compounds. In general, breeding has focussed on important production traits, such as yield and quality. In Asian countries, greater seed yield and yield stability were the primary objectives, besides breeding for quality of the harvested product (Quijada et al., 2007). The seed yield is the resultant of several determining components, i.e., the number of siliques per unit area (determined as the number of siliques per plant and the number of plants per unit area), branch number, number of seeds per silique, and seed size.

The oil content in air-dried seeds varies between 36 and 44 % for *B. rapa*. The yellow seed coat trait has been shown to be associated with lower fiber content and higher oil and protein content (Stringam et al., 1974). Oil content is influenced by the environment, particularly temperature and moisture stress, and soil nitrogen, but there is also genetic variation for oil content in *B. rapa*. Selection for oil content has led to slow but steady improvement. During the last two decades, *Brassica* species became the second largest oilseed crop after soybean (Scarth and Tang 2006), mainly due to the growth of *B. napus* as oil crop. Seed oil quality is determined by its fatty acid composition, while antinutritional factors, particularly glucosinolate levels and protein/fiber proportions, determine meal quality (Quijada et al., 2007).

Rapeseed oil on average has 65% γ-tocopherol and 35% α-tocopherol (Goffman and Becker, 2001). Tocopherols are essential nutrients that act as anti-oxidants in human food. *B. rapa* is a valuable source of diverse health-promoting antioxidant metabolites. Hydrogen-1 Nuclear

Magnetic Resonance (^{1}H NMR) spectroscopy has been used to uncover qualitative and quantitative differences of various cultivars of *B. rapa*; and different cultivars were differentiated by elucidated metabolites, several organic and amino acids, carbohydrates, adenine, indole acetic acid (IAA), phenylpropanoids, flavonoids and glucosinolates (Abdel-Farid et al., 2007). Glucosinolates are sulphur-containing substances that are broken down upon cell disruption by the enzyme myrosinase to give bitter-tasting, toxic, and goitrogenic compounds. The need for modification of the fatty acid composition of the oil and the elimination of glucosinolates from the seed meal attracted great attention in the past from *Brassica* breeders. Erucic acid has been identified as an anti-nutritional compound for human consumption. The first low erucic acid *B. rapa* cultivar Span (Downey et al., 1975) was obtained by introducing the responsible alleles into an adapted cultivar resulting in the development of nutritionally superior Canola (=**Can**adian **o**il, **l**ow **a**cid) cultivars. On the other hand, demand for higher levels of erucic acid for the industrial oil market has encouraged breeders to try to produce cultivars with levels greater than 50%. Rapeseed (*B. napus* and *B. rapa*) methyl ester (biodiesel), which is environmentally friendlier than fossil fuel, has become an important fuel for diesel engines in the EU (Rakow 2004). The EU Directive on biofuels requires member states to have a minimum 5.75% of biofuels in all fuel sold for transport by the end of 2010. Biofuels are important as they can reduce carbon emissions by up to 70%.

The main types of oilseed rape currently grown are “Double low” and “HEAR”. “Double low” (“00”) commercial varieties of both *B. napus* and *B. rapa* dominate the oilseed Brassica production area in developed countries. These varieties are grown for food and characterized as having a very low (<1%) content of erucic acid in the fatty acid profile of the seed storage lipid and a very low content of glucosinolates (<18 µmoles per gram seed at 8.5% moisture) in their seed and meal. The term “canola quality” is normally applied to seed, oil and meal from such varieties. High erucic acid rape (HEAR) varieties are grown specifically for their erucic acid content - typically 50-60 % of oil.

Except for oil seed types, most major types of *B. rapa* are self-incompatible and pure line selection is complicated, labor intensive and often leading to severe inbreeding depression. Therefore, most cultivars are produced by mass or family selection. As a consequence, single cultivars often have a wide genotypic variation, even if they show a uniform morphological appearance. Nevertheless some hybrid varieties are available. The first hybrid cultivar of Chinese cabbage was released by the Takii Seed Co., Japan, in 1950 (Quijada et al., 2007). Hybrid cultivars show a more uniform appearance, and are suitable for simultaneous harvesting. As such, they are welcomed by growers and already have a big market share. Self-incompatibility is used as the system to prevent self-fertilization in case of hybrid seed production.

Chinese cabbage, a vegetable crop of *B. rapa,* has been recorded in ancient China and Korea. Chinese cabbage types are as diverse as *B. oleracea cabbage types.* Heading versus non-heading, shape, size and storability of the head are important traits in breeding vegetable *B. rapa* types. Tightly heading Chinese cabbage is a relatively new crop, first recorded in China in the eighteenth century. This crop may have originated from either a non-heading or a loosely heading type and is now widely cultivated in China, Japan, and Korea (Hirai and Matsumoto, 2007). As Chinese cabbage is cultivated all year round, cultivars with various types of growth characters are bred and used. Early-maturing types can be harvested within 50 days after sowing, while late-maturing Chinese cabbage types require prolonged cultivation for as long as 120 days. Resistances to high and low temperatures are needed for summer and winter cultivation, respectively. Bolting resistance is required in cultivars for spring harvest. Chinese cabbage is susceptible to almost all the pathogens, pests and environmental stresses affecting *B. oleracea* (Dixon, 2007).

Arabidopsis thaliana, another species of the *Brassicaceae* family, is the model system of choice for research in plant biology and genome analysis and the first flowering plant to have its entire genome sequenced (The Arabidopsis Genome Initiative, 2000). It has a small genome size (125 Mbp), a low amount of repetitive DNA and a high gene density (Meinke et al., 1998). The ancestral lineages of *Arabidopsis* and *Brassica* diverged between 12.2 and 19.2 million years ago and the two species share extensive co-linearity and 87% sequence identity between orthologous exons (Love et al., 2005; Yang et al., 2006; The Arabidopsis Genome Initiative, 2000). The genome size of *B. rapa* is about four times (550 Mbp) the size of the Arabidopsis genome (Park et al., 2005). Due to the close phylogenetic relationship of *B. rapa* with *A. thaliana*, many opportunities emerged in transferring knowledge from the model to *B. rapa* for further improvement. Some breeders are now also working with less well known *Brassicaceae* species, such as *Brassica carinata*, as a source of valuable genes for resistance to pathogens, pests and other characteristics of economic importance (Dixon, 2007).

Genetics and genetic linkage map

Although *B. rapa* and *A. thaliana* share considerable DNA identity, the direct transfer of knowledge on gene function based on DNA sequence is not as straightforward as previously anticipated. This is largely due to extensive chromosomal duplications and subsequent deletions in these duplicated regions. The genomes of many flowering plants have undergone one or more rounds of duplication. Analysis of the genome sequence of *A. thaliana* shows large scale gene duplications (Blanc et al., 2003). Comparative genetic mapping often revealed six copies of each *A. thaliana* gene in *Brassica* species (Lagercrantz 1998; Parkin et al., 2003). Genomes of the *Brassica* species are composed of three rearranged variants of an

ancestral genome, basically similar to that of Arabidopsis, and the tribe *Brassicaceae* probably descended from a hexaploid ancestor (known as the triplication theory). Chromosome rearrangements, including fusions and/or fissions, subsequently resulted in diploid *Brassica* species with variation in chromosome number. Still, syntenic regions corresponding to Arabidopsis chromosome segments could be identified within the allopolyploid genome of *B. napus*, each in triplicate (Parkin et al., 2003). The genome triplication has led to an approximately 1.7-fold increase in the *B. rapa* gene number (49,000 - 63,000) compared to that of Arabidopsis (Hong et al., 2008; Yang et al., 2006). Chinese cabbage is considered a typical representative of the *Brassica* A genome. Due to the small genome size relative to the other *Brassica* species, together with the extensive genetics and genomics resources available, it has been adopted as the subject for the first whole genome sequence arising from the Multinational *Brassica* Genome Project (MBGP; http://www.brassica.info/). For its whole-genome sequencing, the *B. rapa* Genome Sequencing Project (*Br*GSP) consortium has developed suitable genomic resources and constructed genetic and physical maps. The Chinese cabbage (*B. rapa* ssp. *pekinensis*) inbred line Chiifu-401- 42 has been selected as the sequencing template (Choi et al., 2007).
Identifying molecular markers closely linked to a target gene is one of the most important steps of map-based cloning. Existence of a high-quality map will be of great importance in the process of map-based cloning. The development of genetic maps will be helpful for applied genetics and breeding of *Brassica* crops. A range of DNA marker types, including Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNA (RAPD), Simple Sequence Repeats (SSRs) and Amplified Fragment Length Polymorphisms (AFLPs), have been produced for *B. rapa*. The simplest and easiest populations to develop are F2 and back cross (BC) populations. Unfortunately they are difficult to maintain in this state unless by vegetative propagation. Recombinant inbred lines (RILs) and doubled haploid (DH) populations are easy to maintain, as these homozygous lines can simply be propagated sexually after self-fertilization. They are therefore considered to be immortal and very useful for repeated genetic analysis. There are more than 20 independent maps available for *B. rapa*, mainly made for F2 populations (12), but also for a few RIL (4) and DH (4) populations, and mostly using Chinese cabbage as one of the parents (Song et al., 1991; Teutonico and Osborn 1994; Nozaki et al., 1997; Kole et al., 2002; Kim et al., 2006; Suwabe et al., 2006; Soengas et al., 2007). The first genetic map of *B. rapa* was developed based on an F2 population of 95 individuals derived from the cross between a Chinese cabbage cultivar and an accession of spring broccoli (Song et al., 1991). The first recombinant inbred population was constructed as a set of 87 RI lines (F6), developed by single plant descent from F2 plants of a cross between *B. rapa* cultivars Per (a biennial winter turnip rape) and R500 (an annual spring yellow sarson) (Kole et al., 1997). The reference genetic linkage map of *B. rapa* was

constructed by Choi et al. (2007), for the Multinational *B. rapa* Genome Sequencing Project. Seventy-eight doubled haploid lines derived from anther culture of the F1 of a cross between two Chinese cabbage inbred lines, 'Chiifu-401-42' (C) and 'Kenshin-402-43' (K) were used to construct the map. The map comprises a total of 556 markers, including 278 AFLP, 235 SSR, 25 RAPD and 18 Expressed Sequence Tag Polymorphism (ESTP), Sequence Tagged Sites (STS) and Cleaved Amplified Polymorphic Sequence (CAPS) markers. The total length of the linkage map is 1,182 cM with an average interval of 2.83 cM between adjacent loci.

QTL mapping

Natural variation is caused by spontaneously arising mutations that have been maintained in nature by evolutionary processes such as artificial and natural selection. Thus, natural variation embraces the enormous diversity present within wild plant species as well as most of the genetic variants that are found in domesticated plants (Alonso-Blanco et al., 2009). Some of the phenotypic differences existing in wild or cultivated plants are due to single-gene (monogenic) allelic variants and cause distinct non-quantitative phenotypes. However, most of the natural variation is quantitative and determined by molecular polymorphisms at multiple loci and genes (multigenic). These are referred to as quantitative trait loci (QTLs) and quantitative trait genes (QTGs). The natural variation present in crop plants has been exploited since their domestication thousands of years ago by the genetic manipulation of developmental traits and physiological features related to adaptation to agriculture. Methods for genetic analysis and mapping of natural quantitative variation were developed a few decades ago for crop species, in which many more studies have been performed than in wild plants. Currently, genomic resources have been developed for important crop plants such as rice (*Oryza sativa*), maize (*Zea mays*), barley (*Hordeum vulgare*), and tomato (*Solanum lycopersicum*), enabling the identification of the genes and nucleotide polymorphisms underlying QTLs involved in domestication, yield, biotic and abiotic stress, and quality traits. QTL analysis in *B. rapa* has been used to identify loci related to a wide range of developmental and morphological traits including heat resistance, disease resistance, linolenic acid content, and flowering time. Genes controlling simply inherited traits, like seed colour, seed erucic acid content and the presence of leaf hairs have been mapped in an F3 population (Teutonico and Osborn, 1994). QTL mapping of leaf aliphatic glucosinolate loci has been carried out in two double haploid (DH) population of *B. rapa* and 16 loci controlling aliphatic glucosinolate accumulation has been identified (Lou et al., 2008).

Morphological traits related to leaf and stem along with flowering characteristics (days to bud formation, days to flower and days from bud formation to flower) have been studied by QTL analysis in an F2 population of *B. rapa* (Song et al., 1995). *Brassica* cultivars are classified as biennial or annual based on their requirement for a period of cold treatment (vernalization) to

induce flowering. Genes controlling the vernalization requirement were identified in a *B. rapa* F2 population derived from a cross between an annual and a biennial oilseed cultivar. Two major (*VFR1, 2*) and one minor (*VFR3*) QTLs for vernalization-responsive flowering time were detected (Teutonico and Osborn, 1995). The two genomic regions containing *VFR1* and *VFR2* showed homology to two regions in *B. napus*, which contain QTLs (*VFN1* and *VFN2*) controlling vernalization responsive flowering time variation in segregating populations derived from annual and biennial oilseed cultivars. The *Brassica* regions containing *VFR2* and *VFN2* were also found to be homologous to a region at the top of chromosome 5 of *A. thaliana*, where several flowering-time genes are located (Osborn et al., 1997), including *FLC*. Diploid *Brassica* species contain three copies of this genomic region. Thus, multiple copies of a gene homologous to a flowering-time gene on At5, such as *FLC*, could contribute to the wide range of variation in flowering time observed in Brassica species (Schranz et al., 2002).

Mutagenesis and Transformation

Transgenic technology provides the means for identifying and isolating genes controlling specific characteristics in one kind of organism, and for moving copies of those genes into another quite different organism, which will then also have those characteristics. This powerful tool enables plant breeders to generate more useful and productive crop varieties containing new combinations of genes and it expands the possibilities beyond the limitations imposed by traditional cross-pollination and selection techniques. Transgenic plants are proving to be powerful tools to study various aspects of plant sciences. They can be used for the functional analysis of plant genes e.g. in model plants and then making a link to their utilization in transgenic crops. For *Brassica* oilseed improvement, total seed oil yield and qualitative differences in oil composition are important breeding aspects. Transgenic *Brassica* with altered seed oil composition has been obtained. In addition, herbicide-, pathogen- and insect-tolerant transgenic *Brassica* cultivars have already been produced (Lim et al., 1998; Moon et al., 2007).

The *A. tumefaciens* DNA delivery system is still the most commonly used strategy to transform plants. This is for both historical reasons (it was the first available DNA delivery system) and a variety of advantages compared to direct DNA transformation methods. The *A. tumefaciens* transformation system is simple, in many cases efficient, and inexpensive (Block, 1993). In transgenic plants, insertional mutagenesis using heterologous maize transposons or *Agrobacterium* mediated T-DNA insertions, has offered valuable tools for the identification and isolation of genes based on a mutant phenotype (Gibson and Somerville, 1993). The gene functions discovered by insertional mutagenesis and silencing strategies along with expression pattern analysis will provide an integrated functional genomics perspective and

will offer unique applications in transgenic crops. More than forty genes have been identified by T-DNA tagging in *Arabidopsis* by 1997 (Azpiroz-Leehan and Feldmann, 1997) and since then many more have been isolated. The strategy of activation tagging was employed for the first time in tobacco cell lines by Walden (Walden, 1994) to generate overexpression mutants and isolate genes involved in plant growth and development. Activation tagging generates a dominant gain-of-function mutation which allows a direct selection of a desired phenotype in primary transformants. With a strong constitutive enhancer such as that derived from the CaMV 35S promoter, gene expression can be increased above normal levels. In this way, also redundant genes might display an overexpression phenotype if their product is limiting or a change in the concentration of gene products creates an imbalance that is manifested as a phenotype. These phenotypes can either directly reveal the gene function or provide a clue to the pathway in which the gene is involved. A transposon-based activation tagging method using the *En/I* Maize transposon system was reported for *Arabidopsis* (Marsch-Martinez et al., 2002). This technique is far more superior than T-DNA based activation tagging because of the higher frequency of gene-activated phenotypes found in transposon-tagged populations compared to T-DNA based populations (Weigel et al., 2000), probably due to methylation and subsequent silencing of multicopy T-DNA insertions (Chalfun et al., 2003).

It would be very useful to have access to a high efficiency mutagenesis tool in *B. rapa* to do functional genomics studies. There is no report about insertional mutagenesis in *Brassica* so far. One of the reasons could be the low frequency of transformation in *Brassica*. *B. rapa* is known as one of the most recalcitrant members of *Brassica* genus to regenerate shoots in vitro (Moon et al., 2007) and production of *B. rapa* transgenic plants is less advanced than others like *B. oleracea* and *B. napus* (Liu et al., 1998; Kuvshinov et al., 1999; Cho et al., 2000). Nonetheless, several *B. rapa* crop types have been transformed via *Agrobacterium*- mediated methods (Liu et al., 1998; Kuvshinov et al., 1999; Cho et al., 2000), but still there is no routine protocol for *A. tumefaciens*- mediated transformation of *B. rapa*. The transformation efficiency may be influenced by several factors, including genotype, explant type, donor plant age, concentration of growth regulators, *Agrobacterium* type and culture parameters (Earle et al., 1996).

Agrobacterium rhizogenes mediated transformation could be an alternative to get transgenic *B. rapa* plants. *A. rhizogenes* is a soil bacterium responsible for the development of the hairy root disease on a range of dicotyledonous plants. This phenotype is caused by genetic transformation in a manner similar to the development of crown gall disease by *A. tumefaciens*. Infection of wound sites by *A. rhizogenes* is followed by the transfer, integration, and expression of T-DNA from the root-inducing (Ri) plasmid and subsequent development of the hairy root phenotype (Christey and Braun, 2005). Ti and Ri plasmids both contain genes encode enzymes for producing organic compounds termed opines, which are produced

by infected plant cells and delivered to the pathogen as nutrients. In addition, Ti plasmids contain three T-DNA genes that direct synthesis of the plant hormones, cytokinin and auxin. In Ri plasmids the *rol* genes (root loci) *A* and *B,* play primary roles in adventitious root induction. *A. rhizogenes*-mediated transformation is a useful, easy and fast technique for introducing interested genes into plant cells, especially for species recalcitrant to transformation this technique is a valuable tool. When tested on *Brassicaceae*, within five to seven days after inoculation with *A. rhizogenes*, the vascular bundles of storage root disks of turnip or radish developed small outgrowths with numerous root hairs (Tanaka et al., 1985; Tepfer 1990). Hairy roots can be induced on a wide range of plants and many can be regenerated into plants, often spontaneously (Christey and Sinclaira 1992). Although the presence of the *rol* genes usually results in an altered phenotype, several studies have shown segregation of Ri and tumor-inducing (Ti) T-DNA, meaning that insertion of Ri and vector T-DNA may occur on different chromosomes during transformation, and segregation in the subsequent generation allows the recovery of transgenic plants with a normal phenotype (Puddephat et al., 2001; Christey and Braun, 2005).

There is variation in Agrobacterium-mediated transformation responses between different *Brassicaceae* species (Poulsen, 1996). Research in Arabidopsis has shown heritable variation for *A. tumefaciens* binding to the plant cell or for T-DNA integration. Further analysis of the plant genes involved, has made use of T-DNA-tagged *A. thaliana* mutant lines that are resistant to transformation by *A. tumefaciens* (*rat* mutants) (Nam et al., 1999). In *B. oleracea*, Cogan et al. (2002) identified QTLs for transgenic and adventitious root production using an *A. rhizogenes*-mediated co-transformation system in conjunction with a doubled haploid (DH) mapping population. The transfer and integration of *Agrobacterium* T-DNA into the plant genome is mediated through expression of the *vir* genes present on the virulence plasmid native to *Agrobacterium* strains. The virulence genes are highly conserved, and the key events in T-DNA transfer are common to the two principal species of *Agrobacterium* used in plant transformation. Consequently, the analysis of plant genes regulating transformation using either *A. rhizogenes* or *A. tumefaciens* will allow an understanding of a process common to both.

Scope of the thesis

At the time when this research started the main aim was to develop an *En/I* based activation tagging population in *B. rapa* to provide possibilities for studying plant gene functions. The first question that had to be answered was, would the *En/I* transposon system be active in *B. rapa*? This question is dealt with in **chapter 2**. Because *B. rapa* is known as one of the most recalcitrant members of *Brassica* genus to regenerate shoots in vitro, and also successful transformation of *B. rapa* by *A. tumefaciens* has been hampered by the lack of efficient plant transformation, it was tried to quantify transformation frequency by hairy root transformation using an *A. rhizogenes* transformation system in a segregating population. An F2 population of *B. rapa* derived from a cross between genotype L58, which is rapid–flowering, self–compatible Cai-xin line, and R-o-18, a doubled haploid spring oil genotype was developed. The construction of a genetic linkage map and the QTL analysis for some morphological traits will be discussed in **chapter 3**. The result of QTL analysis showed that the F2 population displays variation for many traits. By making a RIL population it was tried to fix the recombination events as much as possible to obtain an "immortal" mapping population. Genotyping the recombinant inbred population (F7) and QTL analysis of some morphological traits is the topic of **chapter 4**. In **chapter 5** the same population is used for the QTL analysis of seed metabolites, focusing on putative health-beneficial secondary metabolites. Finally, in **chapter 6** all findings of this research are discussed.

References

Abdel-Farid I.B., Kim H.K., Choi Y.H. Verpoorte R. (2007). Metabolic characterization of *Brassica rapa* leaves by NMR spectroscopy. J Agric Food Chem**.** 55:7936-7943.

Alonso-Blanco C., Aarts M.G.M., Bentsink L.,Keurentjes J.J.B., Reymond M., Vreugdenhil D., Koornneef M. (2009). What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell. 21:1877-1896

Azpiroz-Leehan R., Feldmann K.A. (1997). T-DNA insertion mutagenesis in Arabidopsis: going back and forth. Trends Genet. 13:152-156.

Blanc G., Hokamp K., Wolfe K.H. (2003). A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome. Genome Res. 13:137-144.

Block, M. D. (1993). The cell biology of plant transformation: Current state, problems, prospects and the implications for the plant breeding. Euphytica. 71:1-14.

Chalfun-Juniora A., Mes J.J., Mlynárováa L., Aarts M.G.M., Angenent G.C. (2003). Low frequency of T-DNA based activation tagging in Arabidopsis is correlated with methylation of CaMV 35S enhancer sequences. FEBS Letters. 555:459-463.

Cho H.S., Cao J., Ren J.P., Earle E.D. (2000). Control of Lepidopteran insect in transgenic Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) transformed with a synthetic *Bacillus thuringiensis cry1C* gene. Plant Cell Rep. 20:1-7.

Choi S., Teakle G., Plaha P., Kim J., Allender C., Beynon E., Piao Z., Soengas P., Han T., King G., Barker G., Hand P., Lydiate D., Batley J., Edwards D., Koo D., Bang J., Park B.S., Lim Y. (2007). The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. Theor Appl Genet. 115:777-792.

Christey M.C., Benjamin K.S. (1992). Regeneration of transgenic kale (*Brassica oleracea* var. *acephala*), rape (*B. napus*) and turnip (*B. campestris* var. *rapifera*) plants via *Agrobacterium rhizogenes* mediated transformation. Plant Science. 87:161-169.

Christey M.C., Braun R.H. (2005). Production of hairy root cultures and transgenic plants by *Agrobacterium rhizogenes*-mediated transformation. Methods Mol Biol. 286:47-60.

Cogan N.O.I., Lynn J.R., King G.J., Kearsey M.J., Newbury H.J., Puddephat I.J. (2002). Identification of genetic factors controlling the efficiency of *Agrobacterium rhizogenes*-mediated transformation in *Brassica oleracea* by QTL analysis. Theor Appl Genet. 105:568-576.

Dixon G.R. (2007). Vegetable Brassicas and Related Crucifers. Cambridge, Mass. CABI. pp 1-339.

Downey R.K., Stringam G.R., McGregor D.I., Stefansson B.R. (1975). Breeding rapeseed and mustard crops. In: Harapiak JT (ed) Oilseed and Pulse Crops in Western Canada. A Symposium, Western Coop Fertilizers. pp157–183.

Earle E.D., Metz T.D., Roush R.T., Shelton A.M. (1996). Advances in transformation technology for vegetable Brassica. Acta Hort. 407:161-168.

Gibson S., Somerville C. (1993). Isolating plant genes. Trends Biotechnol. 11:306-313.

Goffman F.D., Becker H.C. (2001). Diallel analysis for tocopherol contents in seeds of rapeseed. Crop Sci. 41:1072-1079.

Gómez-Campo C. (1999). Biology of *Brassica* coenospecies. Elsevier Science, The Netherlands. pp1- 81.

Hirai M., Harada T., Kubo N., Tsukada M. Suwabe K., Matsumoto S. (2004). A novel locus for clubroot resistance in *Brassica rapa* and its linkage markers. Theor Appl Genet. 108:639-643.

Hirai M., Matsumoto S. (2007). *Brassica rapa*. In C. Kole (ed). Genome Mapping and Molecular Breeding in Plants. Springer-Verlag Berlin, Heidelberg. 5:pp185-190.

Hong C.P., Kwon S.J., Kim J.S., Yang T.J., Park B.S., Lim Y.P. (2008). Progress in understanding and sequencing the genome of *Brassica rapa*. Int J Plant Genomics. 2008:582837.

Kim J.S., Chung T.Y., King G.J., Jin M., Yang T.J., Jin Y.M., Kim H.I., Park B.S. (2006). A sequence-tagged linkage map of *Brassica rapa*. Genetics. 174:29-39.

Kole C., Williams P.H., Rimmer SR., Osborn T.C. (2002). Linkage mapping of genes controlling resistance to white rust (*Albugo candida*) in *Bassica rapa* (syn. *campestris*) and comparative mapping to *Brassica napus* and *Arabidopsis thaliana*. Genome. 45:22-27.

Kole P.K., Vogelzang R., Osbom T.C. (1997). Genetic linkage map of a *Brassica rapa* recombinant inbred population. Heredity. 88:553-557.

Kuvshinov V., Koivu K., Kanerva A., Pehu E. (1999). *Agrobacterium tumefaciens*–mediated transformation of greenhouse *Brassica rapa* ssp. *oleifera*. Plant Cell Rep. 18:773–777.

Lagercrantz U. (1998). Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that Brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics. 150:1217–1228.

Li G., Gao M., Yang B., Quiros C.F. (2003). Gene for gene alignment between the Brassica and Arabidopsis genomes by direct transcriptome mapping. Theor Appl Genet. 107:168-180.

Lim H.T., You Y.S., Park E.J., Song Y.N., Park H.K. (1998). High plant regeneration, genetic stability of regenerants, and genetic transformation of herbicide resistance gene (BAR) in chinese cabbage (*Brassica campestris* ssp. *pekinensis*). ISHS Acta Horticulturae. 459:199-208**.**

Liu F., Cao M.Q., Yao L., Li Y., Robaglia C., Tourneur C. (1998). In planta transformation of pakchoi (*Brassica campestris* L. ssp. *chinensis*) by infiltration of adult plants with Agrobacterium. Acta Hortic. 467:187–192.

Lou P., Zhao J., He H., Hanhart C., Del Carpio D.P., Verkerk R., Custers J., Koornneef M., Bonnema G. (2008). Quantitative trait loci for glucosinolate accumulation in *Brassica rapa* leaves.New Phytol. 179:1017-32.

Lou P., Zhao J., Kim J.S., Shen S., Pino Del Carpio D., 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.

Love C.G., Robinson A.J., Lim G.A., Hopkins C.J., Batley J., Barker G., Spangenberg G.C., Edwards D. (2005). Brassica ASTRA: an integrated database for Brassica genomic research. Nucleic Acids Res. 33:D656-659.

Marsch-Martinez N., Greco R., Van Arkel G., Herrera-Estrella L., Pereira A. (2002). Activation tagging using the *En-I* maize transposon system in Arabidopsis. Plant Physiol. 129:1544-1556.

Meinke D.W., Cherry J.M., Dean C., Rounsley S.D., Koornneef M. (1998). *Arabidopsis thaliana*: a model plant for genome analysis. Science. 282:662-682.

Moon H.S., Halfhill M.D., Good L.L., Raymer P.L., Neal Stewart C.Jr. (2007). Characterization of directly transformed weedy *Brassica rapa* and introgressed *B. rapa* with Bt cry1Ac and gfp genes. Plant Cell Rep. 26:1001-10.

Nam J., Mysori K.S., Zheng C., Knue M.K., Matthysse A.G., Gelvin S.B. (1999). Identification of T-DNA tagged Arabidopsis mutants that are resistant to transformation by Agrobacterium. Mol Gen Genet. 261:429-438.

Nozaki T., Kumazaki A., Koba T., Ishikawa K., Ikehashi H. (1997). Linkage analysis among loci for RAPDs, isozymes and some agronomic traits in *Brassica campestris* L. Euphytica. 95:115–123.

Olsson G. (1960). Species crosses with the genus *Brassica,* artificial *Brassica juncea Coss.* Hereditas. 46:173–223.

Osborn T.C., Kole C., Parkin I.A.P., Sharpe A.G., Kuiper M. Lydiate D.J., Trick M. (1997). Comparison of flowering time genes in *Brassica rapa*, *B. napus* and *Arabidopsis thaliana*. Genetics. 146:1123-1129.

Park J.Y., Koo D.H., Hong C.P., Lee S.J., Jeon J.W., Lee S.H., Yun P.Y., Park B.S., Kim H.R., Bang J.W., Plaha P., Bancroft I., Lim Y.P. (2005). Physical mapping and microsynteny of *Brassica rapa* ssp. *pekinensis* genome corresponding to a 222 kbp gene-rich region of Arabidopsis chromosome 4 and partially duplicated on chromosome 5. Mol Genet Genomics. 274:579-588.

Parkin I.A.P., Sharpe A.G., Lydiate D.J. (2003). Patterns of genome duplication within the *Brassica napus* genome. Genome. 46:291-303.

Pereira A. (2000). A transgenic perspective on plant functional genomics. Transgenic Res. 9:245-260.

Poulsen G.B. (1996). Genetic transformation of Brassica. Plant Breeding. 115:209-225.

Prakash O. (1961). Food and drinks in ancient India. Munshi Ram Manohar Lal, Delhi. pp 265-266.

Prakash S., Hinata K. (1980). Taxonomy, cytogenetics and origin of crop Brassicas, a review. Opera. Bot. 55:3-57.

Puddephat I.J., Robinson H.T., Fenning T.M., Barbara D.J., Morton A., Pink D.A.C. (2001). Recovery of phenotypically normal transgenic plants of *Brassica oleracea* upon *Agrobacterium rhizogenes*-mediated co-transformation and selection of transformed hairy roots by GUS assay. Mol Breed. 7:229–242.

Quijada P., Xiaowu Wang J.C., Hirai M., Kole C. (2007). Oil seeds. In C. Kole (ed). Genome Mapping and Molecular Breeding in Plants, Springer Berlin Heidelberg. 2:211-263.

Rakow G. (2004). Species origin and economic importance of Brassica. In E.C. Pua and C.J.Douglas (ed.). Biotechnology in Agriculture and Forestry. Brassica. Springer-Verlag Berlin Heidelberg, New York. 54:3–11.

Rich T.C.G. (1991). Crucifers of Great Britain and Ireland. Botanical Society of the British Isles, London, pp 336.

Scarth R., Tang J. (2006). Modification of Brassica oil using conventional and transgenic approaches. Crop Sci. 46:1225-1236.

Schranz M.E., Quijada P., Sung S-B., Lukens L., Amasino R.,Osborn T.C. (2002). Characterization and effects of the replicated flowering time gene *FLC* in *Brassica rapa*. Genetics. 162:1457-1468.

Soengas P., Hand P., Vicente J.G., Pole J. M., Pink D.A. (2007). Identification of quantitative trait loci for resistance to *Xanthomonas campestris* pv. *campestris* in *Brassica rapa.* Theor Appl Genet. 114:637-45.

Song K., Slocum M.K., Osborn T.C. (1995). Molecular marker analysis of genes controlling morphological variation in *Brassica rapa* (syn. *campestris*). Theor Appl Genet. 90:1-10.

Song K.M., Suzuki J.Y., Slocum M.K., Williams P.M., Osborn T.C. (1991). A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci. Theor Appl Genet. 82:296-304.

Stringam G.R., McGregor D.I., Pawlowski S.H. (1974). Chemical and morphological characteristics associated with seed coat color in rapeseed. In: Proc 4th Int Rapeseed Conf. pp:99–108.

Suwabe K., Tsukazaki H., Iketani H., Hatakeyama K., Kondo M., Fujimura M., Nunome T., Fukuoka H., Hirai M., Matsumoto S. (2006). SSR-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. Genetics. 173:309-319.

Tanaka N., Hayakawa M., Mano Y., Ohkawa H., Matsui C. (1985). Infection of turnip and radish storage roots with *Agrobacterium rhizogenes*. Plant Cell Rep. 4:74-77.

Tanhuanpää P., Schulman A. (2002). Mapping of genes affecting linolenic acid content in *Brassica rapa* ssp. *Oleifera.* Mol Breed. 10:51-62.

Tepfer D. (1990). Genetic transformation using *Agrobacterium rhizogenes*. Physiol. Plant. 79:140-146.

Teutonico R.A., Osborn T.C. (1994). Mapping of RFLP and qualitative trait loci in *Brassica rapa* and

comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*. Theor Appl Genet. 89:885–894.

Teutonico R.A., Osborn T.C. (1995). Mapping loci controlling vernalization requirement in *Brassica rapa*. Theor Appl Genet. 91:1279-1283.

Toxopeus H., Oost E.H., Reuling G. (1984). Current aspects of the taxonomy of cultivated Brassica species. The use of *B. rapa* L. versus *B. campestris* L. and a proposal for a new intraspecific classification of *B. rapa* L. Crucifer Newsletter. 9:55–57.

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

Walden R., Fritze K., Hayashi H., Miklashevichs E., Harling H., Schell J.(1994). Activation tagging: a means of isolating genes implicated as playing a role in plant growth and development. Plant Mol Bio. 26:1521-1528.

Warwick S.I., Black L.D. (1991). Molecular systematics of Brassica and allied genera (subtribe Brassicinae, Brassiceae) - chloroplast genome and cytodeme congruence. Theor Appl Genet. 82:81-92.

Weigel D., Ahn J.H., Blázquez A.M., Borevitz J.O., Christensen S.K., Fankhauser C., Ferrándiz C., Kardailsky I., Malancharuvil E.J., Neff M.M., Nguyen J.T., Sato S., Wang Z.Y., Xia Y., Dixon R.A., Harrison M.J., Lamb C.J., Yanofsky M.F., Chory J. (2000). Activation tagging in Arabidopsis. Plant Physiol. 122:1003-1013.

Yang, T.J., Kim J.S., Kwon S.J., Lim K.B., Choi B.S., Kim J.A., Jin M., Park J.Y., Lim M.H., Kim H. I., Lim Y.P., Kang J.J., Hong J.H., Kim C.B., Bhak J., Bancroft I., Park B.S. (2006). Sequence-level analysis of the diploidization process in the triplicated FLOWERING LOCUS C region of *Brassica rapa*. Plant Cell. 18:1339-1347.

Yu S.C., Wang Y.J., Zheng X.Y. (2003). Mapping and analysis of QTL controlling some morphological traits in Chinese Cabbage (*Brassica campestris* L. ssp. *pekinensis*). Acta Genet Sin. 30:1153-1160.

Zhao J., Wang X., Deng B., Lou P., Wu J., Sun R., Xu Z., Vromans J., Koornneef M., Bonnema G. (2005). Genetic relationships within *Brassica rapa* as inferred from AFLP fingerprints. Theor Appl Genet. 110:1301-1314.

Chapter 2

Agrobacterium-mediated transformation of a transposon construct into *Brassica rapa*

Hedayat Bagheri, Corrie Hanhart, Aline Le Bihan, Andy Pereira and Mark G.M. Aarts

Abstract

A high-efficiency mutagenesis system would be very advantageous in *Brassica rapa* to generate mutations, isolate plant genes and gain insights into its biology. Activation of a heterologous *En/I* transposon was verified by the analysis of transgenic tissue of *B. rapa* Yellow Sarson L143 containing an *En/I* activation tagging construct. The *En-I* activation tagging construct contained the *En (Spm)* transposase coding sequence under control of the CaMV 35S promoter and terminator; a mobile non-autonomous *I (dSpm)* element harboring a tetramer of the CaMV 35S enhancer to act as activator; the *NPTII* gene conferring kanamycin resistance as positive selection marker; and the *SU1* gene that converts the pro-herbicide R7402 into the herbicide sulfonylurea inhibiting or reducing the growth of plants that contain it, as negative selection marker in plants. The construct was introduced into *B. rapa* L143 by *Agrobacterium rhizogenes* hairy-root-mediated transformation. Insertion of the transposon construct and excision of the *I* element from the original construct was confirmed by PCR. Sequencing of cloned PCR products further confirmed the presence of an empty donor site in the donor T-DNA. Transgenic plants of *B. rapa* were obtained by inoculating hypocotyls with *Agrobacterium tumefaciens* AGL0 carrying the binary vector pCAMBIA-2301, harboring a β-glucuronidase (GUS) gene driven by the CaMV 35S promoter. A histochemical GUS assay showed a transformation frequency of 2%. To establish a stable transposon mutant population of *B. rapa, A. tumefaciens* mediated transformation of the *En/I* transposon activation tagging construct was employed. Unfortunately the transformation of *B. rapa* cotyledons and hypocotyls with *A. tumefaciens* AGL1 carrying the the *En/I* transposon activation tagging construct was not successful. This could be either due to the length of the construct, or the expression of the *SU1* gene which may not be viable in green *B. rapa* tissue.

Introduction

Brassica rapa is a crop which is receiving increasingly more interest for functional genomics studies, in view of its economic importance and the genomic advantages as a close relative of the plant reference species *Arabidopsis thaliana*, with genome sequencing in progress (http://www.brassica.info/resource/sequencing.php). While QTL-analysis and genome or EST (expressed sequence tags) sequencing projects are being set up, mutant populations are still scarce. The availability of a mutant resource would be helping researchers in their quest to gain insights into the biology of this commercially important crop. These efforts are critical to understand gene function and, ultimately, the biology of *B. rapa*.

Insertional mutagenesis through T-DNA or transposon tagging, are valuable tools for isolating plant genes. Snapdragon and maize contain well-characterized endogenous transposable elements, which have been used in the past to isolate genes such as *viviparous-1* (involved in seed development), *opaque-2* (encoding a transcriptional regulator), and several genes involved in anthocyanin production and flower development (McLaughlin and Walbot 1987; Gibson and Somerville 1993; Azpiroz-Leehan and Feldmann 1997). Subsequently, these transposons have been used to design heterologous systems for use in plants for which endogenous transposons have not been characterized in sufficient detail. These systems proved to be successful, leading to the isolation of many genes, including the *MALE STERILITY 2* gene of *Arabidopsis* (Aarts et al., 1993), Tomato *dwarf D* (Bishop et al., 1996), and the disease resistance genes *N* from tobacco (Whitham et al., 1996), *L6* from flax (Lawrence 1995), and *Cf-9* from tomato (Jones et al., 1994).

The strategy of activation tagging was employed for the first time by Walden et al. (1994) in tobacco cell lines to generate mutations and to isolate the corresponding genes involved in plant growth and development. Many activation tagging constructs have four copies of the enhancer element of the constitutively active cauliflower mosaic virus (CaMV) 35S promoter. These enhancers can cause transcriptional activation of nearby genes, and because activated genes will be associated with a T-DNA insertion, this approach has become known as activation tagging. Activation tagging generates a dominant gain-of-function mutation which allows a direct selection of a desired phenotype in primary transformants. Later on it was used in Arabidopsis by Weigel and co-workers (Weigel et al., 2000) for the isolation of over 30 dominant mutants with various phenotypes, using resistance to the antibiotic kanamycin or the herbicide glufosinate to select for plants with the activation construct. The frequency with which activation tagging produces visible, dominant activation tagged mutants ranged from 0.07% (Weigel et al., 2000) to 2.2% in *Arabidopsis* (Ichikawa et al., 2003). A T-DNA-based activation system has the disadvantage that it creates a complex integration pattern and chromosomal rearrangement at the insertion site (Marsch-Martinez et al., 2002). A transposon-based activation tagging method using the maize *En/I* transposon was reported for

Arabidopsis (Marsch-Martinez et al., 2002). This technique is far more superior than T-DNA based activation tagging because of the higher frequency of gene-activation phenotypes found in transposon-tagged populations compared to T-DNA based populations (Weigel et al., 2000), probably due to methylation and subsequent silencing of multicopy T-DNA insertions (Chalfun-Junior et al., 2003).

Ac-Ds (*Activator-Dissociation*) and *En-I* (*Enhancer-Inhibitor*) are two systems of maize transposable elements which have been used in transposon tagging. In *Arabidopsis*, the *En-I* system showed a high frequency of independent transpositions of 7.8% to 29.2% (Aarts et al., 1995; Speulman et al., 1999). Transposon tagging is an attractive system to use for transformation-recalcitrant plant species, because just a few primary transformants are sufficient to generate a tagged population (Marsch-Martinez et al., 2002). The *En-I* transposon activation tagging construct, which was used by Marsch-Martinez et al. (2002), carries the *BAR* gene, conferring resistance to the herbicide Basta, and the *SU1* gene from *Streptomyces griseolus* that confers sensitivity to the proherbicide R7402 The later acts as a negative selectable marker to select against the presence of the T-DNA upon transposition of the *I*-element carrying the activation tag. Thus, stable transposants can be selected by applying Basta and R7042 to soil-grown seedlings. For *B. rapa*, such a high-efficiency mutagenesis system will be very advantageous. A collection of many independent transposants can be generated and screened for morphological mutants and mutants altered in e.g. abiotic stress response (drought, salt, mineral deficiency, mineral access) or phytonutrient content. Verified mutants can be used to identify and clone the tagged gene and analyze its function.

Instead of first trying to generate many stable transformants for this transformation-recalcitrant species, a swift prescreening can be performed based on *Agrobacterium rhizogenes*-mediated root transformants. *A. rhizogenes* is a soil bacterium responsible for the development of hairy root disease on a range of dicotyledonous plants. This phenotype is caused by genetic transformation in a manner similar as the development of crown gall disease by *A. tumefaciens*. *A. rhizogenes* is capable of transferring the T-DNA of the Ri (root inducing) plasmid to plants (Christey 2001). Hairy roots can be induced on a wide range of plants and many can be regenerated into plants, often spontaneously. Transgenic plants have been obtained after *A. rhizogenes*-mediated transformation in 89 different taxa, representing 79 species from 55 genera and 27 families, including *B. rapa* (Christey 2001). Hairy roots are easily distinguished by their rapid, highly branching growth on hormone-free medium and plagiotropic root development. The Ri plasmid carries *rol* (*root locus*) and *aux* genes that are transferred to the plant and that are responsible for the hairy root phenotype (Christey and Braun 2005). Plants regenerated from hairy roots often exhibit an altered phenotype characterized by several morphological changes including wrinkled leaves, shortened

internodes, reduced apical dominance, reduced fertility, altered flowering, and plagiotropic roots due to expression of *rol* genes. To be of use in crop improvement, *A. rhizogenes*-mediated co-transformation will have to yield transgenic plants that do not display the deleterious hairy-root phenotype (Puddephat et al., 2001). Where Ri and binary vector T-DNAs insert independently into the genome during co-transformation, meiotic segregation can lead to recovery of phenotypically normal plants bearing binary vector T-DNA only, as was shown by Puddephat et al. (2001) upon *A. rhizogenes*-mediated co-transformation of *B. oleracea*.

Ultimately, the production of stable transgenic plants through *A. tumefaciens* mediated transformation is desired, as it would not need the outcrossing of the Ri-T-DNA to obtain morphologically normal plants. The capability of a plantlet to regenerate from a transformed cell appears to be a critical factor affecting the overall transformation efficiency of *B. rapa*. In general, the regeneration from explants (i.e. isolated or excised tissues or organs from an intact plant) can occur via two major pathways, organogenesis and somatic embryogenesis (Zuo et al., 2002). The organogenesis pathway depends on a high cytokinin:auxin ratio and leads to the direct formation of shoots from the so-called 'organogenic' cells. The formation of somatic embryos, which can directly germinate into adult seedlings, requires a high concentration of 2, 4-D. Other important factors that may influence the infection frequency are the duration of inoculation, ethylene concentration, *Agrobacterium* co-cultivation temperature and co-cultivation time. Ethylene produced from a plant inoculated with *A. tumefaciens* inhibits *virulence* (*vir)* gene expression in *A. tumefaciens* and, as a result, *Agrobacterium*-mediated genetic transformation is inhibited (Nonaka et al., 2008). $AgNO_3$, an anti-ethylene agent, was critical to the regeneration of transformed shoots of *B. oleracea* and was thought to reduce wounding stress and stimulate regeneration from transformed cells (Puddephat et al., 2001). The vir genes in *Agrobacterium* can be also induced by acetosyringone or similar phenolic compounds (Takasaki et al., 1997). Tissue culture involves complex interactions and thus experiments to improve tissue culture rely to a great extent on empirical data. Many parameters have to be investigated such as media, genotype, variation among different explants and the amount of growth regulators applied and physiological condition of the starting material. With all mentioned factors interacting during the formation of transgenic plants, it is obvious that establishing optimal conditions are extremely difficult (Poulsen 1996). In general among the *Brassica* species, *B. oleracea* is the easiest species to regenerate from tissue culture, whereas *B. rapa* is the most recalcitrant one (Narasimhulu et al., 1988).

Here we present the development of a transposon-based activation tagging system in *B. rapa* using hairy root transformation and we report on various methods to obtain stable transformation of *B. rapa* by *A. tumefaciens*.

Material and Methods

Vector constructs and bacterial strains

Four different constructs introduced into *A. rhizogenes* MSU440 were used for this study. The pRedRoot binary vector (Limpens et al., 2004), expressing the *DsRed1* gene encoding the red fluorescent protein as a non-destructive selectable marker. The binary vector pCAMBIA-2301 harbouring the β-glucuronidase (*GUS*) gene and selectable kanamycin resistance gene (http://www.cambia.org/daisy/bios/585.html). An *Ac/Ds* activation tagging construct carrying the green fluorescent protein (*GFP*) gene as visible marker, and an *En/I* activation tagging construct carrying the kanamycin resistance gene as selectable marker (Fig 1). All constructs were used to transform *B. rapa.* The *En/I* construct (made by Marsch-Martinez et al., 2002) contains an *En*-transposase source, a nonautonomous transposable *I* element, which contains a tetramer of the CaMV 35S enhancer sequence, and the *NPTII* kanamycin resistance gene. The construct includes positive and negative greenhouse-selectable markers to select for stable transpositions. The *BAR* gene confers resistance to the herbicide Basta and *SU1* gene, which encodes a protein that converts the pro-herbicide R7402 (DuPont, Wilmington, DE) into the herbicide sulfonylurea inhibiting or reducing the growth of plants that contain it. The binary vectors were introduced into MSU440 by electrotransformation and grown for 2 d at 28 °C under antibiotic selection (50 µg ml^{-1}). Integrity of inverted repeat constructs was checked by mini-prepping and restriction-digestion.

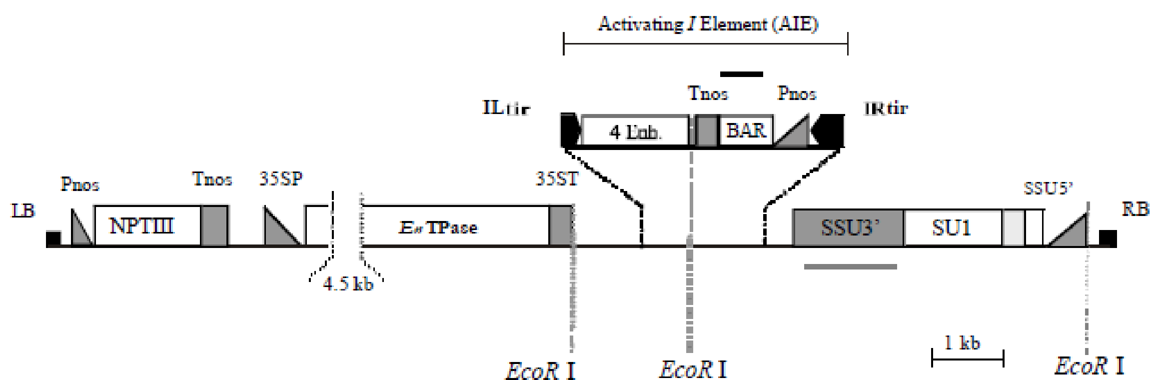


Fig 1 *Schematic representation of the En/I construct (En/I::4X35SEnh) transposon activation tagging construct. LB, Left border; RB, right border; 35SP, 35ST, CaMV 35S promoter and terminator, respectively; EnTPase, En immobile transposase source; ILtir, IRtir, I-element left and right terminal-inverted repeat,respectively; 4 Enh., tetramer of the CaMV 35S enhancer; BAR, Basta resistance gene; Pnos, Tnos, promoter and terminator sequences from the nopaline synthase gene, respectively; SSU5, SSU3, promoter and transit signal peptide to the chloroplast and terminator of the small subunit of Rubisco gene, respectively (reproduced from Marsch-Martinez et al., 2002).*

The *Agrobacterium tumefaciens* strains used in this study were AGL0 and AGL1 (Lazo et al., 1991). Both strains carry pTiBo542DT, a hypervirulent Ti helper plasmid that transforms many dicotyledonous plants at very high frequencies. If these strains are to function in transformation they must also carry a binary vector in which the T-DNA is defined by both left and right border repeats. AGL1 is distinguished from AGL0 by an insertion mutation in the *recA* recombination gene (Lazo et al., 1991), which suppresses recombination of large T-DNA plasmids. AGL0 harbors the binary vector pCAMBIA-2301 and is referred to as AGL0-GUS and the *En/I* transposon construct (AGL0-*En/I*). AGL1 harbors the *En/I* transposon activation tagging construct (AGL1-*En/I*) (Marsch-Martinez et al., 2002).

***Agrobacterium rhizogenes* hairy root transformation**

Brassica seeds were surface-sterilized by incubating for 30 seconds in 70% ethanol, washing in sterile water, then 2 min in 2% hypochlorite (commercial bleach), washing three times in sterile water, and subsequently plated on half MS medium (Murashige and Skoog 1962) without sucrose. Roots of 5-day-old seedling were cut with a schalpel and inoculated with a pellet of *A. rhizogenes* carrying pRedRoot-DsRed placed at the cutting surface as explained by Limpens (Limpens et al., 2004). Infected seedlings were put on the petri dishes having the same medium as germination one with a half-round filter paper (5 seedlings per plate). The Petri dishes were not completely closed by parafilm to enable aeration. Petrie dishes were placed vertically in a growth chamber maintained at 25˚C with a 16h light/8h dark photoperiod at a light intensity of 60 $mEm^{-2}s^{-1}$. Four- day-old seedlings transferred to the medium containing 200 mg/l ticarcillin. The same protocol was employed for *A. rhizogenes* carrying pCAMBIA-2301-GUS and inoculated roots were placed on kanamycin medium. Co-transformation also was done using both *A. rhizogenes* carrying pRedRoot-DsRed and *A. rhizogenes* carrying pCAMBIA-2301-GUS. Inoculated roots containing the *En-I* transposon activation tagging construct were transferred to kanamycin medium with tricarcillin and inoculated roots containing the *Ac/Ds*-GFP construct were transferred to medium with tricarcillin. On selection medium, the roottips were marked on the plate to distinguish transformed and growing roots from untransformed ones.

Plant material and plant growth media

Two varieties of *B. rapa* L58 (Caixin line) and L143 (Yellow Sarson) were used for transformation. Seeds were surface-sterilized for 20 min and sown in half-strength MS-salts (Murashige and Skoog 1962) and vitamins with 0.5% sucrose and 0.6% phyto agar. Seeds were germinated and grown in a cell culture room maintained at 22-24 °C with a 16 h light/8 h dark photoperiod at a light intensity of 45-55 $mEm^{-2}s^{-1}$. Cotyledon and hypocotyl explants

were used for transformation. For each of the different transformation experiments different protocols were used.

Experiment 1:

For callus induction, MS medium supplemented with 3% sucrose, 4 mg/l Kinetin, 2 mg/l NAA and 0.1 mg/l 2,4-D was used. For shoot regeneration MS medium with 1% sucrose, 2 mg/l NAA, 4mg/l BAP and 3mg/l ABA was used. All media were adjusted to pH 5.8 and solidified by 0.6% phyto agar.

Cotyledons and hypocotyls of 5-day-old seedlings were placed on callus induction medium for one week and then transferred to shoot regeneration medium. Every two weeks regeneration medium was refreshed. For plant transformation, *A. tumefaciens* strains (AGL0-GUS and AGL1-*En/I*) were cultivated on LB-agar plates supplemented with 50 mg/l kanamycin and 75 mg/l carbenicillin respectively for 2 days (d) at 27 °C, followed by cultivation on liquid LB medium for 2 d. Bacteria were collected by centrifugation at 3000 × *g* for 15 min, and the pellet was suspended in liquid MS medium to give OD_{600} = 0.5, which was subsequently used in the transformation experiments. For transformation first we tried AGL0-GUS transformation. 10 µl of the *Agrobacterium* suspension was applied on the explant cut surface in the callus induction medium for two days, upon which they were transferred to callus induction medium with tricarcillin (100 mg/l) to prevent overgrowth of *Agrobacterium*. After 5 days they were transferred to shoot regeneration medium with kanamycin (50 mg/l) and tricarcillin (100 mg/l). Six weeks later transformation and/or regeneration was scored. In total 300 explants were used in the first experiment, half with *Agrobacterium* (for transformation) and half without (for regeneration).

Experiment 2:

In this experiment the Radke protocol was followed (Radke et al., 1992). We used 1 mg/l 2,4-D in the induction medium and 3 mg/l BA+1 mg/l Zeatin in the shoot regeneration medium. $AgNO_3$ (5 mg/l) was applied to shoot regeneration medium to reduce the effect of ethylene. For transformation we used AGL0-GUS on 300 explants of both genotypes.

Experiment 3:

As experiment 2, but using 1000 explants, 500 for each genotype. Instead of ALG0-GUS, AGL1-*En/I* was used.

Experiment 4:

As experiment 3, but using 600 explants and using 0.5 mg/l 2, 4-D and 0.5 mg/l BA in the shoot regeneration medium.

Experiment 5:

As experiment 3, but replacing $AgNO_3$ with Aminoethoxyvinylglycine (AVG). We used 1 μM and 5 μM AVG in shoot regeneration medium. Five hundred explants were used for this experiment. *B. rapa* L58 and L143 and Agrobacterium AGL1-*En/I* were used for this experiment.

Experiment 6:

Experiment 3 was repeated but a liquid shoot regeneration medium was used, which was refreshed every two weeks and shaken at low speed for six weeks. *B. rapa* L58 and L143 and Agrobacterium AGL1-*En/I* were used for this experiment.

Experiment 7:

As experiment 3, but to induce direct shoot formation in organogenesis, we used a higher cytokinin:auxin ratio as suggested by Zuo et al., (2002). We applied 5 mg/l BA and 0.5 mg/l NAA in the shoot regeneration medium to resemble the medium Zhang used to get maximum shoot regeneration (Zhang et al., 1998). 400 explants were used.

Experiment 8:

As experiment 3, with alterations of the hormone concentrations in the shoot regeneration medium. Five concentrations of BA were applied (1, 2, 3, 4 and 5 mg/l) in combination with three concentrations of NAA (0.5, 1 and 3 mg/l) and two concentrations of GA3 (0 and 0.2). Fifty explants were used for each treatment.

Experiment 9:

As experiment 3, but seeds were germinated on medium containing 3 mg/l BAP and 4 mg/l NAA. 400 explants were used.

Experiment 10:

As experiment 3, but explants were pre-cultured on medium with 2 mg/l BA, 1 mg/l NAA and 0.2 mg/l GA3 for two days, then inoculated with AGL1-*En/I* and kept for another two days on pre-culture medium before transferring to shoot regeneration medium. 400 explants were used.

Experiment 11:

As experiment 3, but the cotyledon with petiole was used as explant. Petioles were inserted into the agar. 400 explants were used.

Experiment 12:
As experiment 3, but two concentrations, 0.5mg/l and 1mg/l, of thidiazuron (TDZ) were used in the medium. 300 explants were used.

Experiment 13:
As experiment 3, but 0.03 mg/l brassinolide was added to the medium. 300 explants were used.

Experiment 14:
As experiment 3, but acetosyringone (100 µM) was used to re-suspend the bacteria pellet for co-cultivation medium. 300 explants were used for this experiment.

Experiment 15:
Instead of *B. rapa*, tobacco explants (cultivar SR1) were used for transformation. We followed the protocol reported by Dorlhac et al. (1993), using AGL0-GUS, AGL0-*En/I* and AGL1-*En/I*. 10 explants of tobacco leaf were used for transformation with each construct.

Experiment 16:
As experiment 2 but instead of *B. rapa* the related species *B. napus* cv. Westar was used. We used AGL0-GUS, AGL0-*En/I* and AGL1-*En/I* in this experiment. About 60 explants used for AGL0-GUS transformation and more than 300 explants were used for the other two constructs.

Experiment 17:
In planta transformation of *B. rapa* was tried using genotypes L58 and L39. Two weeks cold treatment (4 °C) was applied on the seeds before they were germinated in a growth chamber at 24/20 °C day/night, 16/8 h day/night. After 3 weeks they started flowering. The protocol of *in planta* transformation of *A. thaliana* was followed (Bent 2000) and AGL1-*En/I* was used for dipping. 180 plants of L39 and 140 plants of L58 were used for this experiment. Seeds were collected after ripening of the plants and subjected to selection. As the construct contains genes for kanamycin and basta resistance, both selective agents were (separately) used for screening of transformed seeds. Half of the seeds were germinated on soil and two-leaf-staged seedlings were sprayed in the greenhouse with 1% Basta solution. The other half was screened on solidified MS medium containing 50 mg/l kanamycin.

Experiment 18:

Instead of *A. tumefaciens* transformation, transformation was performed with *A. rhizogenes* to obtain transgenic hairy roots of *B. rapa*. *A. rhizogenes* strain MSU440 was used containing the *En/I* construct. Cotyledon and hypocotyls of accession L58 and R-o-18 were used as explants. Inoculation was as with *A. tumefaciens*.

GUS staining

For histochemical GUS staining one leaflet was put in staining solution including X-Gluc (0.1 mg/ul), K_2HPO_4 (500 mM), KH_2PO_4 (500 mM) with water and was kept at 37°C in dark for overnight. The next day samples were destained in hypochlorite 1% for one hour.

Molecular analysis of excision

Insertion of the *En/I* and *Ac/Ds* activation tagging constructs were confirmed by PCR using BAR-specific primers (Bar F: 5'-ACC ATG AGC CCA GAA CGA CGC-3' and Bar R: 5'-CAG GCT GAA GTC CAG CTG CCA G-3') as the selectable marker is within the transposable element between the terminal-inverted repeats (Fig 1). The PCR reaction comprised a denaturing step of 95°C for 3 min, followed by 40 cycles of 1 min at 95°C, 1 min at 60°C, and 1 min at 72°C, ending with an elongation step of 5 min at 72°C. A subsequent PCR with primers flanking the jumping elements was performed to detect excision of *I* or *Ds* element from the T-DNA in the transformed roots (Fig 3). The following primers were employed for detection of *I* element excision and *Ds* element excision respectively:
35S-T2, 5'-CCA AAA TCC AGT GGG TAC CGA GC-3'; SSU-301-TF, 5'-GTT GGT TGA GAGTCTTGTGGCCT-3'; and AcUp, 5'-CTCAGTGGTTATGGATGGGAGTTG-3'; M13F24, 5'-CGCCAGGGTTTTCCCAGTCACGAC-3'. The PCR reaction comprised a denaturing step of 95°C for 3 min, followed by 35 cycles of 1 min at 95°C, 1 min at 60°C, and 1 min at 72°C, ending with an elongation step of 5 min at 72°C.

Results

A. rhizogenes transformation

One week after transformation transformed roots were visually selected using a fluorescence stereo microscope in case of using *A. rhizogenes* carrying DsRed and the *Ac/Ds* construct carrying 35S::GFP, and by selection of growing roots on antibiotic selection media in case of *A. rhizogenes* carrying *En/I* activation tagging constructs (Fig 2). The putatively transformed roots carrying the GUS construct were evaluated with the histochemical GUS assay.

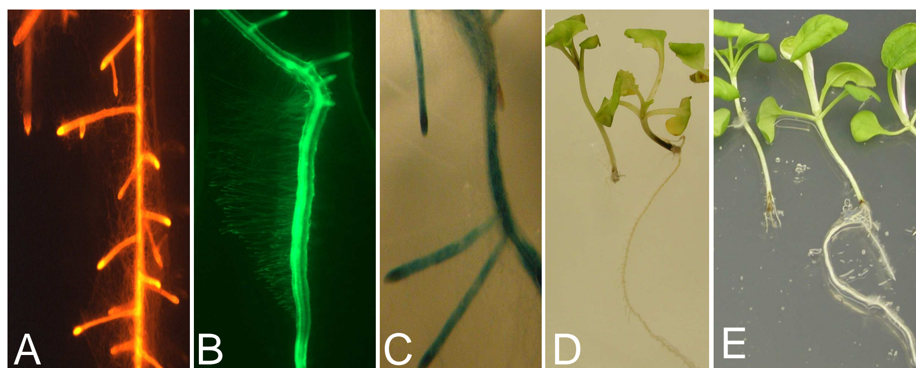


Fig 2 *Selection of roots by A. rhizogenes, transformed with pRedRoot, carrying DsRed (A), the Ac/Ds construct carrying GFP (B), pCAMBIA-2301 carrying GUS (C), En/I activation tagging construct carrying HPT (D), En/I activation tagging construct carrying NPTII (E). A: DsRed expression, B: GFP expression, C: GUS expression, D: Selection of transformed roots on 20 mg/l hygromycin medium (Left: untransformed, Right: transformed), E: Selection of transformed roots on 50 mg/l kanamycin medium (Left: untransformed, Right: transformed).*

Twenty seven plants used for transformation with *A. rhizogenes* carrying DsRed, of which nine plants (33%) showed root DsRed expression. The transformation frequency for transformation with *A. rhizogenes* carrying GUS was 55% (15/27). Roots emerging after co-transformation with both *A. rhizogenes* carrying DsRed and GUS constructs were first tested under a stereo fluorescence microscope. Ten out of fifty plants showed root DsRed expression and about four plants out of these ten plants showed root GUS expression. This does not suggest a higher transformation frequency upon co-infection.

Activation tagging

Twelve plants out of thirty (40%) showed root GFP expression (*Ac/Ds* construct) under the stereo fluorescence microscope. Of the roots transformed with *A. rhizogenes* carrying the *En/I* construct, 58% (26/45) were kanamycin resistant. All roots putatively transformed with transposon activation tagging constructs were cut and used for DNA analysis to test for excision events.

No excision site was detected in material containing the *Ac/Ds* transposon activation tagging construct, but one excision site was observed among 26 roots containing the *En/I* activation tagging construct (Fig 3). Two more excision sites were obtained when repeating the experiment with 50 additional plants of which 27 showed roots to contain the *En/I* construct. The DNA was extracted from the empty donor bands on the gel and cloned using the pGEMT-Easy vector. We could clone only one of the excision sites. Sequencing of the cloned fragment confirmed the presence of an empty donor site in the T-DNA and thus excision of the *I* element. The *I* element which was cloned into the *En/I* activation tagging construct carried its left terminal inverted repeat (TIR) originating from the wx844:*En-1* locus of maize (Pereira et al., 1985) and its right TIR from the *a1-m1*(6078) locus of maize (Pereira and Saedler 1989). The sequence of the empty donor site showed that the *I* element took one base pair of the *waxy 1* gene and three base pairs of the *a1* gene upon excision. Causing such deletions is common for *I* element excisions (Aarts et al., 1993).

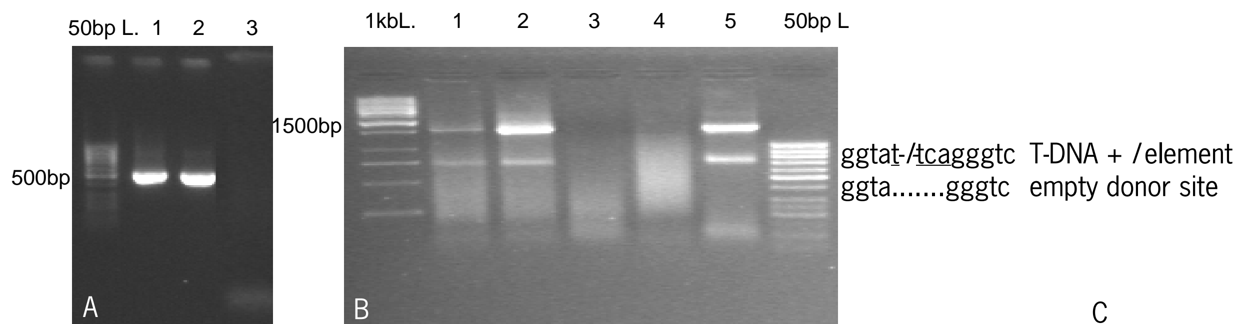


Fig 3 *PCR analysis of T-DNA insertion and I element excision in B. rapa. 50 bp L.: 50 bp ladder;1 kb L.:1 kb ladder; A: PCR using BAR primer shows insertion of En/I activation tagging construct in samples 1 and 2 but not in sample 3; B: PCR using primers flanking the I element shows excision of the I element in sample 1, 2 and 5. Samples 3 and 4 show no excision; C: Flanking DNA sequences of the original I element in the En/I T-DNA (top) and of the empty donor site after I element excision.*

Stable plant transformation

For the first transformation experiment we followed the protocol described by Wahlroos et al. (2003), who reported a transformation frequency of up to 9% for *B. rapa* subsp. *oleifera* using *A. tumefaciens* AGL1 harbouring a binary vector carrying the enhanced green fluorescent protein (eGFP, Clontech, Palo Alto, CA, USA) gene controlled by the CaMV 35S promoter. Initially we used the same genotype as used by Wahlroos et al. (2003), which we received on request. However, based on morphology and upon testing by flow cytometry the provided accession appeared to be *B. napus*, rather than *B. rapa* and we discontinued using this genotype. Instead we used L58 and L143. While the regeneration frequency was about 9%,

only 3% of the explants showed green calli on callus induction medium. Unfortunately they stopped growing and none of them regenerated shoots when transferred to shoot regeneration medium.

In experiment 2 we tested the protocol by Radke et al. (1992), who reported a transformation frequency of 1-9% using cultivars Tobin and Emma of *B. rapa* subsp. *oleifera*. Regeneration and transformation percentages are shown in table 1. Two percent of the L58 explants regenerated transformed shoots when treated with $AgNO_3$, as confirmed by analysis of GUS expression (Fig 4).

Table 1. *Result of regeneration and transformation experiment 2 using AGL0-GUS*

	Regeneration				Transformation			
	+ $AgNO_3$		- $AgNO_3$		+ $AgNO_3$		- $AgNO_3$	
genotypes	Green calli	Regenerated shoots	Green calli	Regenerated shoots	Green calli	Regenerated shoots	Green calli	Regenerated shoots
L143	100	25	90	0	40	0	10	0
L58	89	16	4	0	42	2	10	0

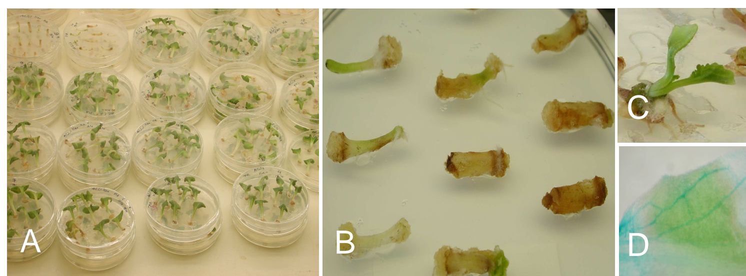


Fig 4 *Transformation of B. rapa. A: cotyledons of L143 on shoot regeneration medium with $AgNO_3$ after inoculation with AGL0-GUS. B: Hypocotyls of L58 on shoot regeneration medium with $AgNO_3$ after inoculation with AGL0-GUS. C: Transformed shoot of B. rapa L58 with AGL0-GUS. D: GUS expression (blue) in a leaflet of a transformed plant of B. rapa L58.*

Experiment 3 followed the protocol of experiment 2, but AGL1-*En/I* was used instead of AGL0-GUS. Only 4% green calli were obtained in this transformation experiment and none regenerated into a transformed shoot. In experiment 4, the concentrations of hormones in the shoot regeneration medium were adjusted, but still only about 4% green calli emerged, none of which regenerated into a shoot. We observed browning of explants which could be a sign

of ethylene production. For experiment 5 we followed experiment 3, but used another ethylene biosynthesis inhibitor AVG (aminoethoxyvinylglycine) instead of $AgNO_3$. This experiment did not yield any green calli. To further elaborate on limiting ethylene production liquid shoot regeneration medium was used. About 4% green calli was observed, which stay longer green than in previous experiments, but no transformed shoot regenerated. A high cytokinin: auxin ratio to induce direct shoot regeneration (experiment 7) was not successful either, as no transformed shoots were formed. In experiment 8 we further experimented with the cytokinin (BA) and auxin (NAA) concentrations in the shoot regeneration medium, in the presence or absence of gibberellic acid (GA3). Results are shown in table 2. 4 mg/l BA produced more green calli than others and 65% regenerated shoots in the absence of transformation, but in the presence of *A. tumefaciens* no transformed shoots were obtained.

Table 2. *Result of regeneration and transformation experiment 8 (using AGL1-En/I), combining five concentrations of BA, three of NAA and two of GA3 (all in mg/l) in the shoot regeneration medium.*

BA(mg/l)	NAA(mg/l)	GA3(mg/l)	Shoot regeneration%	Green calli induced%	Transformed shoot %
1	0.5	0	21	0	0
1	1	0	15	0	0
1	1	0	8	0.5	0
1	1	0.2	6	3	0
1	3	0	3	1	0
1	3	0.2	5	0.5	0
2	0.5	0	15	3	0
2	0.5	0.2	8.3	4.1	0
2	1	0	12	3	0
2	1	0	22	15	0
2	3	0	28	2	0
2	3	0	21	6	0
3	1	0	21	10	0
3	0.5	0.2	25	15	0
3	1	0	19	10	0
3	1	0	35	16	0
3	3	0	42	18	0
3	3	0	40	15	0
4	1	0	52	21	0
4	0.5	0.2	65	35	0
4	1	0	63	37	0
4	1	0	45	24	0
4	3	0	61	28	0
4	3	0	56	26	0
5	1	0	35	18	0
5	1	0	39	15	0
5	1	0	42	17	0
5	1	0	31	14	0
5	3	0	29	15	0

Burnett et al. (1994) reported that the presence of BA or NAA during seed germination medium markedly enhanced subsequent shoot regeneration from the base of the excised cotyledon explants of *Brassica rapa* cv. Horizon. We tested this option (experiment 9), and although about 3% of green calli were obtained, no significant difference was observed compared to previous experiments.

Block (1993) studied the cell biology of plant transformation and showed that during a preculture period dedifferentiation occurs. There are few reports about the effect of preculture on the transformation frequency in *Brassica* (Burnett et al., 1994; Babic et al., 1998; Tang et al., 2003; Tsukazaki et al., 2002). Tsukazaki et al. (2002) investigated the effect of the preculture period on transformation efficiency in hypocotyl explants of *Brassica oleracea* var. *capitata* Matsunami P22. The explants that were precultured for 3 days produced the highest

number of GUS-positive explants. Unfortunately we did not obtain any transformed shoots in this experiment (experiment 10).

Sharma et al. (1991) studied differentiation of shoots and roots from cotyledon explants of *Brassica juncea* (L.) Czern and observed that organogenesis occurred only if the proximal cut end of the petiole was in contact with the medium. In the absence of the petiole, differentiation from the lamina was rare. When this was tried (experiment 11) there was no significant difference between a cotyledon cut in half and cotyledon with petiole. No transformed shoot regenerated. In experiments 12 and 13 we also tested the effect of TDZ (Thidiazuron) (Qin et al., 2006), and brassinolide (Cardoza et al., 2004) in the shoot regeneration medium, but no significant effect was observed.

Agrobacterium related parameters such as virulence induction, transfer and integration are also important in transformation. Both *Agrobacterium* strains AGL0 and AGL1 carry a hypervirulent Ti helper plasmid. Still addition of acetosyringone to the co-cultivation medium was reported by Takasaki (Takasaki et al., 1997) to enhance the T-DNA transfer by inducing the *vir* gene expression of the Ti plasmid. About 4% green calli were obtained in this transformation experiment (14), but no transformed shoot regenerated.

After 14 experiments with no transformed shoots using the AGL1-*En/I* strain-construct combination, we wanted to test if this combination was at all able to transform plants. Therefore we used SR1 tobacco rather than *B. rapa* for transformation in experiment 15. For each explant inoculated with AGL0-GUS, on average ten transformed shoots regenerated, while for AGL0-*En/I* and AGL1-*En/I* this was four and two respectively (Fig 5).

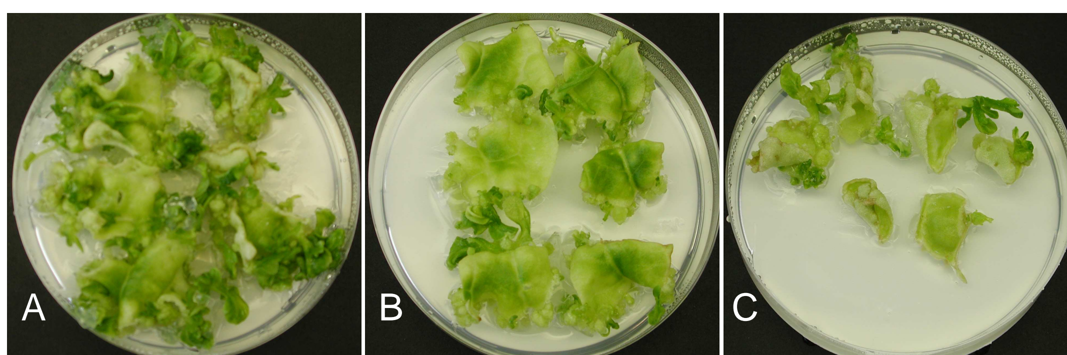


***Fig 5** Regeneration of transformed shoots of tobacco cv. SR1 ten days after co-cultivation with A: AGL0-GUS, B: AGL0 -En/I and C: AGL1- En/I.*

This showed that although the transformation with the *En/I* construct was clearly less efficient than with the GUS construct, in both cases transformants were obtained and thus the strain-construct combinations were all potentially able to give transformed shoots. Of course, tobacco is much easier to transform, so therefore we used explants of *Brassica napus* cultivar

Westar, which is a *Brassica* species, but easier to transform than *B. rapa* (experiment 16). In the shoot regeneration medium 90 % of the explants responded and regenerated shoots, in the absence of *A. tumefaciens*. Also in the presence of AGL0-GUS, 19 % (11/56) of the inoculated cotyledons regenerated transformed shoots. However, only 0.5 % (2/358) of the explants inoculated with AGL0-*En/I* or AGL1-*En/I* produced transformed shoots and they stopped growing after a few days. This indicates that the problem to obtain transformed shoots may not be the transformation protocol, but the construct used, with serious trouble for the cells containing the *En/I* construct to properly regenerate into transformed shoots (Fig 6).

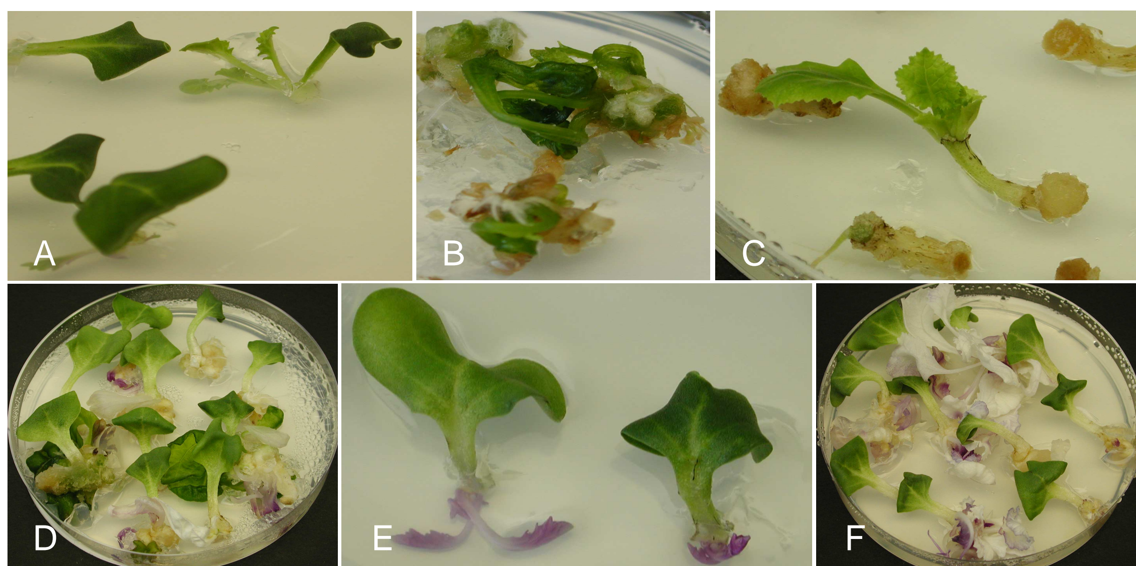


Fig 6 *Brassica napus regeneration and transformation. A: shoot regeneration on cotyledon; B: regeneration of transformed shoots with AGL0-GUS on hypocotyls explants; C: transformed shoot with AGL0-En/I on a hypocotyl explant; D: transformed shoot with AGL1-En/I on cotyledon explants; E and F: regeneration of untransformed shoots on cotyledon explants.*

As this may be caused by the tissue culture phase of the protocol, we tried alternative protocols. An *in planta* transformation protocol was proposed by Cardoza et al. (2004) for *Brassica*, which would negate the need for tissue culture. The *in planta* transformation protocol has been optimized for *Arabidopsis thaliana*, in which flowers are immersed in *A. tumefaciens* cultures, and in which ovules are the targets for transformation (reviewed by Pelletier and Bechtold, 2003). Following *in planta* transformation, seeds are harvested and transformed seedlings are selected. To date, the only crops to be transformed by this method have all been *Brassicaceae*: Pakchoi (*B. rapa* ssp. *chinensis*) (Liu et al., 1998; Xu et al., 2008), radish (*Raphanus sativus* L. *longipinnatus* Bailey) (Curtis and Nam 2001), and canola-quality *B. napus* (Wang et al., 2003), all with transformation frequencies of less than 0.20 %.

About 1200 seeds from L39 and 4000 of L58 were obtained after the infiltration. Plants were selected for resistance to kanamycin or Basta. No green plants were observed after selection for Basta resistance (Fig 7).

Fig 7 *In planta transformation of B. rapa. A: Three weeks after germination of seeds (upon two weeks of cold treatment) plants are ready for in planta transformation; B: The seeds obtained after in planta transformation have been germinated in a greenhouse to be sprayed with the herbicide Basta.*

Finally, rather then trying transformation with *A. tumefaciens,* using direct transformed shoot regeneration, we used *A. rhizogenes* to create transformed hairy roots, which could be used to select transgenic plants from. Sretenović-Rajičić et al. (2006) reported that up to nine percent spontaneous shoot regeneration of excised root cultures grown on the hormone-free medium occurred in *B. oleracea,* but we are not aware of a report of spontaneous shoot regeneration using *B. rapa* hairy roots. When using *A. rhizogenes* containing the *En/I* construct on explants of accessions L58 and R-o-18 we obtained hairy roots after 10 days (Fig 8). Three percent of the hypocotyl explants regenerated untransformed shoots. 100 transgenic hairy roots were transferred to hormone free medium or medium with BA and kept for six weeks. No shoots regenerated on both media, only white and green calli were observed on BA-containing media.

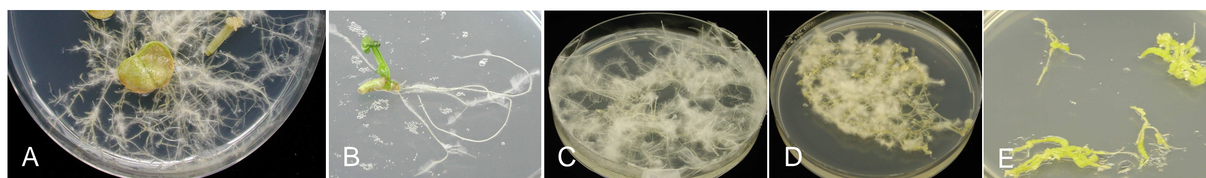


Fig 8 *Hairy root formation on B. rapa explants. A: Induction of hairy root on cotyledon and hypocotyl, B: Untransformed shoot regenerated from a hypocotyl explant inoculated with A. rhizogenes- En/I, C: Hairy roots growing after transferring to free hormone medium; D and E: Hairy roots in BA-containing medium.*

Discussion

Genetic engineering can potentially be used as a method to add specific characteristics to existing varieties. It would be most practical if efficient, genotype independent, and reproducible transformation and regeneration techniques are available. The most widely used method for the introduction of new genes into plants is based on the natural DNA transfer capacity of *A. tumefaciens*. In this work the aim was to make a mutant population of *B. rapa* using *En/I*–based activation tagging through *A. tumefaciens*-mediated transformation. The insertion of *En/I* through *A. rhizogenes*-mediated transformation and subsequent excision of the *I* element demonstrated that the *En/I* transposon system is active in *B. rapa* and could potentially be used in this species for heterologous gene tagging. However, re-insertion of the *I* element was not examined and this is needed for a proper insertional mutagenesis system. We showed that the *En/I* transposon activation tagging system should be preferred in *B. rapa* above an *Ac/Ds*-based system because of the higher transposition frequency. In Arabidopsis, also in contrast to the *Ac* element, a one element *En* system gave a much higher transposition frequency (Cardon et al., 1993). For gene tagging purposes the occurrence of independent transpositions is important, as each independent excision can give rise to independent reinsertions. Although we sequenced excision sites from a few different hairy root samples, we only detected one excision sequence. Normally excision of *En/I* is imperfect (Aarts et al., 1995) and different excisions leave different empty donor sequences, which was not the case in this experiment. Although we took great care not to contaminate samples, we cannot completely rule out cross-contamination of other samples with DNA from one excision site, explaining for the exact same sequence as we found.

Genotype, type of explant, explant pre-treatment, hormone combination of BA and NAA, $AgNO_3$, AVG, acetosyringone are important factors which we included in the stable transformation study to induce shoot regeneration. It has been shown that a large variation in regeneration frequency exists among *B. oleracea* (Tsukazaki et al., 2001) and *B. rapa* (Zhang et al., 1998; Narasimhulu and Chopra 1988) cultivars. In this study we used different *B. rapa* genotypes, which were all self-compatible and rapid cycling. When comparing them genotype L58 showed more green calli than L143 in transformation with *AGL0-GUS* and *AGL1-En/I.*

The positive role of cotyledon tissue on differentiation of shoots in *Brassica*, was reported by several groups (Sharma et al., 1991; Burnett et al., 1994; Teo et al., 1997). In this study we found no significant difference between hypocotyl or cotyledon explants in transformation frequency. A positive effect of acetosyringone has been reported for *B. rapa* transformation (Lim et al., 1998; Takasaki et al., 1997). However, we did not observe this in our transformation study, probably because both *A. tumefaciens* strains are already hypervirulent and virulence cannot be enhanced further in these strains.

Different optimal concentrations of BA and NAA have been reported for shoot regeneration of *B. rapa*. 5 mg/l BA and 0.5 mg/l NAA was reported by Jun II et al. (1995), whereas only 3 mg/l BA (Radke et al., 1992) or 5 mg/l BA and 3 mg/1 NAA (Mukhopadhyay et al., 1992) was reported by others as the best concentrations for shoot regeneration. In our study 4 mg/l BA and 0.5 mg/1 NAA showed more green calli when it compared with other combinations of BA and NAA. This appears to be more due to the BA than the NAA concentration, as also concentrations of 1 or 3 mg/l NAA gave higher percentages of green calli than other treatments.
The production of ethylene in tissue culture is known to increase under stress (Chi et al., 1990). Inclusion of ethylene biosynthesis or ethylene response inhibitors such as AVG and $AgNO_3$ in the shoot regeneration medium was another major factor reported to contribute to increased shoot regeneration efficiency in *Brassica* (Chi et al., 1990; Sethi et al., 1990). These compounds stimulate morphogenesis in many monocotyledon and dicotyledon species including *Brassicas* (Chi et al., 1990; Mukhopadhyay et al., 1992). However, we did not find any significant difference in shoot induction using these ethylene inhibitors.
Our study confirmed that *B. rapa* is a recalcitrant plant in regeneration and especially transformation. While we could optimize the medium to get 65 % shoot regeneration in the absence of *A. tumefaciens*, no transformation was observed in the presence of *A. tumefaciens*. Insertional mutagenesis in *Brassica* has not been described before to our knowledge. One of the reasons could be the low frequency of transformation in *Brassica*. Moon et al. (2007), Tsukazaki et al. (2001) and Narasimhulu and Chopra (1988) previously showed that *B. rapa* is one of the most recalcitrant *Brassica* species in tissue culture by comparing weedy *B. rapa* and introgressed *B. rapa* (Moon et al., 2007), comparing *B. rapa* cultivars with *B. oleracea* (Kuginuki and Tsukazaki et al., 2001), *B. nigra* and amphidiploid *B. juncea*, *B. napus* and *B. carinata* (Narasimhulu and Chopra 1988) genotypes. Zhang et al. (1998) showed large genotypic variation in shoot regeneration of *B. rapa*. When 123 genotypes were tested the regeneration frequency ranged from 0% to 95% (Zhang et al., 1998).
The length of a construct generally has a negative effect on transformation. The T-DNA region of the *En/I* activation tagging construct is 11 kb while the T-DNA size of the GUS construct is 5.3 kb. In tobacco, this decreased the transformation frequency by about 60 to 80%, although tested with a low number of explants. However, as we showed with the *B. napus* transformations, the main inhibiting factor for transformation might be the presence of the *SU1* negative selection marker gene in the *En/I* construct, which reduced the frequency of transformed *Brassica napus* plants from 19% to zero. In Arabidopsis expressing the *SU1* gene, even in the absence of selection, exhibits a different morphology compared to untransformed plants (Tissier et al., 1999). The transformed plants are shorter, bushier, and darker green. Therefore in *Arabidopsis* the *SU1* phenotype served as a visual marker

(Schneider et al., 2005; Marsch-Martinez et al., 2002). In developing a transposon activation tagging population of rice only the *bar* gene has proven to work efficiently (Zhu et al., 2007). Based on our experiments we conclude that the presence and potential expression of the *SUI* gene in green tissues, when controlled by the promoter of the Rubisco small subunit as in the construct we used, leads to a sublethal phenotype in *Brassica*, which prevents proper regeneration of viable shoots.

Hairy root transformation in *B. oleracea* (Sparrow et al., 2004) and *B. napus* (Damgaard and Rasmussen., 1991) has led to direct transformed shoot formation. In our study just some green and white calli were observed on transformed roots, and no green shoots were formed. Again, this can be attributed to expression of the *SU1* gene in green tissue.

In conclusion, we showed that hairy root transformation of *B. rapa* is a fast and easy method to test the action of transformed genes, including the activation of the *En/I* transposon. Omitting the *SU1* gene from the activation tagging construct may allow the production of transgenic *B. rapa* plants containing the *En/I* activation tagging construct, but it will not allow the easy selection against the presence of the *En* transposase gene after excision and reinsertion of the activation tag transposon.

References

Aarts M.G.M., Corzaan P., Stiekema W.J., Pereira A. (1995). A two-element Enhancer-Inhibitor transposon system in *Arabidopsis thaliana.* Mol Gen Genet. 247:555-64.

Aarts M.G.M., Dirkse W.G., Stiekema W.J., Pereira A. (1993). Transposon tagging of a male sterility gene in Arabidopsis. Nature. 363:715-7.

Azpiroz-Leehan R., Feldmann K.A. (1997). T-DNA insertion mutagenesis in Arabidopsis: going back and forth. Trends Genet. 13:152-6.

Babic V., Datla R.S., Scoles G.J., Keller W.A. (1998). Development of an efficient Agrobacterium-mediated transformation system for *Brassica carinata*. Plant Cell Rep. 17:183-188.

Bent A.F. (2000). Arabidopsis in planta transformation. Uses, mechanisms, and prospects for transformation of other sepecies. Plant Phyisiol. 124:1540-1547.

Bishop G.J., Harrison K., Jones J.D.G. (1996). The tomato Dwarf gene isolated by heterologous transposon tagging encodes the first member of a new cytochrome P450 family. Plant Cell. 8:959-69.

Block M.D. (1993). The cell biology of plant transformation: Current state, problems, prospects and the implications for the plant breeding. Euphytica. 71:1-14.

Burnett L., Arnoldo M., Yarrow S., Huang B. (1994). Enhancement of shoot regeneration from cotyledon explants of *Brassica rapa* ssp *oleifera* through pretreatment with auxin and cytokinin and use of ethylene inhibitors. Plant Cell Tiss Org. 37:253-256.

Cardon G.H., Frey M., Saedler H., Gierl A. (1993). Mobility of the maize transposable element *En*/*Spm* in *Arabidopsis thaliana*. Plant J. 3:773–784.

Cardoza V., Stewart C.N. (2004). Invited review: Brassica biotechnology: Progress in cellular and molecular biology. In Vitro Cell Dev Biol-Plant. 40:542-551.

Chalfun-Juniora A., Mes J.J., Mlynárová L., Aarts M.G.M., Angenent G.C. (2003). Low frequency of T-DNA based activation tagging in Arabidopsis is correlated with methylation of CaMV 35S enhancer sequences. FEBS. 555:459-463.

Chi G-L., Barfield D.G., Sim G-N., Pua E.C. (1990). Effect of $AgNO_3$ and aminoethoxyvinylglycine on in vitro shoot and root organogenesis from seedling explants of recalcitrant *Brassica* genotypes. Plant Cell Rep. 9:195-198.

Christey M.C. (2001). Use of Ri-mediated transformation for production of transgenic plants. In Vitro Cell Dev Biol-Plant. 37:687–700.

Christey M.C. Braun R.H. (2005). Production of hairy root cultures and transgenic plants by *Agrobacterium rhizogenes*-mediated transformation. Methods Mol Biol. 286:47-60.

Curtis, I.S., Nam H.G. (2001). Transgenic radish (*Raphanus sativus* L. longipinnatus Bailey) by floral-dip method--plant development and surfactant are important in optimizing transformation efficiency. Transgenic Res. 10:363-71.

Damgaard O., Rasmussen O. (1991). Direct regeneration of transformed shoots in *Brassica napus* from hypocotyl infections with *Agrobacterium rhizogenes*. Plant Mol Biol. 17:1-8.

Dorlhac F.de.B., Vincentz M., Chupeau Y., Vaucheret H. (1993). Co-suppression of nitrate reductase host genes and transgenes in transgenic tobacco plants. Mol Gen Genet. 243:613-621.

Engelen V.F.A., Molthoff J.W.,Conner A.J., Nap J.P., Pereira A., Stiekema W.J. (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. Transgenic Res. 4:288-90.

Gibson S., Somerville C. (1993). Isolating plant genes. Trends Biotechnol. 11:306-13.

Ichikawa T., Kawashima M.N.M., Muto S., Gohda K., Suzuki K., Ishikawa K.H.A., Yoshizumi T., Tsumoto Y., Tsuhara Y., Iizumi H., Goto Y. (2003). Sequence database of 1172 T-DNA insertion

sites in Arabidopsis activation-tagging lines that showed phenotypes in T1 generation. Plant J. 36:421–429.

Jones D.A., Thomas C.M., Hammond-Kosack K.E., Balint-Kurti P.J., Jones J.D. (1994). Isolation of the tomato Cf-9 gene for resistance to Cladosporium fulvum by transposon tagging. Science. 266:789-93.

Jun Il S., Kwon S.Y., Paek K.Y.,Paek K.H. (1995). *Agrobacterium*-mediated transformation and regeneration of fertile transgenic plants of chinese cabbage (*Brassica campestris* ssp. *pekinensis* cv. spring flavor). Plant Cell Rep. 14:620-625.

Kuginuki Y., Tsukazaki H. (2001). Regeneration ability and Agrobacterium-mediated transformation of different cultivars in *Brassica oleracea* and *B. rapa*. J Japan Soc Hort Sci. 70:682-690.

Lawrence G.J., Finnegan E.J., Ayliffe M.A., Ellis J.G. (1995). The *L6* Gene for Flax Rust Resistance is Related to the Arabidopsis Bacterial Resistance Gene *RPS2* and the Tobacco Viral Resistance Gene *N.* Plant Cell. 7:1195–1206.

Lazo G.R., Stein P.A., Ludwig R.A. (1991). A DNA transformation-competent Arabidopsis genomic library in Agrobacterium. Nat Biotechnol. 9:963-7.

Lim H.T., You Y.S.,Park E.J., Song Y.N. (1998). High plant regeneration, genetic stability of regenerants, and genetic transformation of herbicide resistance gene (BAR) in chinese cabbage (*Brassica campestris* ssp. pekinensis). Acta Hort. 459:199-208.

Limpens E., Ramos J., Franken C., Raz V., Compaan B., Franssen H., Bisseling T., Geurts R. (2004). RNA interference in *Agrobacterium rhizogenes*-transformed roots of Arabidopsis and *Medicago trunculataatula*. J Exp Bot. 55:983-92.

Liu F., Cao M.Q., Yao L., Li Y., Robaglia C., Tourneur C. (1998). In planta transformation of pakchoi (*Brassica campestris* L. ssp. chinensis) by infiltration of adult plants with Agrobacterium. Acta Hort. 467:187–192.

Marsch-Martinez N., Greco R., Van Arkel G., Herrera-Estrella L., Pereira A. (2002). Activation tagging using the *En-I* maize transposon system in Arabidopsis. Plant Physiol.129:1544-56.

McLaughlin M., Walbot V. (1987). Cloning of a mutable bz2 allele of maize by transposon tagging and differential hybridization. Genetics.117:771-776.

Moon H.S., Halfhill M.D., Good L.L., Raymer P.L., Neal Stewart C.Jr. (2007). Characterization of directly transformed weedy *Brassica rapa* and introgressed *B. rapa* with Bt cry1Ac and gfp genes. Plant Cell Rep. 26:1001-10.

Mukhopadhyay A., Arumugam N., Nandakumar P.B.A., Pradhan A.K., Gupta V., Pental D. (1992). Agrobacterium-mediated genetic transformation of oilseed *Brassica campestris*: Transformation frequency is strongly influenced by the mode of shoot regeneration. Plant Cell Rep. 11:506-513.

Murashige T., Skoog F. (1962). A revised medium for rapid growth and bioassays with tobacco cultures. Plant Physiol. 15:473 - 497.

Narasimhulu S.B., Chopra V.L. (1988). Species specific shoot regeneration response of cotyledonary explants of Brassicas. Plant Cell Rep. 7:104-106.

Nonaka S., Yuhashi K.I., Takada K., Sugaware M., Minamisawa K., Ezura H. (2008). Ethylene production in plants during transformation suppresses *vir* gene expression in *Agrobacterium tumefaciens*. New Phytol. 178:647-656.

Oldacres A.M., Newbury H.J., Puddephat I.J. (2005). QTLs controlling the production of transgenic and adventitious roots in *Brassica oleracea* following treatment with *Agrobacterium rhizogenes*. Theor Appl Genet.111:479-488.

Pelletier G., Bechtold N. (2003). In planta transformation. Transgenic plants: current innovations and future trends UK. Horizon Scientific Press. pp:65-82.

Pereira A., Saedler H. (1989). Transpositional behavior of the maize *En/Spm* element in transgenic tobacco. EMBO J. 8:1315-1321.

Pereira A., Schwarz-Sommer Z., Gierl A., Bertram I., Peterson P.A., Saedler H. (1985). Genetic and molecular analysis of the *Enhancer* (*En*) transposable element system of *Zea mays*. EMBO J. 4:17-23.

Poulsen G.B. (1996). Genetic transformation of Brassica. Plant Breeding. 115:209 - 225.

Puddephat I.J., Robinson H.T., Fenning T.M., Barbara D.J., Morton A., Pink D.A.C. (2001). Recovery of phenotypically normal transgenic plants of *Brassica oleracea* upon *Agrobacterium rhizogenes*-mediated co-transformation and selection of transformed hairy roots by GUS assay. Mol Breeding. 7:229–242.

Qin Y., Gao L., Pulli H., Guo Y.D. (2006). Shoot differentiation, regeneration of cauliflower and analysis of somaclonal variation by RAPD. Hereditas. 143:91-8.

Radke S.E., Turner J.C., Facciotti D. (1992). Transformation and regeneration of *Brassica rapa* using *Agrobacterium tumefaciens*. Plant Cell Rep. 11:499-505.

Schneider A., Kirch T., Gigolashyili T., Mock H.P., Sonnewald U., Simon R., Flugge U.I., Werr W. (2005). A transposon-based activation-tagging population in *Arabidopsis thaliana* (TAMARA) and its application in the identification of dominant developmental and metabolic mutations. FEBS. 579:4622-4628.

Sethi U., Basu A., Mukherjee G. (1990). Control of cell proliferation and differentiation by modulators of ethylene biosynthesis and action in *Brassica* hypocotyl explants. Plant Sci. 69:225-229.

Sharma K.K., Bhojwani S.S., Thorpe T.A. (1991). The role of cotyledonary tissue in the differentiation of shoots and roots from cotyledon explants of *Brassica juncea* (L.) czern. Plant Cell Tiss Org. 24:55-59.

Sparrow P.A., Townsend T.M., Arthur A.E., Dale P.J., Irwin J.A. (2004). Genetic analysis of *Agrobacterium tumefaciens* susceptibility in *Brassica oleracea*. Theor Appl Genet. 108:644-50.

Speulman E., Metz P.L.J., van Arkel G., Hekkert B.L., Stiekema W.J., Pereira A. (1999). A Two-Component Enhancer-Inhibitor Transposon Mutagenesis System for Functional Analysis of the *Arabidopsis* Genome. Plant Cell. 11:1853-1866.

Sretenović-Rajičić T., Ninković S.,Miljuš-Đukić J., Vinterhalter B., Vinterhalter D. (2006). *Agrobacterium rhizogenes*-mediated transformation of *Brassica oleracea* var. sabauda and *B. oleracea* var. capitata. Biol. Plant. 50:525-530.

Takasaki T., Hatakeyamae K., Ojima K., Watanabe M., Toriama K., Hinata K. (1997). Factors influencing Agrobacterium-mediated transformation of *Brassica rapa* L. Breeding Sci. 47:127-134.

Tang G.X., Zhou W.J., Li H.Z., Mao B.Z., He Z.H., Yoneyama K. (2003). Medium, explant and genotype factors influencing shoot regeneration in Oilseed *Brassica* spp. J Agron Crop Sci. 189:351 - 358.

Teo W., Lakshmanan P., Kumar P., Goh C.J., Swarup S. (1997). Direct shoot formation and plant regeneration from cotyledon explants of rapid-cycling *Brassica rapa*. In Vitro Cell Dev-Plant. 33:288-292.

Tissier A.F., Marillonnet S., Klimyuk V., Patel K., Torres M.A., Murphy G., Jones J.D.G. (1999). Multiple independent defective suppressor-mutator transposon insertions in Arabidopsis: A tool for functional genomics. Plant Cell. 11:1841-1852.

Tsukazaki H., Kuginuki Y., Aida R., Suzuki T. (2002). Agrobacterium-mediated transformation of a doubled haploid line of cabbage. Plant Cell Rep. 21:257-262.

Wahlroos T., Susi P., Tylkina L., Malyshenko S., Zvereva S.,Korpela T. (2003). Agrobacterium–mediated transformation and stable expression of the green fluorescent protein in *Brassica rapa*. Plant Physiol Bioch. 41:773-778.

Walden R., Fritze K., Hayashi H., Miklashevichs E., Harling H., Schell J. (1994). Activation tagging: a means of isolating genes implicated as playing a role in plant growth and development. Plant Mol Bio. 26:1521-1528.

Wang, W.C., Menon G., Hansen G. (2003). Development of a novel Agrobacterium-mediated transformation method to recover transgenic *Brassica napus* plants. Plant Cell Rep. 22:274-81.

Weigel D., Ahn J.H., Blázquez M.A., Borevitz J.O., Christensen S.K., Fankhauser C., Ferrándiz C., Kardailsky I., Malancharuvil E.J., Neff M.M., Nguyen J.T., Sato S., Wang Z.Y., Xia Y., Dixon R.A., Harrison M.J., Lamb C.J., Yanofsky M.F., Chory J. (2000). Activation tagging in Arabidopsis. Plant Physiol. 122:1003-13.

Whitham S., McCormick S., Baker B. (1996). The N gene of tobacco confers resistance to tobacco mosaic virus in transgenic tomato. Proc Natl Acad Sci. 93:8776-81.

Xu H., Wang X., Zhao H., Liu F. (2008). An intensive understanding of vacuum infiltration transformation of pakchoi (*Brassica rapa* ssp. chinensis). Plant Cell Rep. 27:1369-76.

Zhang F.L., Takahata Y., Xu J.-B. (1998). Medium and genotype factors influencing shoot regeneration from cotyledonary explants of Chinese cabbage (*Brassica campestris* L. ssp. pekinensis). Plant Cell Rep. 17:780-786.

Zhu Q.H., Eun M.Y., Han C.D., Kumar C.S., Pereira A., Ramachandran S., Sundaresan V., Eamens A.L., Upadhyaya N.M., Wu R. (2007). Transposon insertional mutants: a resource for rice functional genomics. Rice Functional Genomics. Springer New York. pp: 223-271.

Zuo J., Niu Q.w., Ikeda Y., Chua N.H. (2002). Marker-free transformation: increasing transformation frequency by the use of regeneration-promoting genes. Curr Opin Biotech. 13:173-180.

Chapter 3

Identification of seed related QTLs in a new genetically segregating *Brassica rapa* F2 population

Hedayat Bagheri[1], Dunia Pino Del Carpio[2], Guusje Bonnema[2], Maarten Koornneef[1,3] and Mark G.M. Aarts[1]

[1]Laboratory of Genetics and [2]Laboratory of Plant Breeding, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, the Netherlands, [3]Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, 50829 , Cologne, Germany.

Abstract

Two reciprocal F2 populations of 190 individuals each were developed, made by crossing *Brassica rapa* genotype L58, which is a rapid-flowering, self-compatible Caixin line, with R-o-18, a self-compatible Indian doubled haploid spring oil line. A linkage map was constructed for the L58 × R-o-18 population using 97 AFLP markers and 21 publicly available SSRs, covering a total map distance of 757 cM with an average resolution of 6.4 cM. A total of 15 QTLs for eleven traits including flowering time, total plant height, plant height until first open flower, number of siliques, number of seeds per silique, seed size, seed weight, seed oil content, seed colour, adventitious root and, transgenic hairy root formation were identified. A strong QTL explaining 52 % of the seed colour variation was found on linkage group A9. Seed colour and seed size were controlled by the maternal plant genotype rather than the seed genotype. The seed coat colour QTL co-localized with QTLs for seed size, seed weight, seed oil content, number of siliques and number of seeds per silique but no correlation was detected between seed coat colour and seed oil content when comparing all genotypes. This is probably due to the presence of additional independent loci with opposite effects involved in controlling seed coat colour and seed oil content. Instead, a significant positive correlation was observed between seed coat colour and seed oil content only in the light coloured classes (scored classes of 1- 6 from 13 classes, 1=Yellow till 13=Black). Plants with more siliques have more seeds and higher seed oil content but smaller seeds. One QTL on adventitious root formation of F3 cotyledons and one on transgenic hairy root formation of F3 seedlings were detected on A10. These loci will be important for selection of lines that are more efficient in *B. rapa* regeneration and transformation.

Introduction

Brassica rapa L. belongs to the family of Brassicaceae. This family contains well-known crop species such as *Brassica oleracea*, *B. juncea, B. nigra*, *Raphanus sativus* and *Sinapis alba*, but also *Arabidopsis thaliana*, which is the best studied plant model organism. The haploid *B. rapa* genome, or “A” genome, consists of ten chromosomes (n=10) and is one of the parents of the amphidiploid *B. napus* (U, 1935). Wild *B. rapa* has a wide distribution from the Western Mediterranean region to Central Asia. It is most likely the first *Brassica* crop that was domesticated more than 4000 years ago (Quijada et al., 2007). It is cultivated as a vegetable, fodder and oilseed crop. Domestication led to a wide variety of forms, such as the leafy vegetables of which *B. rapa* ssp. *chinensis* (Pak-choi) and *B. rapa* ssp. *pekinensis* (heading Chinese cabbage) are best known; root vegetables like *B. rapa* ssp. *rapifera* Metzg. (turnip) (Quijada et al., 2007); and oil types like ssp. *oleifera* (Chinese oil turnips) and ssp. *trilocularis* (Yellow Sarson). Zhao et al. (2005) used AFLP fingerprinting on 161 *B. rapa* accessions collected from different parts of the world and distinguished two main groups: one of Asian and one of mainly European origin, and a small group of Indian oil types. Each group contained mostly similar morphotypes, but different morphotypes from the same origin were often more related to each other than to similar morphotypes from different origins.

Greater seed yield and yield stability besides seed quality are the primary objectives in seed crop breeding. Seed contributes directly to the economic success of commercial crops. Seed yield is the resultant of some determining components, i.e., the number of siliques per unit area, branch number, number of seeds per silique, and seed size (Quijada et al., 2007). The seed of *B. rapa* is mainly used for oil production, although the meal remaining after oil extraction is also of economic interest.

Mainly F2 maps are available for *B. rapa* (McGrath and Quiros 1991; Chyi et al., 1992; Teutonico and Osborn 1994; Ajisaka et al., 1995; Matsumoto et al., 1998; Zhang et al., 2000; Lu et al., 2002). The first *B. rapa* genetic linkage map was made from an F2 population of 95 individuals derived from the cross between a Chinese cabbage cultivar and a spring broccoli accession (Song et al., 1991). Doubled haploid (DH) populations, which are commonly used for Brassicas, are difficult to develop for *B. rapa* as some genotypes respond very poorly to DH induction (Kole et al., 1997). In general a higher level of segregation distortion is found in DH population due to preferential selection of genotypes responsive to microspore or anther culture (Suwabe et al., 2004). By now several DH (Lim et al., 1998; Suwabe et al., 2004; Wang et al., 2004; Choi et al., 2007; Lou et al., 2007) and Recombinant Inbred Line (RIL) (Kole et al., 1997; Novakova et al., 1996; Yu et al., 2003) populations of *B. rapa* have been described which have been used for genetic mapping. Available RIL populations have been developed from crossings between the biennial cultivar “Per” and the annual cultivar “R500” (Kole et al., 1997), between a Chinese cabbage and a Mizuna line and between two

Chinese cabbage lines, respectively (Novakova et al., 1996; Yu et al., 2003). The DH population created by Lim et al. (1998), generated from a cross between two morphologically diverse Chinese cabbage inbred lines, “Chiifu” and “Kenshin”, is used as a reference for *B. rapa* genetic mapping and for genome sequencing of *B. rapa* “Chiifu” (http://www.brassica.info/resource/sequencing.php).

More than ten decades ago Smith and Townsend (1907) demonstrated that *Agrobacterium tumefaciens* causes plant tumors. T-(transferred) DNA is transferred from the bacterium to the plant cell and expression of several genes in the T-DNA results in tumor formation. These genes affect the concentrations of the phytohormones auxin and cytokinin in transformed tissue (Akiyoshi et al., 1983). Tumors induced by wild-type *Agrobacterium* contain high concentrations of both cytokinin and auxin relative to untransformed callus. Today the natural gene transfer capacity of this soil bacterium is the most widely used method of delivering transgenes into many species like Brassica (Puddephat et al., 1996). A number of virulence (Vir) proteins are involved in T-DNA transmission from Agrobacterium Ti (tumor inducing) or Ri (root inducing) plasmids into the genomes of higher plants. Certain phenolic and sugar compounds of the plant defense system from wounded plant cells serve as inducers (or co-inducers) of the bacterial *vir* genes (Gelvin, 2000).

Brassica rapa is known as one of the most recalcitrant members of *Brassica* genus to regenerate shoots in vitro (Narashimhulu and Chopra 1988). Also successful transformation of *B. rapa* by *A. tumefaciens* has been hampered by the lack of efficient plant transformation and regeneration techniques. Therefore there is a need for developing efficient transformation methods to overcome genotype dependency.

Hairy root transformation has become a useful, easy and fast technique for introducing interested genes into plant cells. *A. rhizogenes*-mediated root transformation is a fast method to generate adventitious, genetically transformed roots. *A. rhizogenes* is capable of transferring the T-DNA of Ri (root inducing) plasmid to plants (Christey 2001). Especially in species recalcitrant to transformation, such as *Brassica*, this technique is a valuable tool. As virulence genes in T-DNA are highly conserved and important factors involved in T-DNA transfer are common to both *A. tumefaciens* and *A. rhizogenes*, the analysis of plant genes regulating transformation using either *A. rhizogenes* or *A. tumefaciens* will lead to the understanding of process common to both (Cogan et al., 2002). Marker-assisted selection has been used for successful breeding of regeneration and transformation frequency. Koornneef et al (1993) used a genetic analysis to screen for high or low regeneration potential of *Lycopersicon esculentum* plants in breeding. In Arabidopsis, mutant lines that are resistant to *A. tumefaciens* have been identified (Mysore et al., 2000). Cogan et al (2002) also identified quantitative trait loci (QTL) for transgenic and adventitious root production using an *A. rhizogenes*-mediated co-transformation system in conjunction with a *Brassica oleracea*

double haploid (DH) mapping population. Auxin is involved in the process of adventitious root initiation. Endogenous auxin levels and the sensitivity of cells to exogenous auxin are important parameters in this process (Blakesley et al., 1991; Audus 1959).

In this report we describe a new F2 mapping population, as starting material for the development of a new RIL population. The population was derived from a cross between two distinct morphotypes, Cai Xin and Yellow Sarson, both early flowering and self-compatible to ensure rapid propagation after repeated self-fertilization of each line. Thus a mapping population can be obtained that is easy to maintain, once propagated to F8 or beyond through single-seed descent and that can be analyzed for a large number of traits. The Cai Xin parent is L58, a broccoletto vegetable type originating from China (*B. rapa* ssp. *parachinensis*) (first described by Wu et al., 2007). The other parent, R-o-18, a doubled haploid Yellow Sarson oil type line (*B. rapa* ssp. *trilocularis*), which has been used as susceptible parent in F2 populations for genetic analysis of disease resistance to TuMV and *Xanthomonas campestris* (Rusholme et al., 2007; Soengas et al., 2007). This line is also used for microsatellite development (Lowe et al., 2004), and TILLING purposes (McCallum et al., 2000) to identify mutations in selected genes using an organized mutated population (http://www.brassica.info/research/activities/tilling.php). The parents differ in many traits and the population is suited to study characteristics for both vegetable and oil seed types.

Results

Genotyping and the construction of a linkage map

A linkage map was constructed for the L58 × R-o-18 F2 population consisting of 190 individuals using 97 polymorphic AFLP markers and 21 publicly available SSR markers, covering a total map distance of 757 cM with an average distance of 6.4 cM between markers (Figure 1). Selective amplification with the PM (*Pst*I/*MseI*) primer combinations showed more amplification products than with the EM (*Eco*RI/*Mse*I) primer combinations. In total, 300 bands were produced using four PM and seven EM primer combinations. Of these, 158 were polymorphic, which is 52% of the identified bands. 36 AFLP markers were excluded from the analysis because they could not be scored for many F2 plants or they showed an identical segregation as another marker. 25 AFLP markers could not be assigned to a linkage group at the high LOD score. 23 out of 36 SSRs tested showed polymorphisms for the parents and 21 of them were mapped, allowing the assignment of linkage groups to their respective chromosomes as presented in the reference *B. rapa* linkage map (Choi et al., 2007).

Based on the chi-square test for goodness-of-fit to the expected 1:2:1 Mendelian segregation ratio, six SSR markers showed a segregation distortion at $P<0.005$. The corresponding loci are found on chromosomes A1, A2 and A5 (Figure 1).

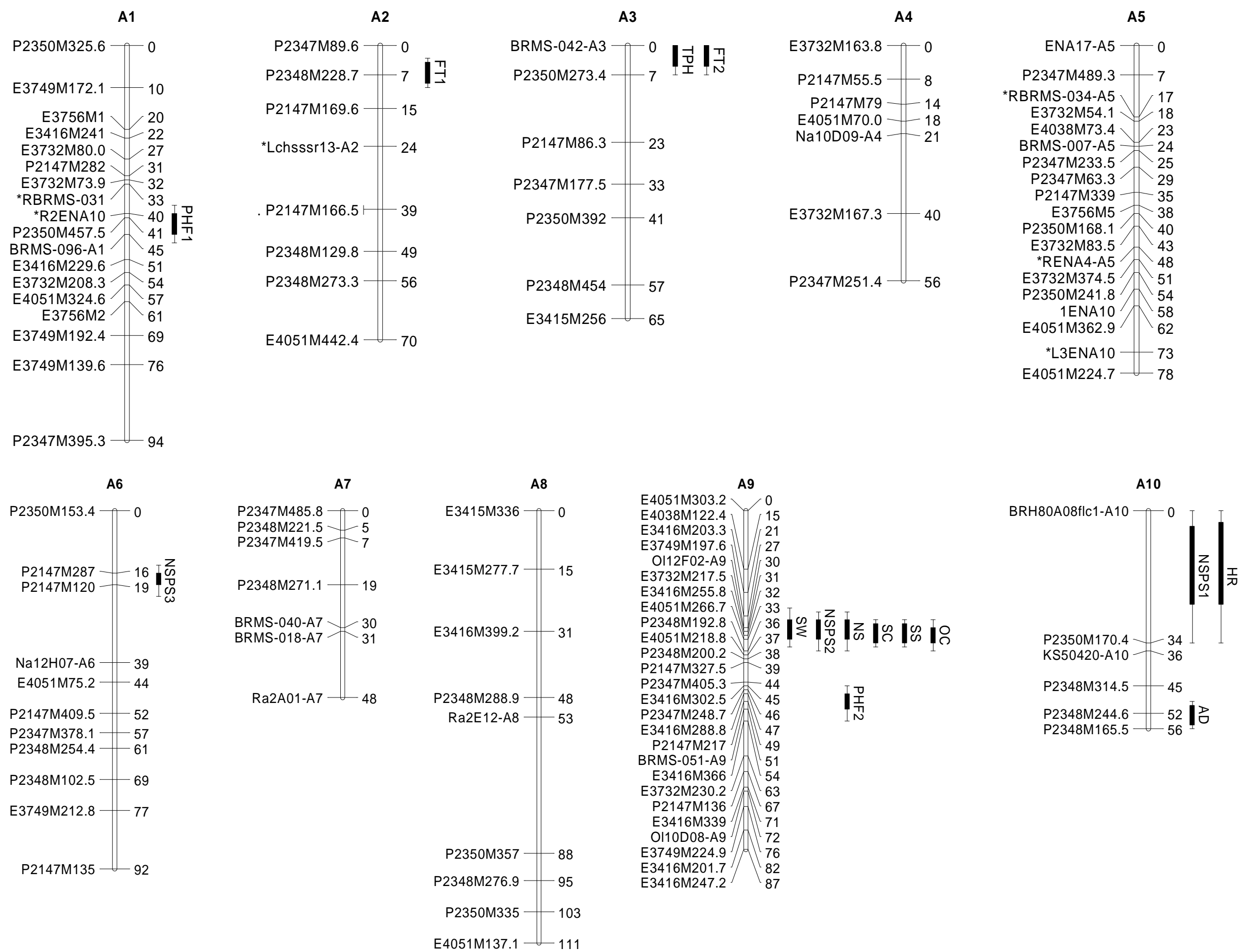


Figure 1 *Genetic linkage map of the F2 B. rapa L58 × R-o-18 F2 population, showing the positions of 97 AFLP and 21 SSR markers distributed over 10 linkage groups. The linkage groups correspond to the 10 chromosomes of the B. rapa reference linkage map (Choi et al., 2007). QTLs mapped to the linkage map for the determined traits are indicated with boxes and whiskers representing 1-LOD and 2-LOD confidence intervals (95%) respectively for significant QTLs. FT= flowering time; TPH= total plant height; PHF= plant height until first open flower; NSPS= number of seeds per silique; OC= seed oil content; SS=seed size; SC=seed colour; NS= number of siliques; SW=seed weight; AD: adventitious root and HR: hairy root formation. Skewed SSR markers in A1, A2 and A5 are indicated with * and L (skewed to L58) or R (skewed to R-o-18).*

Phenotyping of the F2 population

Both parents are rapid flowering. L58 and R-o-18 flower on average after respectively 29 days and 39 days after sowing (Figure 2). The minimum and maximum flowering time for the F2 lines was 30 and 58 days respectively. This was similar in the F2 derived from the reciprocal cross, which was not used for linkage analysis (respectively 30 and 64 days). Different visible phenotypes, as well as seed oil content, were recorded for the L58 × R-o-18 F2 population.

A very prominent phenotype segregating in the population was seed colour. L58 has brown-black seeds and line R-o-18 has yellow seeds. The size and colour of F1 seeds of the reciprocal crosses corresponded to that of the female parents, indicating that these seed phenotypes reflected the genotype of the mother plant (Figure 2).

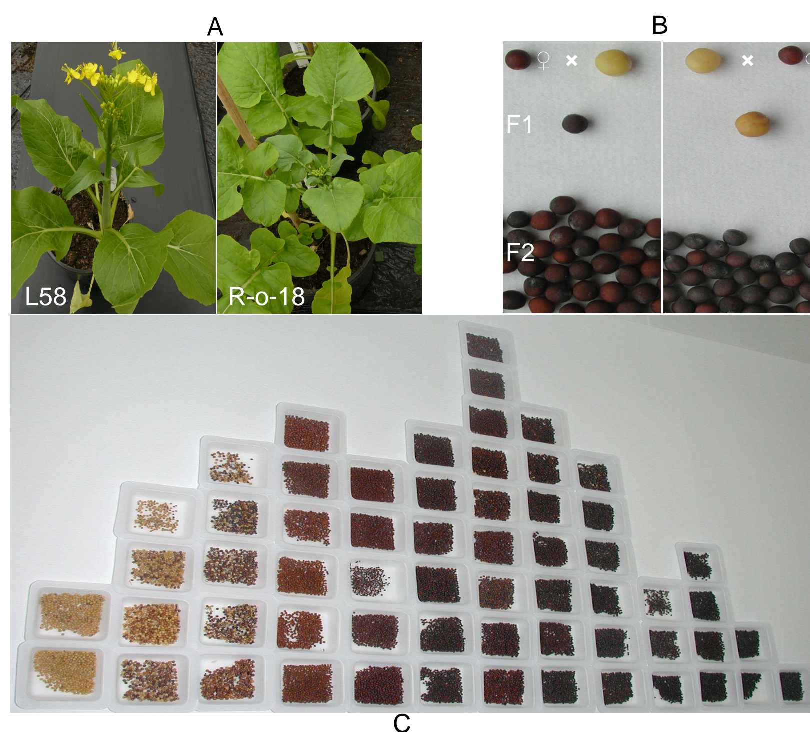


Figure 2 *Morphological phenotypes of plants (A) and seeds (B and C) of parental lines and progeny of the L58 × R-o-18 population. A) L58 (left) and R-o-18 plants (right), 30 days after sowing. B) Seed colour in the parents (Black: L58; Yellow: R-o-18), F1, F2 and reciprocals of the L58 × R-o-18 cross and C) The variation in F3 seeds was divided into 13 different classes (1=Yellow till 13=Black).*

Figure 3 shows the frequency distributions for the number of siliques per plant; number of seeds per silique; seed weight; flowering time; plant height; silique length; seed size; seed colour; and seed oil content. Transgression beyond the parental values was observed for most of the traits except for seed colour. For the number of siliques, seed oil content and flowering time, the transgression is mainly beyond the R-o-18 parent, which has the highest parental value. For the other traits, transgression is to both sides.
Correlation analysis of the measured traits showed that flowering time was highly correlated with total leaf number, plant height until first open flower and leaf number below 20 cm (Table 1). There was no significant correlation between seed colour and seed oil content when data for all lines were examined, however, when only taking the lines classified as light coloured (seed coat colour clusters 1 to 6, Figure 2) there was a highly significant positive correlation (r= 0.62) between seed colour and seed oil content, so genotypes with reddish-brown seed coat had higher oil content than the yellow-seeded genotypes. Silique number and seed number per silique were positively correlated to each other but negatively correlated with seed weight. Seed oil content was positively correlated with the number of siliques. Seed weight and seed size were highly positively correlated, and they were negatively correlated to seed oil content and silique number. In general, plants with more siliques had more, but smaller seeds and higher seed oil content.

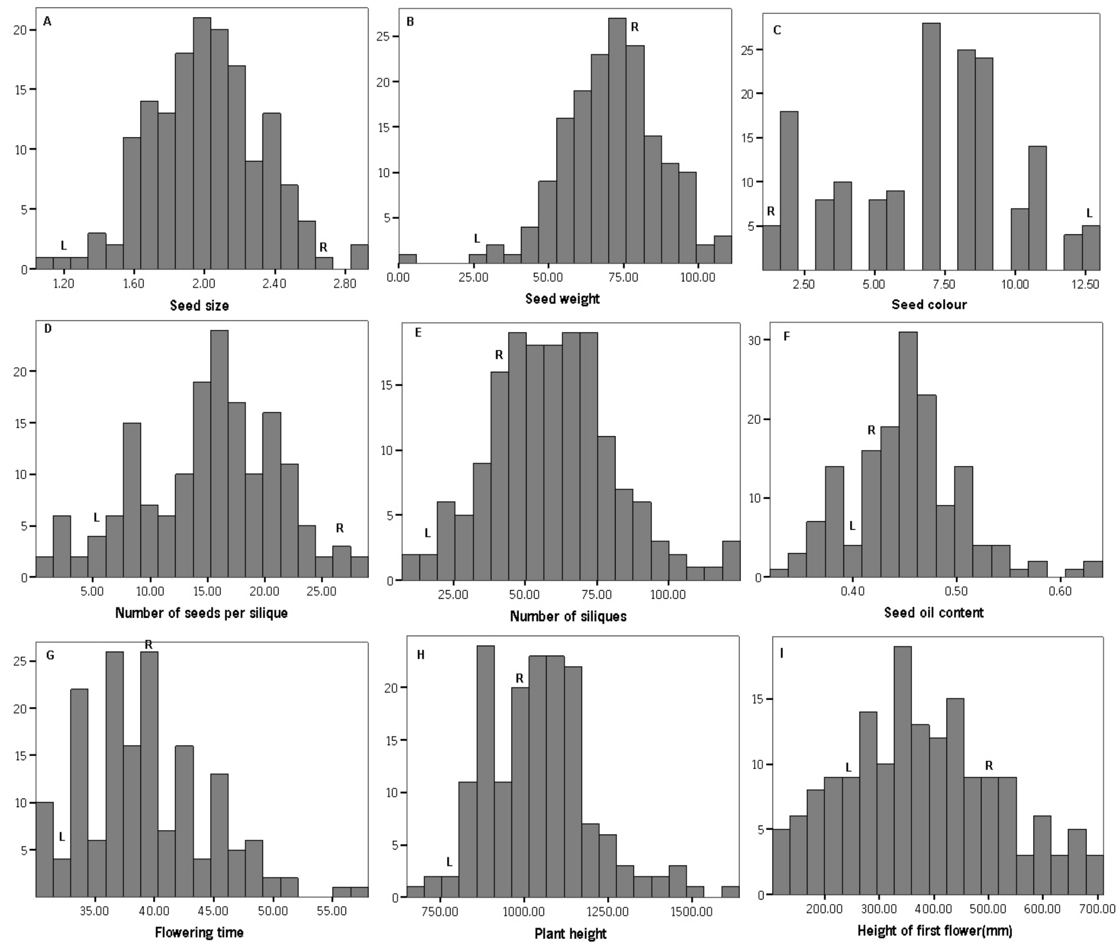


Figure 3 *Frequency distributions of non-normalized data of nine traits in the L58 × R-o-18 F2 population. The vertical axis indicates the number of lines per trait value class and the horizontal axis the different trait value classes. (A) seed size (mm); (B) seed weight (of 10 seeds in mg); (C) seed colour; (D) number of seeds per silique; (E) number of siliques per plant; (F) seed oil content (mg); (G) flowering time (days after sowing); (H) total plant height (mm); and (I) plant height until first open flower (mm). The parental values are the mean of five replicates, indicated with L as L58 and R as R-o-18.*

***Table 1** Pearson correlation analysis of the measured traits.*
*NS: Number of siliques; NSPS: number of seeds per silique; SW: seed weight; OC: seed oil content; PC: seed protein content; SC: seed colour; SS: seed size; FT: flowering time; LNF: leaf number until first flower; TPH: total plant height; SL: silique length; TLN: total leaf number; PHF: plant height until first open flower; BN: branch number. ** means significant at P≤0.01; * significant at P≤0.05.*

Trait	NS	NSPS	SW	OC	PC	SC	SS	FT	LNF	TPH	SL	TLN	PHF
NS	1												
NSPS	0.34**												
SW	-0.37**	-0.31**											
OC	0.42**	0.38**	-0.69**										
PC	0.07	0.04	-0.28**	0.20*									
SC	-0.02	-0.02	0.09	0.03	0.09								
SS	-0.54**	-0.51**	0.95**	-0.68**	-0.22*	0.1							
FT	0.36**	0.13	-0.37**	0.39**	0.12	-0.25**	-0.36**						
LNF	0.02	-0.07	-0.007	0.16*	-0.005	-0.16*	-0.02	0.50**					
TPH	-0.01	-0.11	0.14	0.14	0.03	-0.29**	0.06	0.40**	0.53**				
SL	0.05	0.56**	0.12	0.07	0.01	0.05	0.01	-0.001	0.09	-0.12			
TLN	0.39**	0.18*	-0.28**	0.37**	0.13	-0.17*	-0.32**	0.88**	0.50**	0.31**	0.1		
PHF	0.34**	0.17*	-0.35**	0.32**	0.33**	0.1	-0.33**	0.63**	0.05	0.05	0.03	0.72**	
BN	0.04	-0.28*	0.05	0.02	0.09	0.03	0.01	0.34**	0.33**	0.28**	-0.18*	0.40**	0.16*

This F2 population showed segregation for other morphological traits such as leaf glossiness (cuticular wax composition), leaf shape, leaf chlorophyll content, petiole length, fertility, shattering, carpel number (trilocularis), vivipary, parthenocarpy, fertility, and anthocyanin content (stem), but these traits have not been measured.

QTL analysis

In total 15 QTLs were mapped for eleven different traits (Table 2; Figure 1). The strongest QTL was found for seed colour, mapping to chromosome A9. It co-located with other seed-related QTLs for seed size, seed weight and silique number. Correlation analysis (Table 1) already showed significant correlation between these traits, except however for seed colour. Figure 4 shows the box plots of phenotypic values for each of the three genotypic classes at each QTL region. In case of seed colour, genotypes homozygous for the L58 allele and heterozygotes are not significantly different from each other, and many lines classified in classes 7 to 13 (most black ones) are actually heterozygote for the QTL on chromosome A9 (Figure 5). They are significantly different from genotypes homozygous for the R-o-18 allele though, indicating that the L58 allele is fully dominant over the R-o-18 allele. In classes 1 to 6 (most yellow ones) many lines are homozygous for the R-o-18 allele for the QTL on chromosome 9. Genotypes homozygous for the L58 allele show a significant difference from

genotypes homozygous for the R-o-18 allele and heterozygotes in case of seed weight, seed oil content and number of seeds per silique (Figure 4).

Table 2 *QTLs detected in the F2 population for nine different traits: Flowering time (FT); plant height at first flower (PHF); total plant height (TPH); number (No.) of seeds per silique (NSPS); seed oil content (OC); seed size (SS); seed colour (SC), number of siliques (NS) and seed weight (SW). QTLs are numbered according to decreasing LOD score (LOD). %Expl. var. is the percentage of total phenotypic variance explained by individual QTLs. For each of the QTLs the allelic effect is indicated (Effect). These are calculated as µA-µB (µ= mean), where A and B are F2 carrying L58 and R-o-18 genotypes at the QTL positions, respectively. µA and µB were estimated by MapQTL. Effects are given in days (flowering time), millimeter (plant height, size), milligram per seed (oil content) or without unit (silique or seed number, seed colour).*

Trait	QTL	Linkage group	LOD	Position of peak LOD cM	% Expl. var.	Effect
Flowering time	FT1	A2	4.2	6.8	10.1	3.8
	FT2	A3	4.1	4.2	9.5	-5.2
Plant height until first open flower	PHF1	A1	4.7	41.3	11.3	-131.4
	PHF2	A9	3.4	45.2	8.0	-110.7
Total plant height	TPH	A3	4.9	4.2	13.6	-137.3
No. seeds/silique	NSPS1	A10	3.1	15.0	8.5	-3.0
	NSPS2	A9	2.9	30.1	3.8	4.6
	NSPS3	A6	2.6	17.2	20.7	5.0
Seed oil content	OC	A9	6.6	30.1	20.0	0.1
Seed size	SS	A9	7.0	30.1	19.0	-0.4
Seed colour	SC	A9	26.3	30.1	52.6	+5.4
Seed weight	SW	A9	4.6	30.1	12.0	-0.6
No. siliques	NS	A9	3.1	30.1	8.5	16
Adventitious root formation	AD	A10	5.2	52.0	36.0	-0.3
Transgenic hairy root formation	HR	A10	4.6	15.0	30.0	-0.1

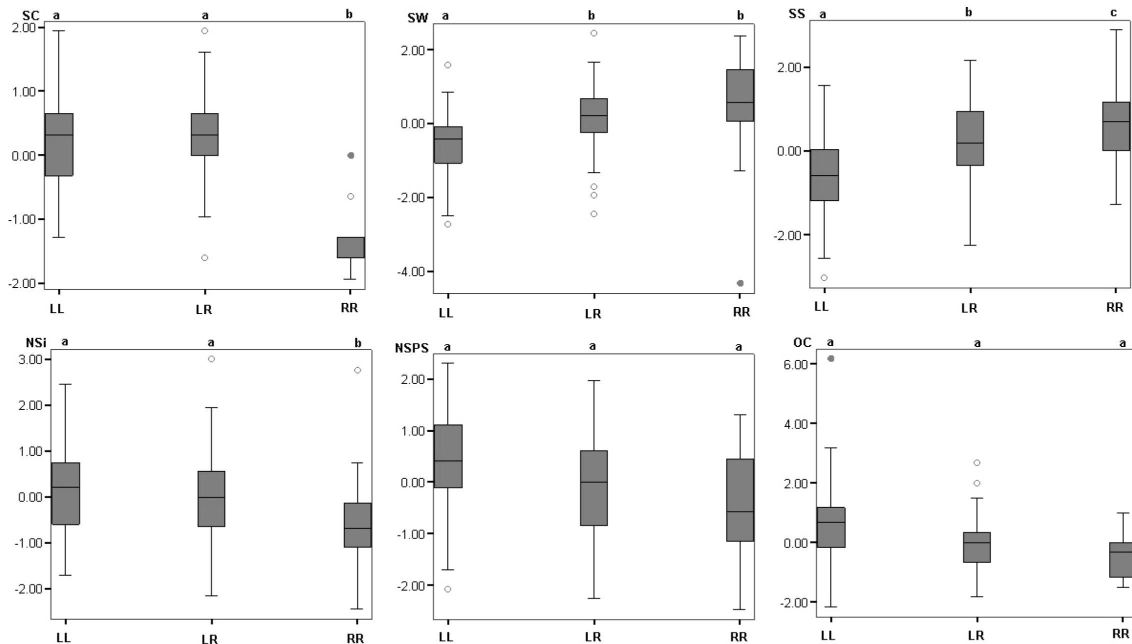


Figure 4 *Box plots of the phenotypic values of the three genotypic classes in F2 for six seed traits for which one co-localizing QTL was found around SSR Ol12F02 on chromosome A9. The X-axis shows the genotypic class (LL= homozygote for L58 allele, LR= heterozygote and RR= homozygote for R-o-18 allele) and the Y-axis the standard score Z (= X-μ/σ, X= trait value, μ= trait mean and σ=trait standard deviation), SC: seed colour. SS: seed size. SW: seed weight. NS: number of silique. OC: seed oil content. NSPS: number of seeds per silique. Box plots show the median, interquartile range, outliers (◦) and extreme cases (•). a, b and c are significance classes at p<0.05 in a t-test comparing mean values of genotypic classes.*

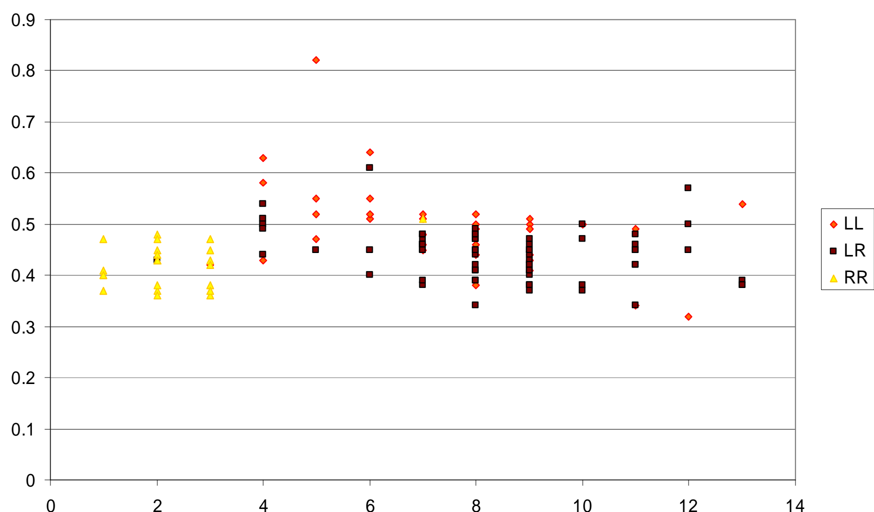


Figure 5 *Frequency of genotypic classes at the SC QTL region in each of the different seed colour classes. The X-axis shows the seed colour class and the Y-axis shows the frequency. LL= homozygote for L58 allele, LR= heterozygote and RR= homozygote for R-o-18 allele.*

Discussion

A segregating F2 population of *Brassica rapa* consisting of 190 individuals was developed from two distinct morphotypes that both are early flowering and self compatible. This population was used for genetic linkage mapping. Quantitative F2 phenotyping was performed for 16 traits. For several traits there is strong transgression of phenotypes beyond the parental values, often far beyond the parent with the highest value, suggesting the presence of several loci controlling a trait, with contrasting allele effects in this population. It makes this population a potentially rich tool for the analysis of genetic variation for *B. rapa* breeding purposes. In total 15 QTLs for 11 of the traits were detected, most of which have not been described before for *B. rapa*. Two new QTLs for "plant height below the first open flower" (PHF1, PHF2) were detected, which mapped to A1 and A9 respectively. PHF is a poor predictor of total plant height (TPH). We did not find significant correlation between both traits (table 1) and also did not find a common QTL (Figure 1). PHF QTLs also do not correspond to any of the plant height QTLs mapped by Lou et al. (2007) in F2/F3 populations of *B. rapa*, but the QTL for total plant height (TPH), which we mapped to the top of A3, may correspond to one QTL (out of three QTLs they identified), which they mapped to the same region.

Correlation analysis showed a significant positive correlation between flowering time and plant height. Although both parents are rapid-cycling, two QTLs were identified for flowering time (FT1, FT2). FT2 co-localizes with the TPH-QTL on the top of chromosome A3, the other maps to chromosome A2. There are four known flowering time genes in *B. rapa; BrFLC1*, *BrFLC2*, *BrFLC3* and *BrFLC5*, all corresponding to an orthologous copy of the *FLC* flowering time gene previously discovered in *A. thaliana* (Koornneef et al., 1994; Michaels and Amasino, 1999) and known to play a major role in determining natural variation for flowering time in *A. thaliana* (Shindo et al., 2005). The *BrFLC* genes are assigned to linkage groups A10, A2, and A3 respectively (Schranz et al., 2002; Kim et al., 2006). Unfortunately the resolution of the F2 population using anchored SSR markers is not sufficiently detailed to confirm co-localization of FT1 with *BrFLC2* or FT2 with *BrFLC3* or *BrFLC5*.

In this F2 population, seed weight, seed size, number of siliques, number of seeds per silique and seed oil content are highly correlated, and therefore it was not unexpected to find that these traits all share one co-localizing QTL on A9. This is the only QTL we detected for the number of siliques, however there are two additional QTLs for the number of seeds per silique, on A6 and A10. Both traits are positively correlated with seed oil content, but they are negatively correlated with seed size and seed weight. A negative correlation between the number of seeds per silique and seed size was previously also established for *A. thaliana* in which several QTLs were found to affect both traits (Alonso-Blanco et al., 1999). When

selecting for lines homozygous for the L58 allele at this locus, one will select for plants with many siliques and many, but small, seeds per silique with high oil content. Breeding for lines with more siliques is one approach to increase seed and oil yield and these loci may be of considerable importance for the rapeseed breeding industry.

Although *B. rapa* is a minor oilseed crop, it is one of the two ancestral species of *B. napus*, which is the second largest seed oil crop in the world. Seed coat colour is a very important trait in oilseed type *Brassica* crops. Yellow seeds are preferred, since yellow-seededness corresponds with significantly higher seed oil contents (Rakow et al., 1999) and the meal of yellow-seeded varieties that remains after extracting oil has higher protein and lower crude fiber contents than that of black-seeded varieties and is of better quality for use as poultry and livestock feed (Tang et al., 1997). We found one major QTL on A9 explaining 52 % of the seed coat colour variation (Figure 1 and Table 2). Probably the same locus for seed coat colour was previously mapped to A9 by Lou et al. (2007) in two *B. rapa* DH populations.

The black seed coat allele is over-dominant over the allele for yellow seed coat. Seed coat colour is a maternal trait (Figure 2), in agreement with the fact that seed pigments are deposited in the testa layers, which are maternal tissues (Vaughan et al., 1976). Ahmed and Zuberi (1971) already reported that a single gene is responsible for the dominant reddish brown seed colour in Indian *B. rapa* Toria's lines. Also Chen and Heneen (1992) showed that a single maternal gene controls seed colour with the black allele dominating over the yellow allele in *B. rapa*. However Schwetka (1982) described six genes to be involved in seed colour, with one having a pleiotropic effect on hilum colour in turnip rape (*B. rapa).* In the model proposed by Van Deynze and Pauls (1993) for *B. napus*, black seed colour is controlled by dominant alleles at three loci, and only homozygous recessive alleles at all of these loci confer yellow seeds.

Seed colour mutants are also described for *A. thaliana* as *transparent testa (tt)* mutants that have yellow or pale brown seeds. All known *tt* mutations are recessive and show maternal inheritance of the seed phenotype. The mutants *tt1, tt2, tt8, tt9, tt10, tt12,* and *tt15* appear to be altered in seed colour only (Debeaujon et al., 2001), whereas the other genes control anthocyanin accumulation in all parts of the plant and often encode flavonoid biosynthesis genes (Lepeniec et al., 2006). Orthologues of the Arabidopsis *TRANSPARENT TESTA 12* genes have already been cloned from *B. napus*, *B. oleracea* and *B. rapa* (Chai et al., 2008). *AtTT12* encodes a membrane-associated MATE transporter that functions as a vacuolar flavonoid/H+-antiporter to mediate proanthocyanidin (PA) accumulation in cells of the seed coat (Marinova et al., 2007). PA polymers are the major constituents of seed coat pigments of *Arabidopsis* and *Brassica*. Yellow-seeded oilseed rape varieties generally have thinner seed coats that result in reduced acid detergent fibers, mainly comprised of cellulose and lignin. Yellow-seededness generally also corresponds to increased seed oil and protein content

(Badani et al., 2006; Xiao and Liu 1982). A direct relationship between seed coat thickness and seed colour in *B. rapa*, independent of seed size, was already reported by Stringam et al. (1974). Like the testa-controlled seed colour, also seed size is maternally controlled. In Arabidopsis most seed pigmentation mutants show reduced seed weight and seed size (Debeaujon et al., 2000), which is different when comparing L58 and R-o-18, of which the yellow-seeded genotype R-o-18 has the largest seeds.

One QTL explaining 20% of the phenotypic variance for seed oil content was identified on A9, in the same region where the seed coat colour QTL was mapped (Figure 1 and Table 2). Oil content at the QTL level is still poorly understood in *B. rapa*. A RAPD (Random Amplified Polymorphic DNA) marker linked to an oleic acid QTL was mapped on A6 by Tanhuanpää et al. (2004) in *B. rapa* ssp. *oleifera*. But, in *B. napus* 6 to 7 QTLs involved in seed oil content have been detected (Gül et al., 2003; Qiu et al., 2006; Zhao et al., 2006). Based on previous work in *B. napus* (Badani et al., 2006; Rakow et al., 1999), we expected a negative correlation between seed colour and oil content, with dark seeded genotypes having lower oil contents. Instead we found a strong positive correlation, but only when examining the six "lighter" seed colour genotypic clusters and not when the whole F2 population was considered. The reason for this is probably that seed colour is a fully dominant trait and thus the dark-seeded genotypes (classes 7-13) are mainly a mix of genotypes homozygous for the L58 allele at A9, or heterozygous, with no bias for darker seeds to be homozygous (Figure 5). This will be different for oil content or the other traits associated with the A9 QTL, which are co-dominant, thus disturbing the correlations. Only when seeds are lighter and largely homozygous for the R-o-18 allele at the A9 QTL, the correlation can be detected. Another reason why the correlation was not detected could be the presence of additional, though undetected, loci with opposite effects that independently control seed oil content. Since the A9 QTL still explains only 20 % of the variance for oil content, compared to 52% for seed colour, there are likely additional loci controlling oil content that went undetected in this F2 population.

The opposite correlation we found between seed colour and oil content, compared to Badani et al. (2006) and Rakow et al. (1999), also suggests that it may not always be advantageous to select for yellow-seededness when breeding for high seed oil content in Brassicas. Seed colour and oil content may however not be controlled by one gene, but by two closely linked genes for which favourable alleles happen to be in coupling phase in most genotypes, but by a rare recombination in the ancestry of R-o-18 happened to end up in repulsion phase when they became fixed in the doubled haploid R-o-18. Ahmed and Zuberi (1971) previously also described *B. rapa* varieties with reddish brown seeds that produced more oil than yellow seed coat varieties, suggesting that similar genotypes occur more frequently. In the F2 population described here, oil yield per plant is mostly depending on seed size, number of siliques and

number of seeds per silique. Since silique number is easy to score, this could be a good morphological marker for oil content improvement of *B. rapa* instead of yellow seed coat colour. In a segregating DH population of *B. napus* oil content showed the strongest correlation to seeds per silique (Zhao et al., 2006), which is correlated with silique number in our population. Still, the possibility remains that also seed size, the number of seeds per silique and the number of siliques per plant are not pleiotropic effects of the same genetic locus but caused by two or more closely linked loci.

Finding markers related to adventitious and transformed root formation is useful to improve regeneration and transformation ability of *B. rapa*. The AD and HR QTLs on A10 are controlling adventitious and transgenic root formation respectively. The lines that are sensitive to *A. rhizogenes* are also good candidates for high sensitivity for stable transformation by *A. tumefaciens*. In addition, hairy root cultures are an efficient means of producing secondary metabolites that are normally biosynthesized in roots (Hu and Du., 2006). Also in *B. oleracea* the QTLs controlling these two traits mapped to the same position within the genome (Oldacres et al., 2005), when they used the same explants for the study.

In conclusion, the F2 population we developed shows substantial variation to map traits of agronomic interest and an immortal recombinant inbred line population of this cross will be an interesting tool for genetic analysis of such traits and identification of genetic loci that can be used in future breeding programs. Detailed analysis involving the generation of near isogenic lines and further fine-mapping should allow more insight in the relationships between seed properties.

Materials and methods

Plant growth and generation of the F2 population

Two reciprocal F2 populations were made by crossing *B. rapa* genotypes L58 and R-o-18. For each population 200 F2 seeds were used as starting material. The seeds of L58 (*B. rapa* ssp. *parachinensis*) were provided by Dr. Xiaowu Wang from the Institute for Vegetables and Flowers of the Chinese Academy of Agricultural Sciences, Beijing, China and seeds of R-o-18 (*B. rapa* var. *trilocularis*) were obtained from Dr. Lars Østergaard, John Innes Centre, Norwich, UK.

Individual plants were grown in October 2006 in separate pots in a temperature-controlled greenhouse with artificial day length extension to 16 hours. After about four weeks the first lines started to flower. The inflorescences were covered with plastic bags to prevent cross-pollination. In case of poor seed set, hand pollinations were performed.

Trait measurement

The F2 population derived from the cross L58 (♀) × R-o-18 (♂) was used to determine the phenotypes for 16 traits: flowering time, plant height until first open flower, leaf number until first open flower, branch number, total leaf number on main axis, total plant height at ripening stage, silique number per plant, silique length, number of seeds per silique, seed weight, seed size, seed colour, seed oil content, total seed protein content, adventitious root formations and transgenic hairy root formation.

Seed size was measured using the ImageJ 1.390 software (http://rsb.info.nih.gov). Seed colour of fully mature F3 seeds was scored by eye and ranked into 13 different classes ranging from yellow (1) to black (13) (Figure 2). Silique length was determined as the average length of three ripe siliques. The mean number of seeds per silique was determined by averaging three ripe siliques. Seed oil was extracted by grinding 10 weighed seeds of each line in 650 µl hexane, shaking the mix for two min. followed by one min. of centrifugation at 14.000 rpm in an Eppendorf microfuge. 600 µl of supernatant was transferred to a new tube and left in the fume hood overnight to evaporate the hexane. The oil content was determined in mg oil per mg seed. The seed remains left after oil extraction were used for total protein measurement, using the Bradford essay as described by Goossens et al. (1999).

Hairy root transformation

F3 lines of the cross L58 (♀) × R-o-18 (♂) were used for adventitious and hairy root formation. Adventitious root formation was done on 5-day-old cotyledons and transformation was done on 5-day-old seedlings. For each experiment 10 seeds were used. Seeds were surface sterilized with 70% ethanol (v/v) for 30 seconds, followed by agitation for 5 min in sodium hypochlorite (2.0% active chlorite). After three rinses in sterile distilled water, seeds

were placed in 15×90 mm petri dishes, each containing 20–25 ml half strength salts and vitamins MS medium, without sucrose, solidified with 0.8% (w/v) agar. Petri dishes were placed vertically in a growth chamber maintained at 25˚C with a 16h light/8h dark photoperiod at a light intensity of 60 $mEm^{-2}s^{-1}$.

A. rhizogenes MSU440 were used for hairy root transformation. The pRedRoot binary vector (Limpens et al., 2004) contains the *DsRed1* gene, coding for the red fluorescent protein as a non-destructive selectable marker. It was electroporated into *A. rhizogenes* MSU440. Transgenic cells were cultivated on LB-agar plates supplemented with 50 mg/l kanamycin for 2 d at 27 °C. Roots of 5-day-old seedlings were cut with a scalpel and inoculated with a pellet of *A. rhizogenes* carrying pRedRoot placed at the cutting surface as explained by Limpens et al. (2004). Infected seedlings were transferred to new 9 cm Petri dishes containing half strength salts and vitamins MS medium, without sucrose, solidified with 0.8% (w/v) agar and a half-round filter paper (5 seedlings per plate). The Petri dishes were not completely closed by parafilm to enable aeration. Petri dishes were placed vertically in a growth chamber maintained at 25˚C with a 16h light/8h dark photoperiod at a light intensity of 60 $mEm^{-2}s^{-1}$. After 2 days co-cultivation seedlings were transferred to medium containing 200 mg/l tricarcillin (Duchefa, NL) to kill the *Agrobacterium*. The number of adventitious and transgenic roots was scored after 10 days on medium. Transformed roots were scored using a stereo fluorescence microscope (Figure 6).

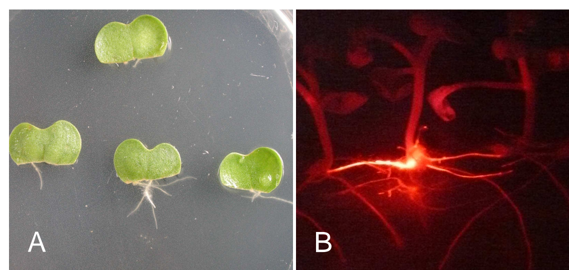


***Figure 6** Adventitious and transformed roots. A: adventitious root formation on cotyledons of B. rapa lines on hormone-free medium. B: B. rapa seedlings with A. rhizogenes carrying pRedRoot show transformed hairy roots expressing DsRed1 using a stereo fluorescence microscope.*

DNA extraction and molecular analysis

Total DNA was extracted from frozen leaves or flower buds as described by Van der Beek et al. (1992) and the AFLP and SSR procedure was performed as described by Vos et al. (1995) and Choi et al. (2007). Pre-amplification and selective amplification were carried out as described by Zhao et al. (2005). For selective amplification seven combinations of EM (*Eco*RI/*Mse*I) (E34M15, E34M16, E37M32, E37M49, E37M56, E40M38, and E40M51) and

four combinations of PM (P*st*I/*Mse*I) (P23M48, P23M50, P21M47and P23M47) primers were used. The *Pst* I and *Eco*RI primers were labelled with IRD-700 at their 5' ends (Zhao et al., 2005). The reaction product of selective amplification was mixed with an equal volume of formamide-loading buffer, denatured for 5 minutes at 94° C, cooled on ice and run on a 5.5% denaturing polyacrylamide gel using a LI-COR (Lincoln, Neb) 4200 DNA sequencer (Myburg et al., 2001). The AFLP gel images were analyzed by the AFLP-Quantar Pro software. All distinguishable bands ranging from 50 bp to 500 bp were used in the data analysis. The AFLP bands were scored as 1 or 0 for presence or absence of the band respectively. All weak and ambiguous bands were scored as "unknown". In addition, 36 public SSR primer pairs (Choi et al., 2007) were used to screen for polymorphisms using the same LI-COR system to run a 5.5% denaturing polyacrylamide gel .

Construction of a genetic linkage map and QTL analysis

Linkage analysis and map construction were carried out using the program JoinMap4 (http://www.kyazma.nl). All markers were grouped using increasing LOD scores (ranging from 5 to 15) to identify ten linkage groups. The regression mapping algorithm was used for linkage analysis. Recombination frequencies were converted to centiMorgan distances using the Kosambi mapping function. MAPQTL 5.0 (http://www.kyazma.nl) was used for QTL analysis. First, an interval mapping test was performed to find putative QTLs. A permutation test with 1000 repetitions was applied to determine the LOD thresholds (at p=0.05). A LOD-score of 2.6 was used as a significance threshold for the presence of a candidate QTL. Multiple-QTL model (MQM) mapping also was performed to locate QTLs after the selection of cofactors. Linkage maps along with QTLs were visualized using Mapchart (Voorrips 2002).

Acknowledgements

This work was financially supported by a personal grant to Hedayat Bagheri from the Ministry of Science, Research and Technology of Iran and the IOP Genomics project IGE050010 on Brassica Vegetable Nutrigenomics. We thank Xiaowu Wang and Lars Østergaard for providing seeds, Xiaowu Wang and Wu Jian for their comments on the manuscript, Joost Keurentjes for his useful advice on data analysis and Corrie Hanhart for her technical assistance.

References

Ahmed S.U., Zuberi M.I. (1971). Inheritance of seed coat colour in *Brassica campestris* L., variety Toria. Crop Sci. 11:309-310.

Ajisaka H., Kuginuki Y., Hida K., Enomoto S., Hirai M. (1995). A Linkage map of DNA markers in *Brassica campestris*. Breed Sci. 45:195-197.

Akiyoshi D.E., Morris R.O., Hinz R., Mischke B.S., Kosuge T., Garfinkel D.J., Gordon M.P., Nester E.W. (1983). Cytokinin/auxin balance in crown gall tumors is regulated by specific loci in the T-DNA. PNAS . 80:407-411.

Alonso-Blanco C., Blankenstijn-De Vries H., Hanhart C.J., Koornneef M. (1999). Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. Proc Natl Acad Sci USA. 96:4710-4717.

Audus L.J. (1959). Plant growth substances. Leonard Hill. London.pp 1-553.

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

Blakesley D., Weston G.D., Hall J.F. (1991). The role of endogenous auxin in root initiation Part I: Evidence from studies on auxin application, and analysis of endogenous levels. Plant Growth Reg. 10:341-353.

Chai Y.R., Lei B., Huang H.L., Li J.N., Yin J.M., Tang Z.L., Wang R., Chen L. (2009). *TRANSPARENTTESTA12* genes from Brassica napus and parental species: cloning, evolution, and differential involvement in yellow seed trait. Mol Genet Genomics**.** 281:109-123.

Chen B.Y., Heneen W.K. (1992). Inheritance of seed colour in *Brassica campestris* L. and breeding for yellow-seeded *B. napus* L. Euphytica. 59:157-163.

Choi S.R., Teakle G.R., Plaha P., Kim J.H., Allender C.J., Beynon E., Piao Z.Y., Soengas P., Han T.H., King G.J., Barker G.C., Hand P., Lydiate D.J., Batley J., Edwards D., Koo D.H., Bang J.W., Park B.S., Lim Y.P. (2007). The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. Theor Appl Genet. 115:777-792.

Christey M.C. (2001). Use of Ri-mediated transformation for production of transgenic plants. In Vitro Cell Dev-Pl. 37:687-700.

Chyi Y-S., Hoenecke M.E., Sernyk J.L. (1992). A genetic linkage map of restriction fragment length polymorphism loci for *Brassica rapa* (syn. *campestris*). Genome. 35:746-757.

Cogan I., Lynn R., King J., Kearsey J., Newbury J., Puddephat J. (2002). Identification of genetic factors controlling the efficiency of *Agrobacterium rhizogenes*-mediated transformation in *Brassica oleracea* by QTL analysis. Theor Appl Genet. 105:568-576.

Debeaujon I., Léon-Kloosterziel K.M., Koornneef M. (2000). Influence of the testa on seed dormancy, germination, and longevity in *Arabidopsis*. Plant Physiol. 122:403-413.

Debeaujon I., Peeters A.J.M., Karen M., Léon K., Koornneef M. (2001). The TRANSPARENT TESTA12 gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. Plant Cell. 13:853-872.

Gelvin S.B. (2000). *Agrobacterium* and plant genes involved in T-DNA transfer and integration. Annu Rev Plant Physiol Plant Mol Biol. 51:223-256.

Gül M., Becker H.C., Ecke W. (2003). QTL mapping and analysis of QTL × nitrogen interactions for protein and oil contents in *Brassica napus* L. Proceedings of the 11th international rapeseed congress. pp:91-93.

Hu Z-B., Du M. (2006). Hairy root and its application in plant genetic engineering. J Int Plant Biol. 48:121-127.

Kim J.S., Chung T.Y., King G.J., Jin M., Yang T-J., Jin Y-M., Kim H-I., Park B-S. (2006). A sequence-tagged linkage map of *Brassica rapa*. Genetics. 174:29-39.

Kole C., Kole P., Vogelzang R., Osborn T.C. (1997). Genetic linkage map of a *Brassica rapa* recombinant inbred population. J Hered. 88:553-557.

Koornneef M., Bade J., Hanhart C., Horsman K., Schel J., Soppe W., Verkerk R., Zabel P. (1993). Characterization and mapping of a gene controlling shoot regeneration in tomato. Plant J. 3:131-141.

Koornneef M., Blankestijn-de V.H., Hanhart C., Soppe W., Peeters T. (1994). The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg *erecta* wild-type. Plant J. 6:911-919.

Lepiniec L., Debeaujon I., Routaboul J.M., Baudry A., Pourcel L., Nesi N., Caboche M. (2006). Genetics and biochemistry of seed flavonoids. Annual review of plant biology. 57:405-430.

Lim Y.P., Kim J.H., Bang J.W., Nam H.G., Cho K.W., Jang C.S. (1998). Genetic mapping of Chinese cabbage (*Brassica rapa* L. var. *pekinensis*).Plant & Animal Genome Conf. pp: 221.

Limpens E., Ramos J., Franken C., Raz V., Compaan B., Franssen H., Bisseling T., Geurts R. (2004). RNA interference in *Agrobacterium rhizogenes*-transformed roots of Arabidopsis and *Medicago truncatula*. J Exp Bot. 55:983-992.

Lou P., Zhao J., Kim J., Shen S., Pino Del Carpio D., Song X., 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.

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

Lu G., Cao J.S., Chen H. (2002). Genetic linkage map of *Brassica campestris* L. using AFLP and RAPD markers. J Zhejiang Univ (Sci). 3:600-605.

Marinova K., Pourcel L., Weder B., Schwarz M., Barron D., Routaboul J.M., Debeaujon I., Klein M.(2007). The *Arabidopsis* MATE transporter TT12 acts as a vacuolar favonoid/H+-antiporter active in proanthocyanidin-accumulating cells of the seed coat. Plant Cell. 19:2023-2038.

Matsumoto E., Yasui C., Ohi M., Tsukada M. (1998). Linkage analysis of RFLP markers for clubroot resistance and pigmentation in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Euphytica. 104:79-86.

McCallum C.M., Comai L., Greene E.A., Henikoff S. (2000). Targeted screening for induced mutations. Nat Biotechnol. 18:455-7.

McGrath J.M., Quiros C.F. (1991). Inheritance of isozyme and RFLP markers in *Brassica campestris* and comparison with *B. oleracea*. Theor Appl Genet. 82:668-673.

Michaels S.D., Amasino R.M. (1999). FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell .11:949-956.

Myburg A.A., Remington D.L., O'Malley D.M., Sederoff R.R., Whetten R.W. (2001). Protocol for high-throughput AFLP analysis using LI-COR IR2 automated sequencers. Biotechniques. 30:348-352, 354, 356-357.

Mysore K.S., Nam J., Gelvin S.B. (2000). An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. Proc Nat Acad Sci.97:948-953.

Narashimhulu S.B., Chopra V.L. (1988). Species specific shoot regeneration response of cotyledons of *Brassicas*. Plant Cell Rep. 7:104-106.

Novakova B., Salava J., Lydiate D. (1996). Construction of a genetic linkage map for *Brassica campestris* L. (syn. *Brassica rapa* L.). Genetika Slechteni. 32:249-256.

Oldacres A.M., Newbury H.J., Puddephat I.J. (2005). QTLs controlling the production of transgenic and adventitious roots in *Brassica oleracea* following treatment with *Agrobacterium rhizogenes*. Theor Appl Genet. 111:479-88.

Puddephat I.J., Riggs T.J., Fenning T.M. (1996). Transformation of *Brassica oleracea* L: a critical review. Mol Breed. 2:185-210.

Qiu D., Morgan C., Shi J., Long Y., Liu J., Li R., Zhuang X., Wang Y., Tan X., Dietrich E., Weihmann T., Everett C., Vanstraelen S., Beckett P., Fraser F., Trick M., Barnes S., Wilmer J., Schmidt R., Li J., Li D., Meng J., Bancroft I. (2006). A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. Theor Appl Genet .114:67-80.

Quijada P., Wang X.W., Cao J.S., Hirai M., Kole C. (2007). Genome Mapping and Molecular Breeding in Plants, Springer Berlin Heidelberg. 2:211-263.

Rakow G., Relf-Eckstein J., Raney P., Gugel R. (1999). Development of high yielding, disease resistant, yellow-seeded *Brassica napus*. 10th International Rapeseed Congress. pp: 67.

Rusholme R.L., Higgins E.E., Walsh J.A., Lydiate D.J. (2007). Genetic control of broad-spectrum resistance to turnip mosaic virus in *Brassica rapa* (Chinese cabbage). J Gen Virol. 88:3177-3186.

Schranz M.E., Quijada P., Sung S.B., Lukens L., Amasino R., Osborn T.C. (2002). Characterization and effects of the replicated flowering time gene *FLC* in *Brassica rapa*. Genetics. 162:1457-1468.

Schwetka A. (1982). Inheritance of seed colour in turnip rape (*Brassica campestris* L.). Theor Appl Genet. 62:161-169.

Shindo C., Aranzana M.J., Lister C., Baxter C., Nicholls C., Nordborg M., Dean C. (2005). Role of FRIGIDA and FLC in determining variation in flowering time of *Arabidopsis thaliana*. Plant Physiol. 138:1163-1173.

Smith E.F., Townsend C.O. (1907). A plant-tumor of bacterial origin. Science. 25:671-673.

Soengas P., Hand P., Vicente J.G., Pole J.M., Pink D.A.C. (2007). Identification of quantitative trait loci for resistance to *Xanthomonas campestris* pv. *campestris* in *Brassica rapa*. Theor Appl Genet. 114:637-45.

Song K.M., Suzuki J.Y., Slocum M.K., Williams P.H., Osborn T.C. (1991). A linkage map of *Brassica rapa* (syn. campestris) based on restriction fragment length polymorphism loci. Theor Appl Genet. 82:296-304.

Stringam G.R., McGregor D.I., Pawlowski S.H. (1974). Chemical and morphological characteristics associated with seed coat colour in rapeseed. Proc.4th Int. Rapeseed Congr. pp: 99-108.

Suwabe K., Iketani H., Nunome T., Ohyama A., Hirai M., Fukuoka H. (2004). Characteristics of

microsatellites in *Brassica rapa* genome and their potential utilisation for comparative genomics in cruciferae. Breed Res. 54:85-90.

Suwabe K., Tsukazaki H., Iketani H., Hatakeyama K., Fujimura M., Konodo M., Nunome T.,Fukuoka H., Hirai M., Matsumoto S. (2004). Joint meeting of the 14th Crucifer Genetics Workshop and the 4th ISHS Symposium on Brassicas. pp:143.

Tang Z.L., Li J.N., Zhang X.K., Chen L., Wang R. (1997). Genetic variation of yellow-seeded rapeseed lines (*Brassica napus* L.) from different genetic sources. Plant Breeding. 116:471- 474.

Tanhuanpää P.K., Vilkki J.P., Vilkki H.J. (2004). Mapping of a QTL for oleic acid concentration in spring turnip rape (*Brassica rapa* ssp.*oleifera*). Theor Appl Genet. 92:952-956.

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

U N. (1935). Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jpn J Bot. 7:389-452.

Van der Beek J.G., Verkerk R., Zabel P., Lindhout P. (1992). Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: cf9 (resistance to *Cladosporium fulvum*) on chromosome 1. Theor Appl Genet. 84:106-112.

Van Deynze A.E., Beversdorf W.D., 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., Pauls K.P. (1994). The inheritance of seed colour and vernalization requirement in *Brassica napus* using doubled haploid populations. Euphytica. 74:77-83.

Vaughan J.G., Phelan J.R., Denford K.E. (1976). Inheritance of seed coat colour in *Brassica juncea*. Can J Plant Sci. 59:635-637.

Voorrips R.E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 93:77-8.

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

Wang M., Zhang F.L., Meng X.D., Liu X.C., Zhao X.Y., Fan Z.C. (2004). A linkage map construction for Chinese cabbage based on AFLP markers using DH population. Acta Agric Boreali-Sin. 19:28-33.

Wu J., Schat H., Sun R., Koornneef M., Wang X., Aarts M.G.M. (2007). Characterization of natural variation for zinc, iron and manganese accumulation and zinc exposure response in *Brassica rapa* L. Plant Soil. 291:167-180.

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

Yu S.C., Wang Y.J., Zheng X.Y. (2003). Mapping and analysis QTL controlling heat tolerance in *Brassica campestris* L. ssp. *pekinensis*. Acta Hort Sin. 30:417-420.

Zhang L.G., Wang M., Chen H. (2000). Construction of RAPDs molecular genetic map of Chinese cabbage. Acta Bot Sin. 42:485-489.

Zhao J., Becker H.C., Zhang D., Zhang Y., Ecke W. (2006). Conditional QTL mapping of oil content in rapeseed with respect to protein content and traits related to plant development and grain yield. Theor Appl Genet. 113:33-38.

Zhao J., Wang X., Deng B., Lou P., Wu J., Sun R., Xu Z., Vromans J., Koornneef M., Bonnema G. (2005). Genetic relationships within *Brassica rapa* as inferred from AFLP fingerprints. Theor Appl Genet. 110:1301-1314.

Chapter 4

Genetic mapping and QTL analysis of plant morphology and seed related traits in a new *Brassica rapa* recombinant inbred line population

Hedayat Bagheri[1], Inge van Oorschot[1], Corrie Hanhart[1], Tanja Jansen-van den Bosch[2], Rolf Mank[2], Joost Keurentjes[1], Maarten Koornneef[1,3] and Mark G.M. Aarts[1]

[1]Laboratory of Genetics, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; [2]Keygene N.V., Agro Business Park 90, 6708 PW Wageningen, The Netherlands; [3]Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, 50829 , Cologne, Germany.

Abstract

A recombinant inbred population (F7) of 160 lines was made by repeated selfing of F2 plants derived from a cross between genotype L58, which is a rapid-flowering, self-compatible Cai Xin line, and R-o-18, a self-compatible Indian doubled haploid spring oil line. The Complexity Reduction of Polymorphic Sequences (CRoPS®) technology was applied for single-nucleotide polymorphism (SNP) detection. The recently developed Illumina® BeadXpress™ platform combined with the GoldenGate assay was used for genotyping of the RIL population. A linkage map was constructed for this population using 86 SNP markers and 6 publicly available SSRs, covering a total map distance of 400 cM with an average resolution of 4.3 cM. A total of 26 QTLs for 19 traits, including flowering time, total plant height, plant height until first open flower, leaf number until first open flower, total leaf number, cuticular wax appearance, branch number, number of siliques, silique length, silique beak length, silique carpel number, seed weight, seed colour, seed shattering, and seed ripening time were detected. This study shows that the developed *B. rapa* RIL population can successfully be used to study the genetic basis of many breeding related traits.

® *The AFLP®, KeyGene®, CRoPS® technologies are covered by patents and/or patent applications owned by Keygene N.V. AFLP, KeyGene and CRoPS, are (registered) trademarks of Keygene N.V.*

Introduction

With the smallest genome size in the Brassica genus and the rapid life cycle of some of its genotypes, *B. rapa* can be a model crop for genetic analysis in this family. For genetic analysis, so-called immortal mapping populations are very useful in the quantitative genetic analysis to identify loci that impact traits of interest. Immortal populations are characterized by their homozygous state, allowing propagation through seeds while maintaining the genotype of the parents in the progeny. In this study two distinct morphotypes of *B. rapa*, Cai Xin L58 and Yellow Sarson R-o-18, were used to develop a Recombinant Inbred Line (RIL) population. Recombinant Inbred Line populations are made by repeated propagation of progeny of an F1 plant over many generations by single seed descent until each line has a high degree of homozygosity, typically until F7/F8 (Koornneef et al., 2004). As both parents are early flowering and self-compatible, the developed RIL population has a rapid propagation and is easy to maintain through single seed descent. Doubled haploid (DH) populations are also immortal. They are made by induction of haploid embryos through culture of microspores from an F1 plant, followed by spontaneous genome doubling of regenerated plantlets (Sebastian et al., 2000). Although DH production is common for *Brassica* species (Pink et al., 2008) it is not easy for *B. rapa* as most genotypes respond very poorly to DH induction (Kole et al., 1997). In addition, normally a higher degree of segregation distortion is found in DH population compared to sexually propagated populations due to the preferential selection of genotypes responsive to microspore or anther culture (Voorrips et al., 1997; Suwabe et al., 2004).

For genetic mapping, the highly abundant single-nucleotide polymorphisms (SNPs) are useful genetic markers. SNPs are the most frequent type of genetic polymorphism and can be readily detected using appropriate genotyping platforms. However in unsequenced crops with duplicated genome sequences the efficiency of SNP discovery is low if not hampered (Li et al., 2004). The AFLP-based Complexity Reduction of Polymorphic Sequences (CRoPS) technology, as a novel approach for large-scale polymorphism discovery in complex genomes, can be used in crops with a complex genome (van Orsouw et al., 2007). Recently the Golden Gate assay has been implemented for analysis on the Illumina® BeadXpress™ platform. The GoldenGate assay was a very efficient tool for high-throughput genotyping of polyploid wheat (Akhunov et al., 2009). In the Illumina Golden Gate assay, cyanine-3 (Cy3) and cyanine-5 (Cy5) labeled PCR products, containing different address sequences are hybridized to small glass rods (VeraCode "beads") that are coupled to oligos complementary to the Illumicodes. The VeraCode beads have etched bar codes that correspond to a specific Illumicode address sequence. Once the beads enter the reader, a laser scans the beads and detects the bar code on each bead as well as the Cy3 or Cy5 fluorescent signal. Each bar code corresponds to a specific Illumicode address which corresponds to a specific genomic

sequence provided in the SNP design file. The ratio of the intensity of the Cy3 and Cy5 fluorescence is used to determine the allelic state (genotype) at the SNP.

Several seed characteristics, such as seed number, the number of siliques per unit area, branch number, and seed size, determine the value of crops grown for seeds like oilseed rape (Bekkaoui and Wilf, 2003; Quijada, 2007). The seed of *B. rapa* is mainly used for oil production, although the meal remaining after oil extraction is also of economic interest. A yellow seed coat colour is an important trait in Brassica oilseed crops as it is generally associated with thinner seed coat and lower cellulose content, improving the quality of the meal (Tang et al., 1997; Badani et al., 2006). In chapter 3 we reported on a strong QTL explaining more than 50% of seed colour variance (named SC), which mapped to chromosome A9 in the F2 population of the same cross from which the RIL population that is studied here is made.

Some morphological related traits such as shattering, carpel number (trilocularis), plant architecture, leaf glossiness (cuticular wax) and vivipary are of interest in plant breeding. Plant architecture is of major agronomic importance, strongly influencing the suitability of a plant for cultivation, its yield and the harvest efficiency (Reinhardt and Kuhlemeier, 2002). Height of the plant at first open flower and leaf number at first open flower are factors that contribute to Brassica plant architecture. Brassica vegetables and oil types generally differ considerably in height. Pod shattering is an undesirable character in crop breeding. Grain loss during harvesting due to shattering decreases seed yield. A loss of 20% of the seed yield has been reported for *B. napus* (Price et al., 1996). Shattering is caused by carpel abscission. Pod shatter resistance is a complex trait, partly recessive and difficult to assess because it is only seen at maturity (Morgan et al., 2003). As far as we are aware there is no report about Brassica loci controlling shattering.

The outer surfaces of *Brassica* leaves are coated with cuticular waxes which protect the plant from a variety of environmental pressures. Genotypic control of wax formation has been studied in many plants; in most cases the normal or glaucous form is dominant and the glossy or glabrous form recessive. The difference in appearance has been variously associated with the quantity, chemical composition or structure of the epicuticular wax (Baker 1974). Glossy leaves often have a reduced wax load. The wax is involved mostly in reduction of water loss and in defense against pathogens (Teece et al., 2008). Wax accumulation also potentially contributes to drought adaptation (Jefferson 1994). Epicuticular wax is hydrophobic and composed of long-chain fatty acids, alcohols, aldehydes and very long-chain wax esters (Wen and Jetter, 2009). Arabidopsis *eceriferum (cer)* mutants defective in normal epicuticular wax deposition have been identified by their glossy appearance (Koornneef et al. 1989; Beittenmiller 1996).

Vivipary is the phenomenon of seeds germinating on the mother plant before harvest. In

cereals, but also in oilseed rape, this can lead to a large economic loss, as it causes severe damage to grain quality. Resistance to vivipary is therefore a very important trait in breeding programs (Zhang et al., 2008). Vivipary is caused by the absence of dormancy, which normally prevents seeds to germinate prematurely. Too much dormancy, causing the inability to germinate, to germinate with a delay or to germinate non-uniformly, also has a negative affect on seed quality (Millar et al., 2006). Genetic and physiological studies showed that abscisic acid (ABA) is a key player in establishing and maintaining seed dormancy (Koornneef et al. 2002). Also in *B. napus* vivipary is found, causing decreased seed viability and vigor (Ruan et al., 2000). Recently a major QTL explaining 50.78% of the total phenotypic variance for vivpary has been mapped to chromosome N11 of *B. napus* (Feng et al., 2009).

Brassica rapa is known as one of the most recalcitrant members of *Brassica* genus to regenerate shoots in vitro (Narashimhulu and Chopra 1988) and routine high frequency transformation of *B. rapa* by *A. tumefaciens* has been hampered by the lack of efficient plant transformation and regeneration techniques. There is still a need to develop efficient transformation methods that can be used for many genotypes.

In this report we describe a new RIL mapping population, derived from a cross between two distinct morphotypes, Cai Xin line L58 and Yellow Sarson line R-o-18. A new high-throughput genotyping platform was designed for *Brassica* and used for genotyping the RIL population. The RIL population was scored for several plant morphology, seed-related and root formation traits, which led to the identification of several QTLs.

Results

Genotyping and the construction of a linkage map

Genotyping of 160 *B. rapa* lines was performed using the Illumina BeadXpress Reader to analyse a GoldenGate assay developed for *Brassica* SNP detection. Of the 384 SNPs present in the GoldenGate assay, 120 were polymorphic, 222 markers gave 100% homozygous or heterozygous scores and 42 SNPs did not generate a product or could not be scored unambiguously. A linkage map was constructed for the L58 × R-o-18 F7 RIL population using 86 polymorphic SNP markers and 6 publicly available SSR markers, covering a total map distance of 399.8 cM with an average resolution of 4.34 cM between markers (Figure 1). Biased location of markers makes grouping difficult in JoinMap and extra linkage groups appear. The SSR markers were included in the mapping to allow the assignment of linkage groups to their respective chromosomes as defined by the *B. rapa* reference linkage map (Choi et al., 2007). Six linkage groups (A2, A4, A5, A6, A9 and A10) correspond to defined *B. rapa* chromosomes (Choi et al., 2007). Linkage groups G1, G2, G3, G4 and G5 could not yet be assigned to a chromosome. *B. rapa* has 10 chromosomes, suggesting that two linkage groups corresponde to one chromosome.

The RIL map is shorter than the F2 map of the same cross (Chapter 3) and another *B. rapa* RIL map (Kole et al., 1997). Based on the chi-square test for goodness-of-fit to the expected 1:1 Mendelian segregation ratio, 28% of the markers showed segregation distortion. For highly skewed loci (9.7% of the markers showed >3:1 ratio) the R-o-18 allele was in excess. Skewed markers were clustered in regions of the A5, A6 and G3 linkage groups.

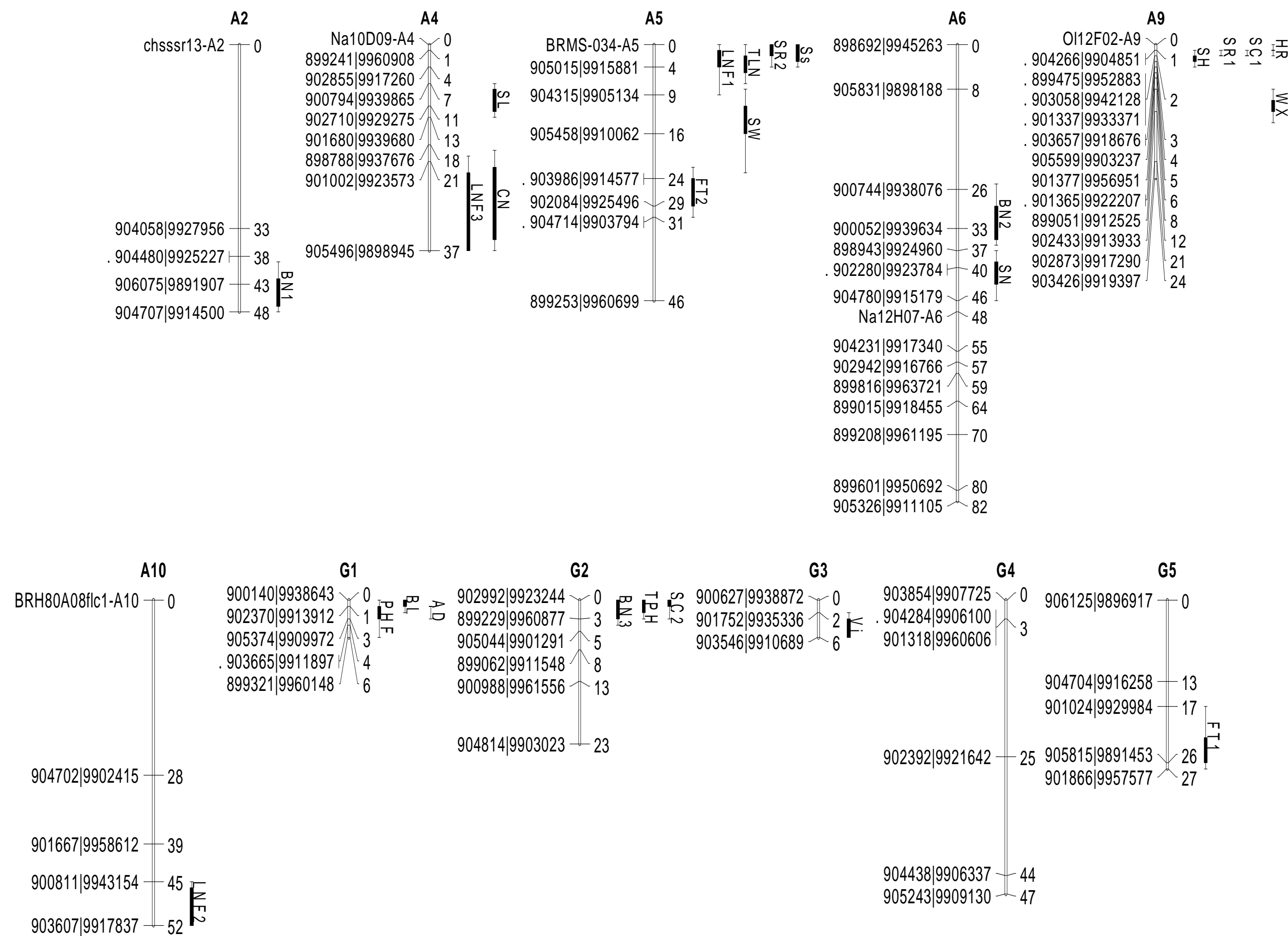
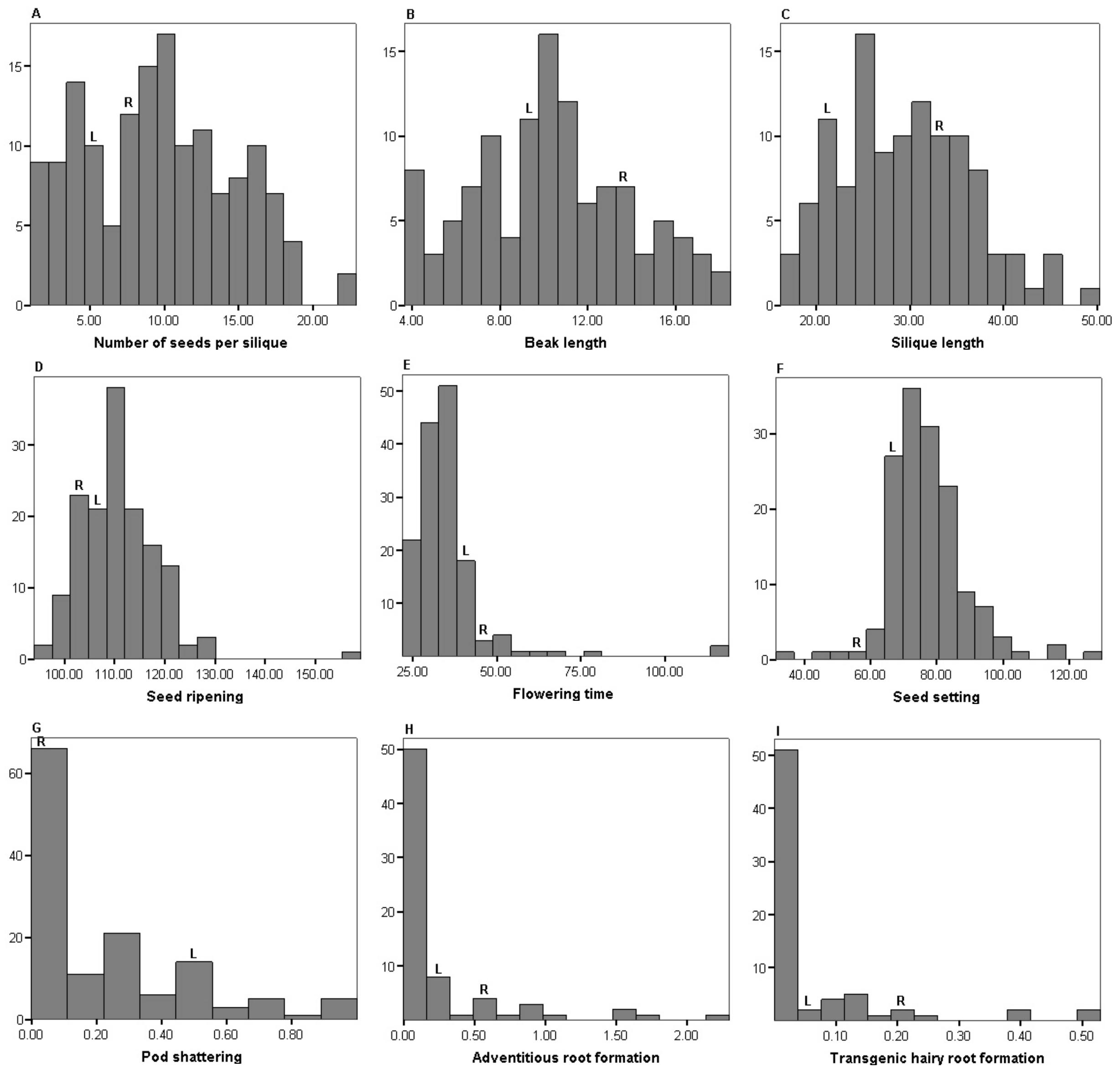


Figure 1 *Genetic linkage map of the B. rapa L58 × R-o-18 RIL population, showing the positions of 86 SNP and 6 SSR markers distributed over 11 linkage groups. Six linkage groups (A2, A4, A5, A6, A9 and A10) correspond to defined chromosomes of the B. rapa reference linkage map (Choi et al., 2007). Linkage groups G1, G2, G3, G4 and G5 could not yet be assigned to a chromosome. QTLs mapping to the linkage map for the determined traits are indicated with boxes. Whiskers represent 1-LOD and 2-LOD confidence intervals (95%), respectively, for significant QTLs. FT: flowering time; TPH: total plant height; PHF: plant height until first open flower; SC: seed coat colour; SN: silique number; SW: seed weight; SH: shattering; WX: wax; SR: seed ripening; Ss: seed setting ; LNF: leaf number until first open flower; BN: branch number; CN: carpel number; SL: silique length; BL: beak length; TLN: total leaf number; Vi: seed vivipary; AD: adventitious root formation; and HR; transgenic hairy root formation.*

Phenotyping the RIL population



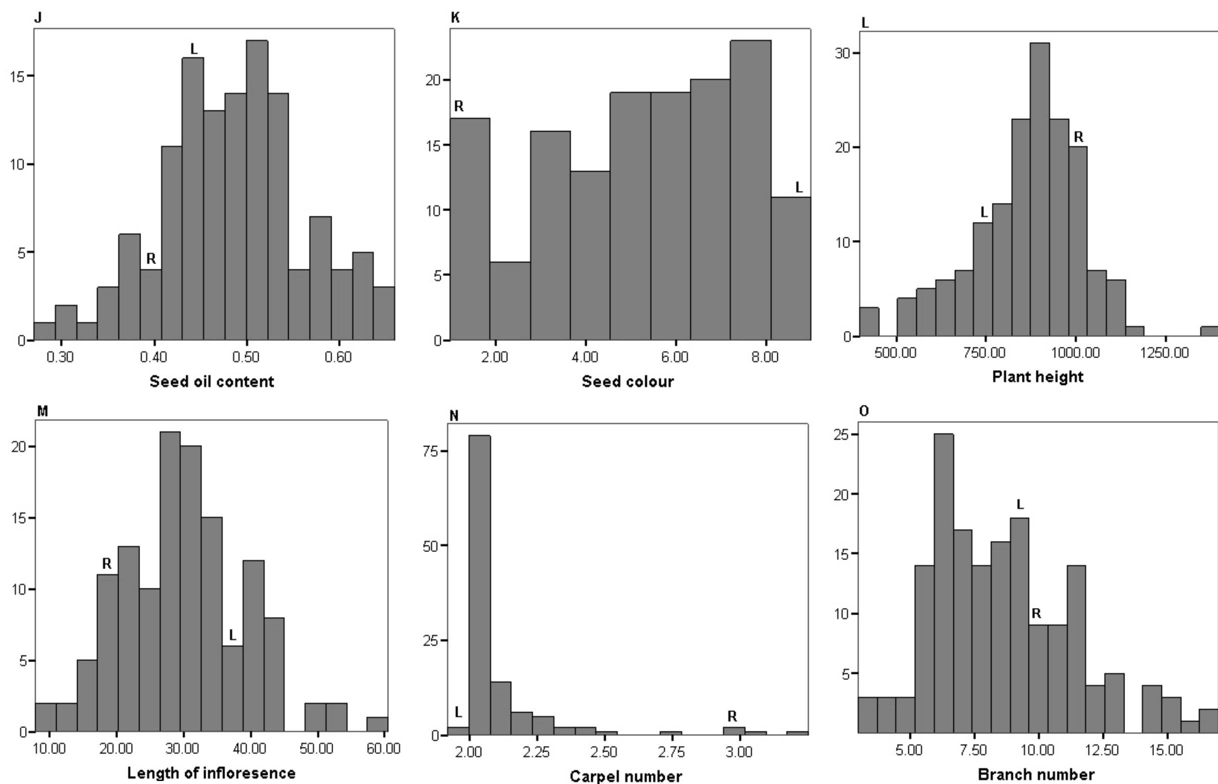


***Figure 2** Frequency distributions of non-normalized data of some traits in the L58 × R-o-18 RIL population. The vertical axis indicates the number of lines per trait value class and the horizontal axis the different trait value classes. (A) number of seeds per silique (B) beak length (mm); (C) silique length (mm); (D) seed ripening (days to ripen); (E) flowering time (days to flower); (F) seed setting (days from flowering until ripening); (G) pod shattering ; (H) adventitious root formation; (I) transgenic hairy root formation; (J) seed oil content; (K) Seed colour (from yellow to black); (L) plant height; (M) length of inflorescence; (N) carpel number; and (O) branch number. The parental values are the mean of five replicates, indicated with L as L58 and R as R-o-18.*

In total 21 traits, consisting of 20 morphological phenotypes and seed oil content, were recorded for the RIL population. Figure 2 shows the frequency distributions of some measured traits including silique length, beak length, seed ripening, flowering time, seed setting, pod shattering, adventitious root formation and transgenic hairy root formation, seed oil content, seed coat colour, plant height, infloresence length, carpel number and branch number. Transgression beyond the parental values was observed for most of the traits except for seed colour.

Correlation analysis of all measured traits showed that flowering time was highly positively correlated with total leaf number and leaf number until first open flower (Table 1). Silique number, seed number per silique, shattering and silique length were positively correlated. Silique length was highly positively correlated with the number of seeds per silique. In

general, plants with more siliques had longer siliques with more seeds and higher seed oil content. Pod shattering was positively correlated with number of seeds per silique, silique number and seed coat colour, but negatively with seed ripening. Branch number was correlated with leaf number and flowering time. Cuticular wax showed a significant correlation with plant height until first open flower. Seed vivipary was positively correlated with seed set and negatively with flowering time. The number of carpels showed a significant positive correlation with seed weight.

***Table 1** Pearson correlation analysis of all measured traits in the L58 × R-o-18 RIL population.*
*FT: flowering time, TPH; total plant height, PHF: plant height until first flower, LNF: leaf number until first flower, WX: wax cuticular, BN: branch number, TLN: total leaf number, SN: silique number, SL: silique length, SH: shattering, BL: beak length, NSPS: number of seed per silique, SW: seed weight, CN: carpel number, Vi: vivipary, Infl: length of infloresence, SR: seed ripening, Ss: seed setting, AD: adventitious root formation, HR: transgenic hairy root formation, SC: seed colour, SO: seed oil content. ** means significant at P≤0.01; * significant at P≤0.05.*

Trait	BL	SL	SN	SR	FT	Ss	PHF	LNF	BN	TLN	TPH	InfL	SH	NSPS	SW	CN	WX	Vi	SC	SO	AD
SL	0.41**	1																			
SN	0.19*	0.27**	1																		
SR	-0.17	-0.18	-0.19*	1																	
FT	-0.01	-.21*	0.1	0.1	1																
Ss	-0.1	0.02	-0.21**	0.66**	-0.11	1															
PHF	0.17	0.12	-0.1	-0.12	0.27**	-0.29**	1														
LNF	-0.02	-0.20*	0.06	0.01	0.86**	-0.62**	0.31**	1													
BN	-0.05	-0.13	0.02	-0.16	0.38**	-0.40**	0.35**	0.56**	1												
TLN	-0.02	-0.21*	0.07	0.01	0.88**	-0.64**	0.29**	0.98**	0.56**	1											
TPH	0.08	0.11	-0.08	-0.13	0	-0.08	0.63**	0.01	0.16*	0.01	1										
InfL	0.02	0.03	0.16	-0.02	0.13	-0.16	0	0.05	0.14	0.09	0.28**	1									
SH	0.18	0.11	0.42**	-0.26**	0.06	-0.23**	-0.14	0.03	0.02	0.05	-0.02	0.17	1								
NSPS	0.41**	0.47**	0.55**	-0.22*	0.02	-0.16	-0.11	0.02	-0.07	0.03	0	0.26*	0.53**	1							
SW	0.12	0.05	0.1	-0.05	0.11	-0.13	0.05	0.12	0.09	0.11	0.11	0.02	-0.02	0.01	1						
CN	0	-0.15	-0.02	-0.09	0.06	-0.11	0.17	0.17	0.03	0.15	0.12	-0.01	0.04	0.1	0.27**	1					
WX	0.1	0.08	0.01	-0.05	-0.01	-0.02	0.29**	0.04	-0.03	0.02	0.25**	-0.11	-0.18*	-0.25**	0.02	-0.01	1				
Vi	0.1	0	0.14	0	-0.37**	0.33**	-0.18	-0.39**	-0.26*	-0.39**	0.01	0.07	-0.01	0.04	0.19	-0.14	0.01	1			
SC	0.02	-0.08	0.06	-0.1	0.05	0.02	0.01	-0.12	0	-0.11	0	0.02	0.24**	0.14	-0.07	0.14	-0.23**	-0.09	1		
SO	-0.04	-0.07	0.34**	0	0.28**	-0.20*	0	0.22*	0.07	0.22*	-0.11	-0.04	0.11	0.24**	-0.17	0.02	-0.02	-0.22*	0.18*	1	
AD	0.28**	-0.24	-0.03	0.1	-0.05	0.05	-0.07	-0.05	0	-0.05	-0.09	0	0.22*	0.04	-0.1	-0.2	-0.08	0.27**	0.04	-0.15	1
HR	0.23*	-0.15	-0.14	0.03	-0.04	0	0.06	-0.08	-0.09	-0.07	0.02	-0.07	0	0	-0.02	-0.11	0.1	0.2	-0.15	-0.11	0.12

QTL analysis

In total 26 QTLs were mapped for 19 different traits (Table 2; Figure 1). The strongest QTLs were found for seed colour (SC1 and SC2), mapping to chromosome A9 and linkage group G2 and explaining 55.4% and 14.3% of the seed coat colour variance respectively. The dark coloured allele at SC1 is coming from L58 (black seeded), but the darker allele at SC2 comes from R-o-18 (yellow seeded). SC1 co-located with pod shattering (SH) and seed ripening (SR) QTLs on A9. 16% of the variance for pod shattering was explained by this locus. 14% of the variance for carpel number was explained by a CN QTL on A4. AD and HR QTLs on G1 and A9 were responsible for adventitious and transgenic hairy root formation.

***Table 2** QTLs detected for 19 different traits in the L58 × R-o-18 RIL population. QTLs are numbered according to decreasing LOD score (LOD). %Expl. var. is the percentage of total phenotypic variance explained by individual QTLs. For each of the QTLs the allelic effect is indicated (Effect). These are calculated as μA-μB (μ= mean), where A and B are RILs carrying L58, respectively R-o-18 genotypes at the QTL positions. μA and μB were estimated by MapQTL. Effects are given in days (flowering time), millimeter (plant height, size), milligram per seed (oil content) or without unit (silique or seed number, seed colour).*

Trait	QTL	Linkage group	LOD	Position of peak LOD cM	% Expl. var.	Effect
Silique length	SL	A4	3.81	26.2	14.4	+7.05
Beak length	BL	G1	4.23	0.0	15.8	-2.90
Silique number	SN	A6	3.67	39.7	10.4	+15.91
Shattering	SH	A9	4.83	3.1	16.1	+0.22
Carpel number	CN	A4	3.7	16.0	14.0	-0.18
Seed colour	SC1	A9	32.42	1.76	55.4	+3.70
	SC2	G2	12.02	0.0	14.3	-2.10
Seed weight	SW	A5	3.60	15.9	15.2	-4.33
Seed ripening	SR1	A9	5.00	1.9	13.5	-5.96
	SR2	A5	3.40	0.0	9.1	+5.05
Seed setting	Ss	A5	2.90	0.0	9.0	+8.0
Vivipary	Vi	G3	2.40	5.5	14.0	-0.22
Total plant height	TPH	G2	4.73	0.0	13.0	-130.00
Plant height until first flower	PHF	G1	3.67	3.4	10.5	-100
Total leaf number	TLN	A5	3.10	3.7	9.1	-4.06
Leaf number until first flower	LNF1	A5	3.12	3.7	8.0	-3.85
	LNF2	A10	2.56	0.0	6.6	-3.58
	LNF3	A4	2.21	52.00	5.3	+3.06
Wax	WX	A9	3.80	12.31	11.0	-0.50
Branch number	BN1	A2	4.54	42.26	10.5	+1.86
	BN2	A6	3.93	32.9	9.1	+1.75
	BN3	G2	3.48	0.0	7.9	+1.77
Flowering time	FT1	G5	3.90	26.4	9.8	-5.00
	FT2	A5	3.54	28.9	8.9	-6.00
Adventitious root formation	AD	G1	3.56	1.32	20.4	-0.41
Hairy transgenic root formation	HR	A9	6.00	1.1	33.0	-0.12

Discussion

A segregating recombinant inbred population of *Brassica rapa* consisting of 160 lines was developed from two distinct morphotypes, both early flowering and self compatible, and used for genetic linkage mapping.

CRoPS technology, as a novel approach for large-scale polymorphism discovery in complex genomes, was applied for SNP discovery (van Orsouw et al., 2007). Of about 1300 putative SNPs, 384 SNPs with a high probability of detection in genotyping assay were selected. The GoldenGate genotyping assay as analyzed using the Illumina® BeadXpress^{TM} platform was successfully employed for SNP genotyping of the *B. rapa* RIL population. This genotyping platform has the potential to genotype up to 384 polymorphic sites in 96 individuals in a single reaction. One third of the SNPs present in the GoldenGate assay were polymorphic between parental lines. The linkage map includes 86 SNPs and 6 publicly available SSRs, grouped in 11 linkage groups. The short length of the map could be due to the biased location of markers which makes the grouping difficult using the JoinMap mapping software, which is why more than 10 linkage groups are found. Over 30 markers could not be mapped so far, which suggests that currently part of the genome is not covered by the genetic map. Since not all linkage groups were assigned to a reference chromosome, there still is the possibility that a chromosome is not represented in the linkage map. Other explanations, like suppression of recombination, or strong genetic distortion of markers, are less likely. Although 28% of the loci show significant segregation distortion from the expected 1:1 ratio, only 9.7% of the markers show highly distorted segregation (>3:1). Even though we made an effort not to select during the single seed descent propagation, this is hard to avoid, especially for lines that are at the low fertility, germination or growth side of the spectrum. Generally, environmental and artificial selection over several generations, causing segregation distortion, is common to many mapping populations (Voorrips et al., 1997; Wang et al., 2003; Kianian and Quiros. 1992).

QTL analysis was performed for 21 traits. For several traits there is strong transgression of phenotypes beyond the parental values, often far beyond the parent with the highest value, suggesting the presence of several loci with contrasting allele effects controlling a trait in this population. It makes this population a potentially rich tool for resolving genetic variation for *B. rapa* breeding purposes. In total 26 QTLs for 19 of the studied traits were detected. Seed coat colour is a very important trait in oilseed type *Brassica* crops. Yellow seeds are preferred, since yellow-seededness was found to correspond with higher protein and lower crude fiber contents compared to black-seeded varieties and are of better quality for use as poultry and livestock feed (Tang et al., 1997; Badani et al., 2006). We already showed in the F2 population of *L58* × *R-o-18* that seed coat colour is a maternal trait and the black seed coat

allele is over-dominant over the allele for yellow seed coat. Two strong QTLs for the seed colour (SC1 and SC2) were detected on A9 and G2 explaining about 70% of the seed coat colour variation. The dark allele at SC1 is derived from the dark-seeded L58 parental genotype, but for SC2 the darker allele comes from R-o-18, indicating that there are modifiers in R-o-18 that might be epistatic to the yellow alleles at SC1. The SC1 locus was also detected in the F2 generation of this population (SC in Chapter 3) and had been reported before by Lou et al. (2007). Contrary to what we found in the F2 generation, the QTL for seed oil content on A9 in same region as SC1 did not exceed the LOD threshold level and since no additional QTL were detected, we did not map a seed oil content QTL in the RIL population. For the analysis of seed oil content, both in the F2 as in the RIL population, we used a crude method to extract seed oil. This method can be sensitive to differences in extraction, which could explain for the inability to detect the A9 seed oil content QTL in the analysis of the RIL population. Seed oil has a significant positive correlation with silique number and the number of seeds per silique, which is in agreement with our previous observations for the F2 population (Chapter 3).

The pod shattering QTL (SH) on A9 is located in the same region as SC1. Correlation analysis confirmed that easy shattering lines have black seeds. Therefore selection for yellow-seeded lines might improve shattering resistance when both traits segregate. Shattering has a significant positive correlation with the number of seeds per silique and the silique number. The QTL for number of seeds per silique in this region was not significant. There is little exploitable variation in pod shattering resistance within the *B. napus* gene pool (Morgan et al., 2003). Introducing pod shattering resistance alleles from *B. rapa* into a *B. napus* breeding program could be useful. In Arabidopsis the MADS-box transcription factor *FRUITFULL* (*FUL*) (Gu et al., 1998;) mediates valve development by inhibiting the action of the *SHATTERPROOF* (*SHP*) (Liljegren et al., 2004) valve margin identity genes. Constitutive expression of the *FUL* gene in *B. juncea* is sufficient to inhibit valve margin formation and seed dispersal (Østergaard et al., 2006).

R-o-18 is a *B. rapa* Yellow Sarson type which was previously named *B. trilocularis* (Gómez-Campo, 1999); because many fruits have three or four carpels. The allele of the carpel number (CN) QTL that maps to chromosome A4, explaining 14% of carpel variance, corresponding to higher carpel number, originated from R-o-18. Carpel number showed only a significant positive correlation with seed weight and therefore seed size. As R-o-18 has bigger and heavier seeds than L58, the CN QTL could be directly involved in determining seed size and weight or it may be linked with loci controlling seed size and weight.

A silique length QTL (SL) is located at chromosome A4. This QTL is different from the QTL for beak length (BL) mapped to G1, indicating both are genetically independently controlled. In *B. rapa* DH populations Lou et al. (2007) reported two genomic regions on A1 and A7

affecting silique length and three loci on A5, A7 and A9 controlling the beak length. The SL QTL is a new locus on A4 but like the BL QTL located on G1, which we could not assign to a reference chromosome, it is not clear if this is a new locus too. The number of seeds per silique is highly positively correlated with other silique related traits such as silique length, beak length and silique number. Therefore SL and BL are likely to have an overall effect on silique related traits.

Seed ripening and seed set traits are highly correlated. Co-location of the SR2 and Ss QTLs reveals there is one QTL on top A5 probably active in early stages of seed setting. The second QTL of seed ripening (SR1) appears to control seed ripening at later stages of seed development. There is no report about seed vivipary QTLs in *B. rapa*. We detected one QTL on G3 explaining 14% of preharvest sprouting variance. Correlation analysis revealed a significant negative relation between vivipary and seed oil content. In *B. napus* vivipary also resulted in lower seed oil (Ruan et al., 2000).

We detected one QTL on A9 for cuticular wax appearance (WX). As we only recorded the appearance difference of the leaves and did not analyze the chemical composition of the epicuticular wax, it is not clear if this QTL explains the quantity, chemical composition or structure of the epicuticular wax.

Branch number QTLs are found on A2, A6 and G2. Branch number has no correlation with silique number, but is highly correlated with leaf number. As every branch originates from the bud at the leaf base, this correlation is expected. The outgrowth of buds to branches is strongly controlled by hormonal balance, involving cytokinins, auxins and stringolactones (Gomez-Roldan et al., 2008).

Finding markers related to adventitious and transformed root formation would be useful to select for regeneration and transformation improved *B. rapa* genotypes. The AD and HR QTLs on G1 and A9 are respectively controlling adventitious and transgenic root formation indicating that these are independent traits. In *B. oleracea* hypocotyls it has been shown that performance for transgenic root production is associated with performance for adventitious (non-transgenic) root production (Oldacres et al., 2005), contrary to what we describe here. As we used different tissues for the adventitious and hairy root production, it is possible that these traits are independent. Correlation analysis also did not show a significant correlation between these two traits. AD and HR mapped on A10 in the F2 population (Chapter 3). Sensitive genotypes of *B. oleracea* in adventitious root production tend also to perform well in transgenic roots production induced by *A. rhizogenes* but the presence of the bacterium had no significant effect on adventitious root production (Oldacres et al., 2005). Therefore it would be beneficial to test in the future if the production of non-transgenic roots is enhanced by treatment with *A. rhizogenes*.

The *B. rapa* immortal population we described has two important characteristics which are an advantage for maintaining the population and for several phenotypic screens: The population is fully self–compatible and plants rapidly complete their life cycle. In addition, the population is sufficiently large to have ample statistical power to detect QTLs. The rapid-cycling properties, and the diverse origin of both parents (vegetable and oilseed rape), make this population a potentially very interesting source for genetic analysis of a range of agronomic traits, both for leafy vegetables and for oilseed properties. The diverse origin of the parents ensured there was sufficient genetic variation in the population to map many morphological and seed related QTLs. Some of the QTLs we describe here are reported for the first time. In few cases we could not detect the same QTLs which were detected in the F2 generation. For example, in the F2 we detected two QTLs for flowering time on A2 and A3. In RIL we could not find these, but instead mapped QTLs on G5 and A5. Also the AD and HR QTLs related to adventitious and transgenic root formation, which mapped on A10 in the F2, mapped to G1 and A9 in the RIL population. We do not think this is due to phenotyping or mapping mistakes in either of the two analyses, but simply reflects the contribution of genotype × environment interaction to the phenotypes. Although both F2 and RIL populations grew in the same greenhouse, they grew in very different seasons. Some QTLs may have been lost due to increased segregation distortion due to inadvertent selection during single seed descent over 6 generations. Finally, some differences in QTL detection may be caused by differences in the accuracy of phenotyping both populations, with more experience gained at the time the RIL population was scored. Generally, QTLs with a high LOD score were detected in both populations. Probably due to a higher number of recombination events in the RIL and a better scoring of recombination by using SNPs in the RIL, compared to dominant AFLPs in the F2, we were able to detect an occasional additional minor QTL, as was the case for the SC2 locus, the second locus for seed coat colour in addition to the strong SC1 locus.
In the future the map will be improved by adding additional markers, including the AFLP markers scored in the F2, which are expected to enlarge the map, to combine linkage groups corresponding to the same chromosome and to add several of the markers that could not be mapped as yet.

Materials and methods

Plant growth and generation of the RIL population

The F2 cross of *B. rapa* which was made between genotypes L58 (♀) × R-o-18 (♂) (chapter 3) was propagated through single seed descent until the F7 generation. Individual plants were grown in April 2007 in separate pots in a temperature-controlled greenhouse with artificial day length extension to 16 hours. After about four weeks the first lines started to flower. The inflorescences were covered with plastic bags to prevent cross-pollination. In case of poor seed set, hand pollinations were performed. The F7 population of 160 individuals obtained in June 2009 was considered as a recombinant inbred line (RIL) population for genetic studies.

Trait measurement

The RIL population was used to determine the phenotypes of 21 traits; seed related traits including: seed colour, seed weight, seed oil, seed ripening, seed setting, seed vivipary and morphological traits including flowering time, total height, plant height until first flower, branch number, silique length, beak length, silique number, number of seeds per silique, number of siliques in main stem, carpel number, pod shattering, total leaf number, leaf number until first open flower, length of inflorescence and cuticular wax. Seed colour was a very prominent phenotype segregating in the population. L58 has brown-black seeds and line R-o-18 has yellow seeds.

Seed colour of fully mature F8 seeds was scored by eye and ranked into nine different classes ranging from yellow (1) to black (9). Silique length was determined as the average length of three ripe siliques. The mean number of seeds per silique was determined by averaging three ripe siliques. Seed vivipary was scored as no (0), medium (1) and high (2) based on a visual estimation of the number of seeds with radicles when harvested. Cuticular wax was scored by eye from 1 (glossy), 2 (normal) till 3 (non-glossy) (Figure 3). Shattering was scored at harvesting time as 0 (no open siliques), 0.5 (few open siliques) and 1 (many open siliques). Seed oil was measured as explained in chapter 3. Seed ripening was scored as days after sowing, when more than 50% of siliques were yellow and ready for harvest. Seed setting was calculated from reducing flowering time from seed ripening time.

The correlation analysis was done for all traits to find significant relation between the traits.

Figure 3 *Phenotyping of RIL population of B. rapa L58 × R-o-18. A: Seed vivipary, B: silique (SL)and beak length (BL) variation, C: pod shattering, D: carpel number variation, E: glaucus plant appearing a bit more dark green at top, and the glossy plant at the bottom.*

Adventitious root formation was tested on 5-day-old cotyledons and hairy root transformation was performed using 5-day-old seedlings as explained in Chapter 3. *A. rhizogenes* MSU440 was used for hairy root transformation (Limpen et al., 2004).

Construction of a genetic linkage map and QTL analysis

Genotyping was done using a 384 SNP loci assay. DNA was extracted from frozen leaves according to a modified CTAB procedure (Van der Beek et al. 1992). Because large amounts of DNA are required for the GoldenGate assay, DNA was amplified with the Genomiphi-kit (*Illustra*™ *GenomiPhi*™ *V2 DNA Amplification Kit*, GE Healthcare UK). For SNP discovery the CRoPS® technology of Keygene N.V. was deployed. This technology was developed as a novel approach for single nucleotide polymorphism discovery between two or more samples (van Orsouw *et al*. 2007). CRoPS analysis was performed comparing two *Brassica rapa* lines (Kenshin and Chiifu) and more than 1300 putative SNPs at high quality were discovered. The SNP-harboring sequences were sent to Illumina® for processing by the Illumina Assay Design Tool (ADT). ADT generates scores for each SNP that could vary from 0 to 1; SNPs with the scores above 0.6 have a high probability to be converted into a successful genotyping assay. A total of 384 SNPs were selected, all having ADT scores above 0.6. A total of 100-500 ng of genomic (GenomiPhi) DNA per plant was used for Illumina SNP genotyping at Keygene N.V. using the Illumina BeadXpress™ platform and the GoldenGate Assay following the manufacturer's protocol. The fluorescence images of an array matrix carrying Cy3- and Cy5-labeled beads were generated with the two-channel scanner. Raw hybridization intensity data processing, clustering and genotype calling were performed using the genotyping module in the BeadStudio package (Illumina, San Diego, CA, USA).

Part of the DNA was used for the SSR-detection as described by (Choi et al., 2007) using the LI-COR system (Lincoln, Neb) 4200 DNA sequencer (Myburg et al., 2001). All weak and

ambiguous bands were scored as “unknown”. Linkage analysis and map construction were carried out using the program JoinMap4 (http://www.kyazma.nl/). All markers were grouped using increasing LOD scores (ranging from 5 to 15) to identify linkage groups. The regression mapping algorithm was used for linkage analysis. Recombination frequencies were converted to centiMorgan distances using the Kosambi mapping function. MAPQTL 5.0 (http://www.kyazma.nl/) was used for QTL analysis. First, an interval mapping test was performed to find putative QTLs. A permutation test with 1000 repetitions was applied to determine the LOD thresholds (at p=0.05). A LOD-score of 2.8 was used as the significance threshold for the presence of a candidate QTL. Multiple-QTL model (MQM) mapping also was performed to locate QTLs after the selection of cofactors. Linkage maps along with QTLs were visualized using Mapchart (Voorrips 2002).

Acknowledgements

This work was financially supported by a personal grant to Hedayat Bagheri from the Ministry of Science, Research and Technology of Iran and the IOP Genomics project IGE050010 on Brassica Vegetable Nutrigenomics.

References

Akhunov E., Nicolet C., Dvorak J. (2009). Single nucleotide polumorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. Theor Appl Genet. 119:507-517.

Choi S.R., Teakle G.R., Plaha P., Kim J.H., Allender C.J., Beynon E., Piao Z.Y., Soengas P., Han T.H., King G.J., Barker G.C., Hand P., Lydiate D.J., Batley J., Edwards D., Koo D.H., Bang J.W., Park B.S., Lim Y.P. (2007). The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. Theor Appl Genet. 115:777-792.

Feng F., Liu P., Hong D., Yang G. (2009). A major QTL associated with preharvest sprouting in rapeseed (*Brassica napus* L.). Euphytica. 169:57-68.

Gómez-Campo C. (1999). Biology of *Brassica* coenospecies. Elsevier Science, The Netherlands.

Gomez-Roldan V., Fermas S., Brewer P.B., Puech-Pagès V., Dun E.A., Pillot J.P., Letisse F., Matusova R., Danoun S., Portais J.C., Bouwmeester H., Bécard G., Beveridge C.A., Rameau C., Rochange S.F. (2008). Strigolactone inhibition of shoot branching. Nature. 455:180-194.

Gu Q., Ferrándiz C., Yanofsky M.F., Martienssen R. (1998). The *FRUITFULL* MADS-box gene mediates cell differentiation during Arabidopsis fruit development. Development. 125:1509-1517.

Ji H.S., Chu S.H., Jiang W., Cho Y.I., Hahn J.H., Eun M., McCouch S.R., Koh H.J. (2006). Characterization and mapping of a shattering mutant in rice that corresponds to a block of domestication genes. Genetics. 173:995-1005.

Kianian S.F., Quiros C.F. (1992). Generation of a *Brassica oleracea* composite RFLP map: linkage arrangements among various populations and evolutionary implications Theor Appl Genet. 84:544-554.

Koornneef M., Alonso-Blanco C., Vreugdenhil D. (2004). Naturally occurring genetic variation in *Arabidopsis thaliana.* Annu Rev Plant Biol. 55**:**141-172.

Koornneef M., Bentsink L., Hilhorst H. (2002). Seed dormancy and germination. Curr Opin Plant Biol. 5:33-36.

Koornneef M., Hanhart C.J., Thiel F. (1989). A genetic and phenotypic description of eceriferum (cer) mutants in *Arabidopsis thaliana*. J Hered. 80:118-123.

Li W., Zhang P., Fellers J..P, Friebe B., Gill B.S. (2004). Sequence composition, organization, and evolution of the core Triticeae genome. Plant J.40:500-511.

Liljegren S.J., Ditta G.S., Eshed Y., Savidge B., Bowman J.L. Yanofsky M.F. (2000). *SHATTERPROOF* MADS-box genes control seed dispersal in Arabidopsis. Nature. 404:766-770.

Limpens E., Ramos J., Franken C., Raz V., Compaan B., Franssen H., Bisseling T., Geurts R. (2004). RNA interference in *Agrobacterium rhizogenes*-transformed roots of Arabidopsis and *Medicago truncatula*. J Exp Bot. 55:983-992.

Matilla A.J., Matilla-Vazquez M.A. (2008). Involvement of ethylene in seed physiology. Plant Sci. 175:87-97.

Morgan C., Bavage A., Bancroft I., Bruce D., Child R., Chinoy C., Summers J., Arthur E. (2003). Using novel variation in Brassica species to reduce agricultural inputs and improve agronomy of oilseed rape-a case study in pod shatter resistance. Plant Genet Resources. 1:59-65.

Myburg A.A., Remington D.L., O'Malley D.M., Sederoff R.R., Whetten R.W. (2001). Protocol for High-Throughput AFLP Analysis Using LI-COR IR2 Automated Sequencers. Biotechniques. 30:348-52, 354, 356-7.

Narashimhulu S.B., Chopra V.L. (1988). Species specific shoot regeneration response of cotyledons of *Brassicas*. Plant Cell Rep. 7:104-106.

Oldacres A.M., Newbury H.J., Puddephat I.J. (2005). QTLs controlling the production of transgenic and adventitious roots in *Brassica oleracea* following treatment with *Agrobacterium rhizogenes*. Theor Appl Genet. 111:479-88

Østergaard L., Kempin S., Bies D., Klee H., Yanofsky M. (2006). Pod shattering-resistant *Brassica* fruit produced by ectopic expression of the *FRUITFALL* gene. Plant Biotechnol. 4:45-51.

Pink D., Bailey L.,McClement S., Hand P., Mathas E., Buchanan-Wollaston V., Astley D.,King G., Teakle G. (2008). Double haploids, markers and QTL analysis in vegetable brassicas. Euphytica.164:509-514.

Price J.S., Hobson R.N., Neale M.A., Bruce D.M. (1996). Seed losses on commercial harvesting of oilseed rape. J Agr Eng Res. 65:183--191.

Quijada P., Wang X.W., Cao J.S., Hirai M., Kole C. (2007). Genome mapping and molecular breeding in plants, Springer Berlin Heidelberg. 2:211-263.

Reinhardt D., Kuhlemeier C. (2002). Plant architecture. EMBO J. 9:846-851.

Ruan S.L., Duan X.M., Hu W.M. (2000). Occurrence of seed vivipary in hybrid rape (*Brassica napus* L.) and its effect on seed quality. J Zhejiang Univ (Agric & Life Sci). 26:573-578.

Sebastian R.L., Howell E.C., King G.J., Marshall D.F., Kearsey M.J. (2000). An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled-haploid mapping populations. Theor Appl Genet. 100:75-81.

Tang Z.L., Li J.N., Zhang X.K., Chen L., Wang R. (1997). Genetic variation of yellow-seeded rapeseed lines (*Brassica napus* L.) from different genetic sources. Plant Breeding. 116:471- 474.

Teece M.A., Zengeya T., Volk T.A., Smart L.B. (2008). Cuticular wax composition of salix varieties in relation to biomass productivity. Phytochemistry. 69:396-402.

Van der Beek J.G., Verkerk R., Zabel P., Lindhout P. (1992). Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: Cf9 (resistance to *Cladosporium fulvum*) on chromosome 1. Theor Appl Genet. 84:106-112.

van Orsouw N.J., Hogers R.C.J., Janssen A., Yalcin F., Snoeijers S., Verstege E., Schneiders H., van der Poel H., van Oeveren J., Verstegen H. van Eijk M.J.T. (2007). Complexity reduction of polymorphic sequences (CRoPSTM): a novel approach for large-scale polymorphism discovery in complex genomes. PLOS One. 11:e1172.

Voorrips R.E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 93:77-78.

Voorrips R.E., Jongerius M.C., Kanne H.J. (1997). Mapping of two genes for resistance to clubroot (*Plasmodiophora brassicae*) in a population of doubled haploid lines of brassica oleracea by means of RFLP and AFLP markers. Theor Appl Genet .94:75-82.

Wang Y.J., Wu X.L., He C.Y., Zhang J.S., Chen S., Gai J.Y. (2003). A soybean genetic map constructed after the population being tested and adjusted. Scientia Agricultica Sinica. 11:1254-1260.

Wen M., Jetter R. (2009). Composition of secondary alcohols, ketones, alkanediols, and ketols in *Arabidopsis thaliana* cuticular waxes. J Exp Bot. 6:1811-1821.

Zhang X-Q., Li C., Tay A., Lance R., Mares D., Cheong J., Cakir M., Ma J., Appels R. (2008). A new PCR-based marker on chromosome 4AL for resistance to pre-harvest sprouting in wheat (*Triticum aestivum* L.). Mol Breeding. 22:227-236.

Chapter 5

Genetic analysis of the concentrations of health-related compounds in *Brassica rapa*

Hedayat Bagheri[1], Inge van Oorschot[1], Joost Keurentjes[1], Hye Kyong Kim[2], Steffi Fritsche[3], Christian Jung[3], and Mark G.M. Aarts[1]

[1]Laboratory of Genetics, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands; [2]Division of Pharmacognosy, Section Metabolomics, Institute of Biology, Leiden University, Einsteinweg 55, 2300 RA, Leiden, The Netherlands; [3]Plant Breeding Institute, Christian Albrechts University, Olshausenstrasse 40, D-24098, Kiel, Germany.

Abstract

There is a wide variation for nutritional traits in *Brassica rapa* but the genetic basis of this variation is largely unknown. A recombinant inbred population of 160 lines, made by crossing genotype L58, which is a rapid-flowering, self-compatible Cai Xin line, with R-o-18, a self-compatible Indian doubled haploid spring oil line, was profiled to detect quantitative trait loci (QTL) controlling seed tocopherol and seedling metabolite variation as detected by High Performance Liquid Chromatography (HPLC) and Nuclear Magnetic Resonance (NMR) analysis. The two parental lines differed considerably in their metabolite profile. L58 had a higher level of glucosinolates and phenylpropanoids whereas levels of sucrose, glucose and glutamate were higher in the R-o-18 parent. QTLs related to seed tocopherol (α-, γ-, δ-, α-/γ-, and total tocopherol) concentrations were detected on chromosomes A6 and A9 explaining 9.5-36.6% of the seed tocopherol concentration variation. QTLs were identified for 148 of the 238 signals detected by NMR analysis. NMR signals in organic/amino acid, sugar/glucosinolate and aromatic regions could be assigned to 19 different compounds in the respective classes. This analysis shows that QTL analysis of metabolites is feasible and useful for the future detection of markers and/or genes involved in the biosynthesis of nutritional compounds in *B. rapa*.

Introduction

Brassica rapa is an important crop with a variety of forms: oilseed, leafy vegetable and turnip (Zhao et al. 2005) and with a wide distribution in the world. It is a valuable source of diverse health-promoting metabolites (*e.g.* antioxidants, vitamins, glucosinolates). Plants in general produce an amazing diversity of low molecular mass natural compounds (Pichersky and Gang, 2000). Over 100,000 metabolites have been detected (Wink, 1988), for which the structures of close to 50,000 have been elucidated (De Luca and St-Pierre, 2000). Many of these compounds are part of secondary metabolic pathways, which are not directly involved in the central metabolic processes of the plant but play very important roles in ecological interactions like in plant defense against pathogens and herbivores and in response to abiotic factors. Primary metabolites, such as carbohydrates, vitamins, amino and organic acids are found in all plants and they are directly involved in normal growth, development, and reproduction.

Tocopherols are essential nutrients that humans can only obtain via food. Tocopherol content and composition can be determined accurately by High Performance Liquid Chromatography (HPLC) (Schledz et al., 2001; Endrigkeit et al., 2009). They are lipid-soluble amphipathic molecules that act as anti-oxidants (Porfirova et al., 2002; Collin et al., 2008). The α-, β-, γ-, and δ- tocopherols produced by plants are jointly known as vitamin E, α-tocopherol is the most interesting component from a nutritional point of view, because it is selectively taken up in the human liver and the activity of α-tocopherol is 2–50 times higher than that of the others (Gilliland et al., 2006). Seeds generally provide the bulk of the micronutrients to the human diet (Gilliland et al., 2006) and also tocopherols are mainly found in plant seeds (Eenennaam et al., 2003). On average rape seed oil contains 65% γ-tocopherol and 35% α-tocopherol. Generally, no ß-tocopherol is found in *B. napus* and only very low amounts (<1%) of δ-tocopherol (Goffman and Becker, 2001). The ratio of the content of α- to γ-tocopherol varied from 0.32-1.40 in rapeseed (Goffman and Becker, 1998). Since γ-tocopherol has ten-fold lower biological activity than α-tocopherol, increasing the α-tocopherol fraction can improve the nutritional value of rape seed (Goffman and Becker 2001).

There are different aims for tocopherol breeding in plants: the increase of total tocopherol, the increase of α-tocopherol for nutritional purposes, and the increase of δ-and γ-tocopherol for oil stability. γ-tocopherol is known to be the direct precursor of α-tocopherol (Schultz 1990). The enzyme γ-tocopherol methyltransferase (gTMT) catalyzes the conversion from δ- to β- and from γ- to α-tocopherol, respectively (Endrigkeit et al., 2009). Between 5 and 7 loci with additive and/or epistatic effects have been mapped for γ-tocopherol and total tocopherol content and for the α-/γ-tocopherol ratio in a Double Haploid (DH) population of *B. napus* (Marwede et al., 2005). *B. rapa* is also related to the plant reference species *Arabidopsis thaliana*. Fourteen QTLs affecting seed tocopherol content and composition have been

identified in two Arabidopsis RIL populations (Gilliland et al., 2006), while mutation studies revealed the tocopherol synthesis genes, *VTE1* to *VTE5* (Porfirova et al., 2002; Cheng et al., 2003). Recently the first gene from *B. napus* involved in tocopherol biosynthesis (*BnaA.VTE4.a1*) has been cloned (Endrigkeit et al., 2009).

Apart from tocopherol, there are many more secondary plant metabolites in the plant metabolome that are suggested to have an effect on health. For instance, it is known that a diet rich in cruciferous vegetables can significantly reduce the risk to develop a range of cancer types by an intake of as little as 10 g per day (Kohlmeier and Su, 1997). The development of high throughput technologies and untargeted methods that use Liquid Chromotography combined with Mass Spectrometry (LC-MS) has been instrumental in revealing a large genetic variation for many metabolites, also within species (Keurentjes et al., 2006). Glucosinolates, for instance, are sulfur containing plant metabolites with anti-carcinogenic properties. They form a group of more than 100 plant secondary metabolites present primarily in the Brassicacea family; each plant species contains a blend of different glucosinolates in significant amounts (Fenwick and Heaney, 1983; Fahey et al., 2001). This blend is largely responsible for the typical flavor and odor of Brassicaceae species. There are also significant differences within the crop species for their glucosinolate profiles (Ciska et al., 2000). Glucosinolates can be grouped into three chemical classes: aliphatic, indole and aromatic, according to whether their amino acid precursor is methionine, tryptophan or an aromatic amino acid (tyrosine or phenylalanine), respectively (Giamoustaris and Mithen 1996). Aliphatic glucosinolates are the most prominent glucosinolates found in *Brassica* vegetables (Mithen et al. 2003). The concentration and chemical structure can vary considerably, depending on the genotype, stage of development, tissue type and environmental conditions (Cartea and Velasco, 2008).

More than 90 different aliphatic glucosinolates have been identified among plants (Fahey et al., 2001) of which 8-16 are found in *B. rapa* (He et al., 2000; Padilla et al., 2007; Lou et al., 2008). QTL mapping of leaf aliphatic glucosinolate loci has been carried out in two doubled haploid (DH) populations of *B. rapa* and 16 loci controlling aliphatic glucosinolate accumulation were identified (Lou et al., 2008).

Flavonoids form a second class of health promoting compounds. These secondary metabolites are synthesized from phenylpropanoid and more than 6000 different flavonoids have been reported (Harborne and Williams, 2000). Based on the oxidation level of the C-ring, different classes of flavonoids are distinguished, including flavonols, flavones, flavanols, and anthocyanins (Lepiniec et al., 2006). Flavonoids serve different functions in plants such as pigmentation of flowers, insect interaction, protection against UV light, plant defence, lignin formation and antioxidant properties (Koes et al**.**, 1993); and are important health related compounds in Brassica (Schijlen et al., 2004). The major flavonoid compounds in *A. thaliana*

are the flavonols, predominantly kaempferol glycosides (Veit and Pauli, 1999; Lepiniec et al., 2006). Altered seed pigmentation mutants in Arabidopsis, such as *transparent testa* (*tt*) mutants, are often impaired in flavonoid accumulation (Koornneef 1990).

Amino acids like threonine, valine and isoleucine are essential food ingredients for humans, while alanine, asparagine and glutamine can be produced in the human body. Glutamate is very important in amino acid metabolism in plants since it is the substrate for the synthesis of glutamine from ammonia; in addition the α-amino group of glutamate is transferred to all other amino acids (Forde and Lea, 2007). Transfer of the α-amino group of glutamate to oxaloacetate creates aspartate which is a precursor of asparagine (Lea et al., 2007), threonine, and isoleucine (Azevedo et al., 2006). The α-amino group can be also transferred to pyruvate and thus form alanine (Forde and Lea, 2007).

Carbohydrates such as glucose are other important primary metabolites. Glucose, as primary product of photosynthesis, is a major energy source and structural/storage component. Starch and cellulose are made up of alpha and beta glucose (glucose isomers) respectively. In the Arabidopsis mutant *gin1* (glucose-insensitive), glucose repression of cotyledon greening and expansion, shoot development, floral transition, and gene expression is impaired (Zhou et al., 1998). There is no report about QTL analysis of glucose in *B. rapa*.

To get unambiguous structural information about a metabolite, Nuclear Magnetic Resonance (NMR) and particularly proton NMR (^{1}H NMR analysis) is probably among the most selective analytical options available, as it is a non-destructive method and can simultaneously detect all proton-bearing compounds (Choi et al., 2006). Although it has a lower sensitivity compared to Mass Spectrometry (MS) (Moco et al., 2007), ^{1}H NMR spectroscopy has been used to uncover qualitative and quantitative differences of various cultivars of *B. rapa*. Different cultivars could be distinguished by elucidated metabolites, several organic and amino acids, carbohydrates, adenine, indole acetic acid (IAA), phenylpropanoids, flavonoids and glucosinolates (Abdel-Farid et al., 2007).

We have used this technique to analyze the genetic variation for a range of (secondary) metabolites in a *B. rapa* seedlings of a recently develop RIL population of 160 lines (chapter 4). In addition, a targeted approach detecting tocopherols was used to analyze variation for these compounds in seeds of the same population.

Results

HPLC result of seed Vit E measurement

To determine the extent of variation in the predominating vitamine E compounds, tocopherols, we analyzed seeds of the parental lines R-o-18 and L58 and all individual lines of the RIL population. L58 showed higher levels than R-o-18 for all tocopherols except for δ-tocopherol. Transgression beyond the parental values was observed for all measured tocopherols. In case of δ-tocopherol transgression was beyond R-o-18 only. Some lines showed very high α-tocopherol concentration in comparison to the other components. Table 1 and Figure 1 show the descriptive statistics of tocopherol analysis.

Table 1. Overview of tocopherol analysis, in parents and in the RIL population of L58×R-o-18. Tocopherol concentration is given in mg per g of seed.

Tocopherol	α-	γ-	δ-	total	α-/γ- ratio
L58	0.317	0.273	0.004	0.595	1.159
R-o-18	0.250	0.233	0.019	0.502	1.076
Max value	0.885	0.530	0.034	1.199	8.266
Min value	0.161	0.086	0.003	0.325	0.537
Mean value(±SD)	0.365(0.123)	0.294 (0.092)	0.015 (0.006)	0.675(0.149)	1.428 (0.943)

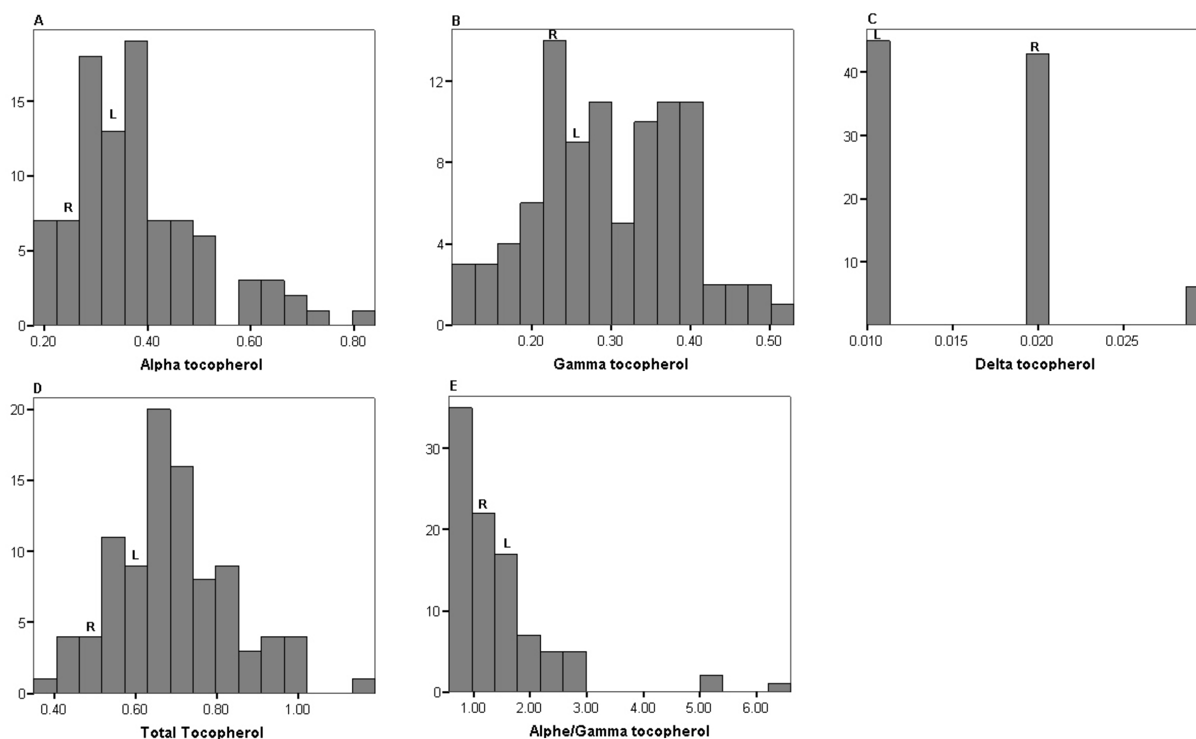


***Figure 1** Frequency distributions of non-normalized data of Vit E in seeds of the L58 × R-o-18 RIL population. The vertical axis indicates the number of lines per trait value class and the horizontal axis the different trait value classes. (A) alpha tocopherol (mg/g); (B) gamma tocopherol (mg/g); (C) delta tocopherol (mg/g); D) total tocopherol (mg/g); (E) alpha/gamma tocopherol ratio; The parental values are the mean of three replicates, indicated with L as L58 and R as R-o-18.*

Correlation analysis of Vit E components revealed that total tocopherol was highly positively correlated with α- and γ- tocopherol. δ-tocopherol was negatively correlated with the ratio of α-/γ- tocopherol while its correlation with α- and γ-tocopherol was not significant (Table 2). Correlation analysis of Vit E components with seed coat colour (chapter 4) showed that seed coat colour had a significant positive correlation with δ-tocopherol but a significant negative correlation with α-, α-/γ- and total tocopherol (Table 3).

Table 2 *Pearson correlation analysis of Vit E components and seed coat colour in the L58 × R-o-18 RIL population. AL: α-tocopherol; Ga: γ-tocopherol; De: δ-tocopherol; To: total tocopherol; ALGa: α / γ tocopherol ratio; and SC: seed coat colour. ** means significant at P≤0.01; * significant at P≤0.05.*

Trait	SC	AL	Ga	De	ALGa
AL	-0.44**	1			
Ga	0.18	-0.09	1		
De	0.231*	-0.16	0.15	1	
ALGa	-0.373**	0.69**	-0.65**	-0.23*	1
To	-0.249*	0.78**	0.54**	-0.01	0.19

* Correlation is significant at the 0.05 level (2-tailed).

** Correlation is significant at the 0.01 level (2-tailed).

NMR results of seedling metabolites detection

To assess the variation in metabolites present in the *B. rapa* RIL population we performed NMR analysis on young seedlings. Usaually an NMR spectrum consists of hundred signals. Among these, 19 compounds in the organic/amino acid, sugar/glucosinolate and aromatic regions of NMR spectra were annotated by ^{1}H-NMR (Table 2), but for many unknown signals further 2D NMR measurements are required to be identified. 1H-NMR data of RIL were subjected to principal component analysis (PCA). PCA of seedling metabolites revealed that the two parental lines were quite different, especially in PC2. PC2 was mostly composed of progoitrin, polypropanoids and organic compounds, while PC1 mostly corresponded to neoglucobrassicin (Figure 2).

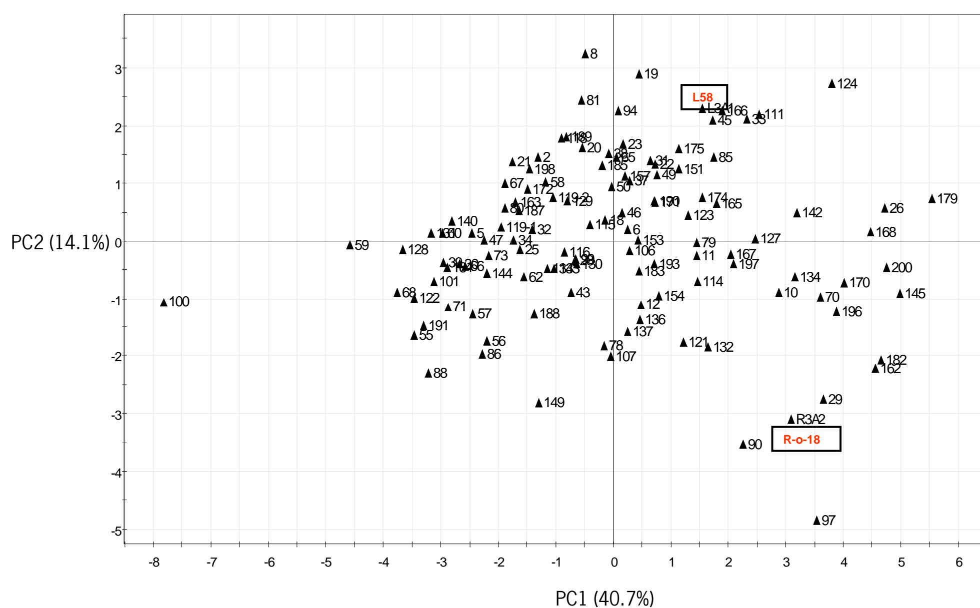


Figure 2 *PCA analysis of seedling metabolites of the L58 × R-o-18 RIL population. Parental values are indicated in box.*

The PCA analysis (Figure 2) revealed that L58 had a higher level of glucosinolates and phenylpropanoids whereas the levels of sucrose, glucose and glutamate were higher in R-o-18. The major phenylpropanoid was sinapoyl glucose. There were at least two more aliphatic glucosinolates, one very similar to progoitrin. Their exact identity can be determined upon two-dimensional (2D) spectrum analysis. In the aromatic area, there were signals from indole compounds (tryptophan or indole acetic acid (IAA)), but this also needs to be confirmed by 2D spectra. As an example, in Figure 3 the ^{1}H NMR spectra of L58 are shown, expanded for organic/amino acid, sugar/glucosinolate, and aromatic regions.

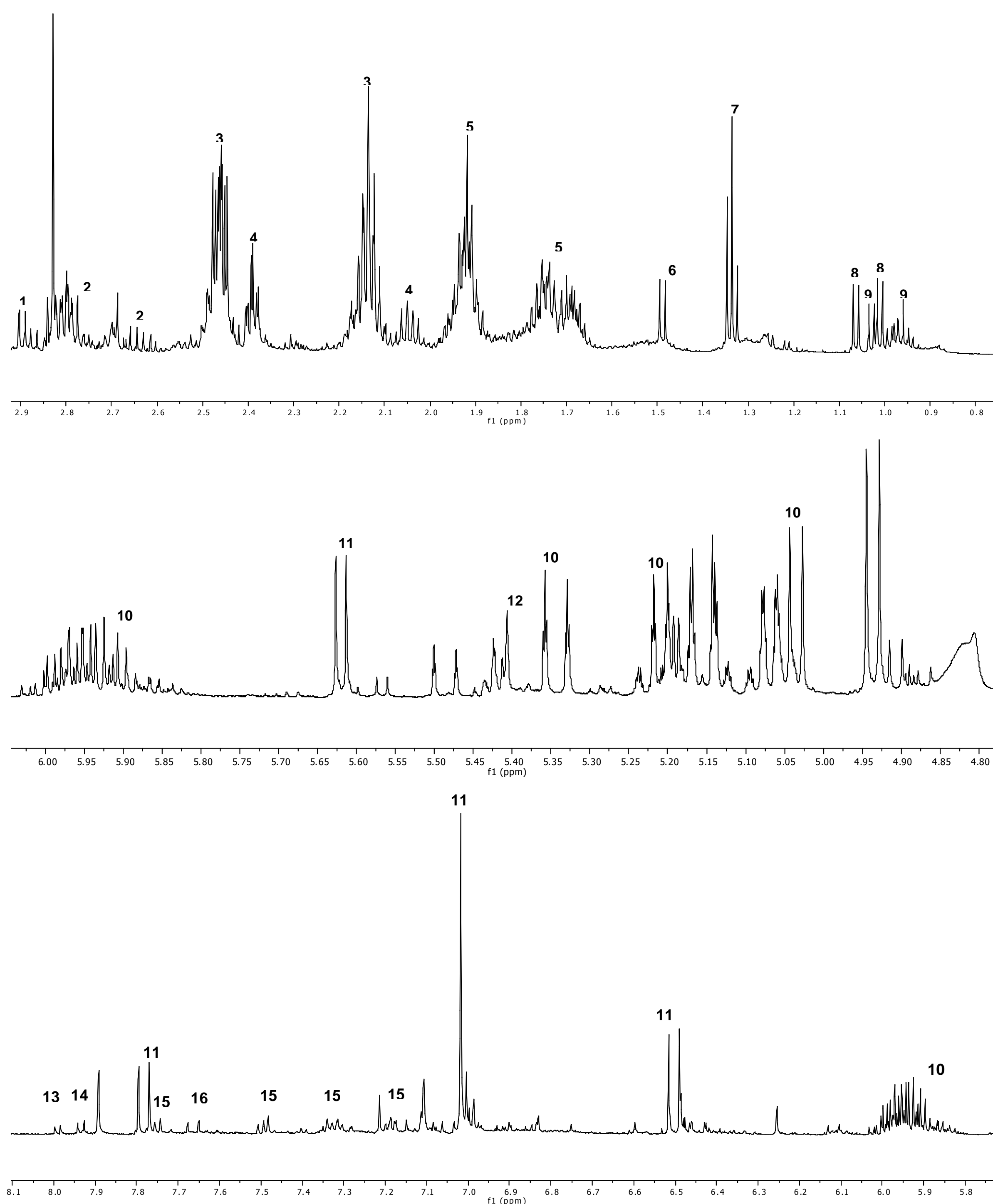


Figure 3 *[1] H NMR spectra of L58 expanded for organic/amino acid (δ 0.8- δ 4.0), sugar/glucosinolate (δ 4.5- δ 6.0), and aromatic regions (δ 6.2- δ 8). The X-axis shows the chemical shift in ppm or δ. 1: asparagine, 2: malate, 3: glutamine, 4: glutamate, 5: citrulline, 6: alanine, 7: threonine, 8: valine, 9: isoleucine, 10: progoitrin, 11: sinapoyl glucose, 12: glucose, 13: flavonoid 1, 14: flavonoid 2, 15: neoglucobrassicin, and 16: phenylpropanoid.*

QTL analysis of seed Vit E

Each tocopherol component was subjected to QTL analysis and QTLs related to seed tocopherol (α-, γ-, δ-, α-⁄γ-, and total tocopherol) concentrations were detected on chromosomes A6 and A9 (Figure 4, Table 3). About 46% of the variance for α-tocopherol was explained by two QTLs (AL1 and AL2, respectively on chromosomes A9 and A6). Two QTLs were found for total tocopherol (To1 and To2), explaining almost 42% of the vit E variance. WhileTo2 co-located with AL1 on A9, To1 did not co-locate with AL2 on A6, but instead co-located with the only QTL for γ-tocopherol (Ga) on A6. The QTL for delta tocopherol (De) maps to the same region of A9 where also AL1 and To2 were mapped. In this region also the seed coat colour QTL (SC1) was mapped.

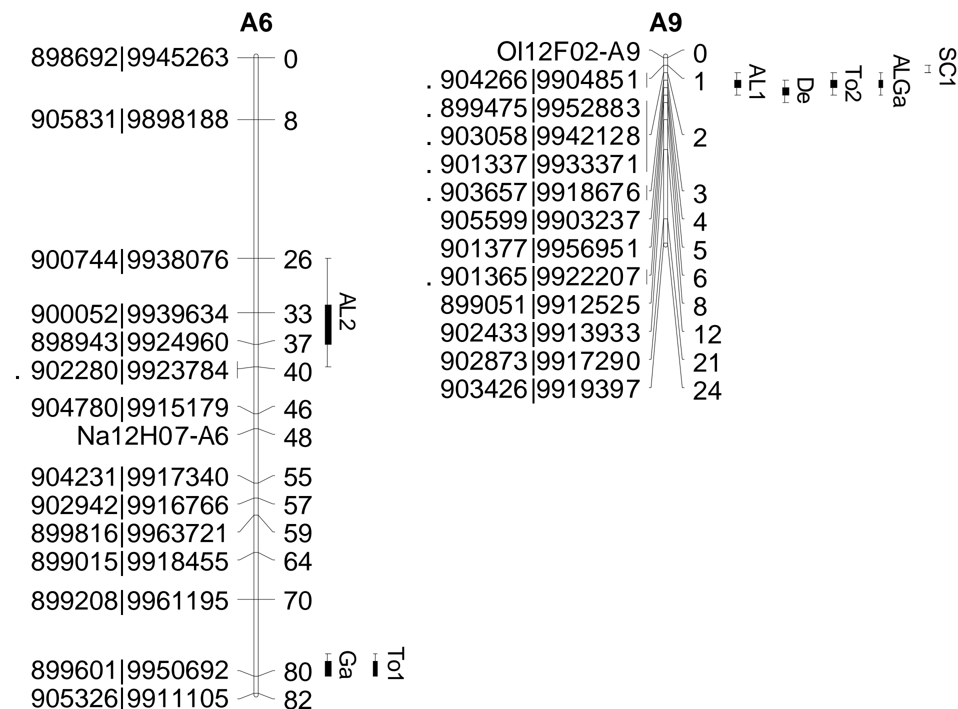


***Figure 4** Vitamin E related QTLs identified in the B. rapa L58 x R-o-18 RIL population. QTLs mapped to the linkage groups A6 and A9 are indicated with boxes and whiskers representing 1-LOD and 2-LOD confidence intervals (95%) respectively for significant QTL. AL: α-tocopherol; De: δ-tocopherol; To: total tocopherol, and ALGa: α-/γ- tocopherol ratio. Also the location of seed colour QTL SC1 is shown.*

Table 3 *QTLs related to Vitamin E concentration as detected in seeds of the B. rapa L58 × R-o-18 RIL population. Peak position indicates the location of the highest LOD score for each QTL. % Expl. var. is the percentage of total phenotypic variance explained by individual QTLs. The allelic effect of each QTL is indicated (Effect), which is calculated as μA-μB (μ= mean), where A and B are RILs carrying L58, respectively R-o-18, genotypes at the QTL position. Effects are given in mg/g or without unit (ratio of α/γ).*

Trait	QTL	Linkage group	LOD	Peak position (cM)	% Expl. var.	Effect
α-tocopherol	AL1	A9	9.15	4.3	36.6	-0.15
	AL2	A6	3.35	36.7	9.5	+0.08
γ-tocopherol	Ga	A6	4.75	81.7	21.0	+0.09
δ-tocopherol	De	A9	2.83	5.14	13.0	+0.004
Total tocopherol	To1	A6	7.96	81.7	27.0	+0.17
	To2	A9	4.68	1.9	14.5	-0.12
α/γ tocopherol ratio	ALGa	A9	5.14	2.37	22.3	-0.92

QTL analysis of seedling metabolites

Genetic analysis of 238 signals detected in the NMR spectra enabled the identification of QTLs for 148 signals (summarized in Table 3). Six QTLs related to malate, citrulline and choline were identified in the organic region of the spectrum. The strongest QTL was belonging to the class of phenylpropanoids explaining 41% of the variation in seedlings. This QTL is one of six QTLs detected for phenylpropanoid compounds. The major phenylproponoid was sinapoyl glucose showing QTLs on A6 and A9. In the amino acid region at least 12 QTLs controlling variation for alanine, asparagine, glutamine, isoleucine, threonine and valine were detected. One of them on G7 explained more than 30% of asparagine variance. QTL analysis of the glucosinolate region of the NMR spectrum detected six significant loci, with the highest one for progoitrin. About 40% of progoitrin and more than 30% of asparagine variation in seedlings was explained by co-locating QTLs on G7. Alpha and beta glucose QTLs on G8 and G7 explain 19% and 15% of the seedling glucose variation. In the aromatic region three QTLs for flavonoids and three QTLs for sinapoyl glucose were detected. Additional QTLs were detected for compounds of which the chemical identity is not known.

***Table 4** Overview of QTLs related to seedling metabolites detected in the B. rapa L58 × R-o-18 RIL population. Peak position indicates the location of the highest LOD score for each QTL. % Expl. var. is the percentage of total phenotypic variance explained by individual QTLs.*

compound	Linkage group	LOD	Peak position (cM)	% Expl. Var.
Alanine	G11	3.4	27.4	12.5
Asparagine	G7	11.2-12.3	44.1	33.0-36.2
	A5	4.0- 4.6	24.2	10.0-11.7
Glutamine	A4	3.3	0	11
	G7	3.4	24.8	9.7
	G8	3.8-7	12.5, 23.18	12.5-23.0
	G11	2.5-5.7	27.4	8.4-15.0
Isoleucine	G8	3.1-4.6	23.1	11.4-16.7
Threonine	G7	4.6	3.3	16.8
Valine	G8	4.0-4.9	23.1	15.0-16.5
	A9	3.4	20.5	10.7
Glutamate	G8	4.3-7.0	12.5, 4.8, 8.4	15.7-24.5
	G7	3.4-3.6	24.8	10.4-11.2
	G11	2.9-3.6	27.4	09.0-11.2
Citrulline	A6	2.5-3.3	8.4	9.3-12.4
Choline	G7	3.4	24.8	12
	A6	1.9	56.9	6.5
Malate	A10	5.3-8.6	23.6, 13.5	19.2-27.2
	A6	3.3	59.0	9.4
Phenylpropanoids	A6	5.0	8.4	18.1
	A9	4.0	1.5	12
	A10	6.5-13.2	7.5, 13.5	20.2-41.3
Progoitrin	A5	4.6-5.3	24.2	11.0-12.5
	G7	9.7-14.4	44.1	28.3-39.5
	G8	4.4	23.1	14.5
Neoglucobrassicin	A5	2.9-4.8	0	11.1-18.0
	G8	3.6-5.8	4.8, 46.4	10.6-17.7
Glucose (alpha)	G8	5.3	12.5	19
Glucose(beta)	G7	4.2	24.8	15.6
Flavonoid 1 (Kaempferol glycoside)	A10	5.1	13.5	18.5
Flavonoid 2 (Kaempferol type)	A6	3.9	79.6	14.4
Flavonoid 3	A5	2.9	0	11
Sinapoyl glucose	A6	2.6-4.8	70.1, 8.4, 0	09.0-17.6
	A9	2.7-5.6	1.5	09.1-17.5
Unknown	A5	4.3-35	30.6	24.0-80.8
Unknown	G8	4-13	4.8	34.0-41.0
Unknown	G8	4.4	8.4	16
Unknown	G7	3.	3.0	12.5
Unknown	G8	3.9	2.9	13.3
Unknown	G8	2.7	0	10.3
Unknown	G8	9.5-15.2	4.8	31.6-45.4
Unknown	G11	3.3-4.4	23.1	11.0-14.2

	G8	3.6-4	27.4	12.0-12.7
Unknown	G7	5.5	44.1	19
Unknown	G7	4.4	24.8	16
Unknown	A6	3.8-4.8	0	12.0-17.2
Unknown	G11	2.6-3.9	27.4	08.0-12
Unknown	G8	3-6.2	12.5, 23.1	10-19.4
Unknown	G11	2.7-3.2	27.4	10-10.4
Unknown	A6	5-5.9	8.4	16.0-21.0
Unknown	A5	3.6	0	13.8
Unknown	G4	3.2	5.5	12
Unknown	G4	3	0	11.2
Unknown	A10	3.5-5.5	13.5	13.6-17.3
Unknown	G7	3.1	3.3	11.8
Unknown	A5	2.9-5.1	0	11.4-19.0
Unknown	G7	6.7-16.1	44.1	22.6-46.5
Unknown	A10	5.1-13.9	7.5	15.1-36.7
Unknown	A5	3.1-4.6	0	12.0-17.1
Unknown	G8	3.2-4.6	12.5	12.3-18.0
		4	8.4	14.7
Unknown	G11	3.6	17.2	12.2
Unknown	G8	2.7-3.1	8.4, 23.1	10.3-12.0
Unknown	G10	3.4-4.5	15.0	12.7-16.4
	A5	2.5-2.7	0	10.0-11.0
Unknown	A6	4.6	79.6	16.8
Unknown	A10	2.6-4.6	52.0	10.0-17.0
	A5	4.4-23.4	24.2	16.0-60.0
Unknown	A6	4.3	6.4	15.8

Discussion

The determination of the concentration of metabolites in a recombinant inbred line population enables the mapping of quantitative trait loci for the individual metabolites and discovery of co-locating and possibly co-regulated compounds. The obtained map positions and insights in the genetic control of the various metabolites can help improving the breeding for healthy compounds in crop plants.

Tocopherols are important dietary nutrients as they constitute the essential vitamin E. The presence of variation in vitamin E seed concentration in the RIL population under study allowed the detection of QTLs for all tocopherols. Considerable transgression of tocopherol levels beyond the parental values revealed that both parents contribute alleles with positive effects at the different loci, in case of the two major QTL for alpha and total tocopherol respectively on A6 and A9. This observation also indicates a potential for improvement of vitamin E content and composition through classical breeding. As α-tocopherol is highly positively correlated to the total tocopherol concentration and one of their respective QTLs maps to the same position, it suggests that the concentration of α-tocopherol, and not of the intermediate γ-tocopherol, has the major contribution to the overall tocopherol concentration. However, the second locus for total tocopherol concentration co-locates with the γ-tocopherol QTL on A6, showing that indeed both tocopherols with highest concentrations contribute to the total tocopherol concentrations. The absence of a significant correlation between α-, γ- and δ concentrations and the finding that these were controlled by different QTLs indicates their independent genetic regulation, which is in agreement with findings of Marwede et al. (2004) in canola. Thus, with three independent loci controlling α- and γ-tocopherol it should be possible to enhance the concentration of both. This will have a negative effect on δ-tocopherol concentration though, since the co-locating De and AL1 loci have opposite allele effects and although both QTLs may be caused by different genes, the corresponding alleles may be difficult to separate genetically. A similar antagonistic effect was seen for Arabidopsis, where over-expression of the *AtVTE3* gene, encoding the tocopherol biosynthetic enzyme 2-methyl-6-phytylbenzoquinol methyltransferase, increases accumulation of δ-tocopherol but decreases the γ-tocopherol content in seeds (Van Eenennaam et al., 2003).

The seed colour locus SC1 is also located on the top of A9 and correlation analysis indeed showed yellow seeds to contain higher amounts of α- and total tocopherol and lower amounts of δ- tocopherol. Since there is no biochemical or genetic reason to suggest a common biochemical basis of biosynthesis of tocopherol and the flavonoids contributing to seed colour, a close linkage of different genes rather than one common gene is the most likely explanation for this correlation.

An adequate diet must contain essential amino acids. There is not much known about genetic variation in amino acid content in Brassica. Here we identified four QTLs for the essential amino acids isoleucine and valine, co-locating on G8. The isoleucine biosynthesis pathway is almost parallel to valine biosynthesis, except for its first step which involves a threonine deaminase and dehydrase. This locus possibly corresponds to the gene encoding the biosynthetic threonine dehydratase (TD) isozyme, as what has been isolated from tomato and potato (Samach et al., 1991; Hofgen at al., 1995). Non-essential amino acids are equally important as the essential amino acids in our body. Eight QTLs for non-essential amino acids were identified in the RIL population. These were all independent, except for one, which was shared between alanine and glutamine. Glutamate also shared a QTL with this co-located region. As glutamate is the substrate for the glutamine synthesis and the α-amino group of glutamate can be transferred to pyruvate to form alanine (Forde and Lea, 2007), this locus can regulate the accumulation of glutamate along with glutamine and alanine synthesis and is somewhere upstream in the biosynthesis pathway of all these compounds.

Six glucosinolates (progoitrin and neoglucobrassicin) related QTLs localized on A5, G7 and G8. Previously five QTLs related to progoitrin were mapped to chromosomes A1, A3, A4, A8 and A10 by Lou et al. (2008) in *B. rapa* DH population. He used forty-day-old leaves for metabolite analysis and there is likely to be difference in the regulation of glucosinolate biosynthesis in mature leaves, compared to young seedlings, which may even still carry glucosinolates originally present in the seed. In any case the progoitrin QTL on A5 is a new locus to regulate progoitrin accumulation in young seedlings. Additional markers are needed to assign linkage groups G7 and G8 to the *B. rapa* reference map (Choi et al., 2007) and thus rule out similarity of these loci with loci mapped by Lou et al. (2008). Extensive studies about aliphatic glucosinolates in Arabidopsis have led to the identification of genes encoding AOP (2-oxoglutarate-dependent dioxygenase) and MAM (methyl-thioalkylmalate synthase), controlling the modification of side-chain moiety and elongation respectively (Field et al., 2004; Heidel et al., 2006). The regulation of aliphatic glucosinolate biosynthesis enzymes is controlled in Arabidopsis by the R2R3 myb-like transcripton factors, myb28 and myb29 (Hirai et al., 2007).

Flavonoids are most commonly known for their antioxidant activity. Three QTLs were detected for flavonoids (kaempferol derivatives) on A5, A6, and A10. The last QTL on A10 co-located with malate and phenylpropanoids QTLs. As flavonoids are synthesized via the phenylpropanoid pathway, the locus on A10 is probably involved in upstream phenypropanoid biosynthesis.

Alpha glucose QTL was mapped on the same position as the glutamate locus while beta glucose QTL was mapped together with glutamine, glutamate, and choline related QTLs. Possibly glucose QTLs have role in carbon metabolism during amino acid synthesis.

The markers which we detected here can be used to establish a diagnostic marker assisted selection in plant breeding. Gaining Knowledge about the genetics behind the accumulation of healthy compounds is necessary while new opportunities for pathway elucidation and identification will be created. Genetic improvement of healthy compounds improves nutritional quality which is an important objective in crop breeding programs.

Materials and methods

Seed preparation for HPLC

F7 seeds derived from one plant per line of recombinant inbreds of the *B. rapa* L58 × R-o-18 population were used for Vitamin E (tocopherol) measurement (two replicates of lines with two technical replicates from the same plant). To validate the data, 20 lines chosen at random were re-measured after measuring all samples. For the tocopherol extraction 10-40 mg seeds were ground in 2-ml reaction tubes with a Geno/Grinder 2000 (SPEX-Sample Prep, Metuchen, USA) using *n*-heptane and 3.0 - 4.0 mm metal beads. The samples were incubated at -20 °C for 2 h. Further applications and HPLC analyses were performed as described by (Dähnhardt et al., 2002; Falk et al., 2003; Schledz et al., 2001).

Quantification of the tocopherols was done by fluorescence detection (excited at λ=290 nm, emission at λ=328 nm). To identify the individual tocopherols, the retention times were compared with standard substances (Merck).

Seedling preparation for NMR

Thirty seeds per RIL of the *B. rapa* L58 × R-o-18 were used. Seeds were surface sterilized with 70% ethanol (v/v) for 30 seconds, followed by agitation for 5 min in sodium hypochlorite (2.0% active chlorite). After three rinses in sterile distilled water, 30 seeds of each individual (for every experiment) were placed in 15×90 mm petri dishes, each containing 20–25 ml half strength MS salts and vitamins, without sucrose and solidified with 0.8% (w/v) agar. Petri dishes were placed vertically in a growth chamber maintained at 25˚C with a 16h light/8h dark photoperiod at a light intensity of 60 $mEm^{-2}s^{-1}$. Five-day-old seedlings without roots were harvested and freeze-dried.

Seedling Extraction and NMR Analysis

20 mg seedlings (dry weight) were extracted with a mixture of 500 µl methanol-d_4 and 500 µl D_2O (KH_2PO_4 buffer, pH 6.0) containing 0.05% TSP (trimethyl silyl propionic acid sodium salt, w/v) by ultra-sonication for 20 min. After centrifugation, 800 µl supernatant were transferred to an NMR tube. ^{1}H NMR spectra were recorded at 25°C on a 600 MHz Bruker AV600 spectrometer equipped with a cryoprobe operating at a proton NMR frequency of 600.13 MHz. CD_3OD was used as the internal lock. Each ^{1}H NMR spectrum consisted of 128 scans using following parameters: TD=51200, Spectrum width=16.02 ppm, 0.25 Hz/point, pulse width (PW) = 30° (6.6 µsec), acquisition time = 1.70 sec. and relaxation delay (RD) = 2.0 sec. A pre-saturation sequence was used to suppress the residual H_2O signal with low power selective irradiation at the H_2O frequency at µ 4.869 (2915.9 Hz) by 60.59 dB during the recycle delay. Free Induction Decay (FID)s were Fourier transformed with LB = 0.3 Hz and the spectra were zero filled to 32 K points. The resulting spectra were manually phased

and baseline corrected, and calibrated to TMSP at 0.0 ppm, using Topspin (version 2.1, Bruker).
The ^{1}H NMR spectra were automatically reduced to ASCII file. Spectral intensities were scaled to the internal standard (TSP) area and reduced to integrated regions of equal width (0.04 ppm) corresponding to the region of δ 0.3–δ 10.0. The region of δ 4.75–δ 4.90 and δ 3.28–δ 3.34 was excluded from the analysis because of the residual signal of HDO and CD_3OD, respectively. Bucketing was performed by AMIX software (Bruker). Principal component analysis (PCA) was performed with the SIMCA-P software (v. 12.0, Umetrics, Umea, Sweden) with scaling based on the Pareto method.

Acknowledgements

This work was financially supported by a personal grant to Hedayat Bagheri from the Ministry of Science, Research and Technology of Iran and the IOP Genomics project IGE050010 on Brassica Vegetable Nutrigenomics. We thank Maarten Koornneef for his comments on the manuscript, Corrie Hanhart for her technical assistance and Jens Hermann for assistance with the tocopherol analysis.

References

Abdel-Farid I.B., Kim H.K., Choi Y.H. Verpoorte R. (2007). Metabolic characterization of *Brassica rapa* leaves by NMR spectroscopy. J Agric Food Chem**.** 55:7936-7943.

Azevedo R.A., Lancien M., Lea P.J. (2006). The aspartic acid metabolic pathway, an exciting and essential pathway in plants. Amino Acids. 30:143-162.

Cartea M.E., Velasco P. (2008). Glucosinolates in Brassica foods: bioavailability in food and significance for human health. Phytochemistry Rev. 7:213-229.

Cheng Z., Sattler S., Maeda H., Sakuragi Y., Bryant D.A., DellaPenna D. (2003). Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes. Plant Cell. 15:2343-2356.

Choi Y.H., Kim H.K., Linthorst H.J.M., Hollander J.G., Lefeber A.W.M., Erkelens C., Nuzillard J.M., Verpoorte R. (2006). NMR Metabolomics to revisit the Tobacco Mosaic virus Infection in *Nicotiana tabacum* Leaves. J Nat Prod. 69:742-748.

Choi S.R., Teakle G.R., Plaha P., Kim J.H., Allender C.J., Beynon E., Piao Z.Y., Soengas P., Han T.H., King G.J., Barker G.C., Hand P., Lydiate D.J., Batley J., Edwards D., Koo D.H., Bang J.W., Park B.S., Lim Y.P. (2007). The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. Theor Appl Genet. 115:777-792.

Ciska E., Martyniak-Przybyszewska B., Kozlowska H. (2000). Content of glucosinolates in cruciferous vegetables grown at the same site for two years under different climatic conditions. J Agric Food Chem. 48:2862-2867.

Collin V.C., Eymery F., Genty B., Rey P., Havaux M. (2008). Vitamin E is essential for the tolerance of *Arabidopsis thaliana* to metal-induced oxidative stress. Plant Cell Environ. 31:244-257.

Dähnhardt D., Falk J., Appel J., van der Kooij T.A., Schulz-Friedrich R., Krupinska K. (2002). The hydroxyphenylpyruvate dioxygenase from Synechocystis sp. PCC 6803 is not required for plastoquinone biosynthesis. FEBS Lett. 523:177-181.

De Luca V., St-Pierre B. (2000). The cell and developmental biology of alkaloid biosynthesis. Trends Plant Sci. 5:168-173.

Endrigkeit J., Wang X., Cai D., Zhang C., Long Y., Meng J., Jung C. (2009). Genetic mapping, cloning, and functional characterization of the BnaX.VTE4 gene encoding a -tocopherol methyltransferase from oilseed rape. Theor Appl Genet. In press, doi: 10.1007/s00122-009-1066-6.

Fahey J.W., Zalcmann A.M., Talalay P. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochem. 56:5-61.

Falk J., Andersen G., Kernebeck B., Krupinska K. (2003). Constitutive overexpression of barley 4-hydroxyphenylpyruvate dioxygenase in tobacco results in elevation of the vitamin E content in seeds but not in leaves. FEBS Lett. 540:35-40.

Fenwick G.R., Heaney R.K. (1983). Glucosinolates and their breakdown products in cruciferous crops, food, and feeding stuffs. Food Chem. 11:249-271.

Field B., Cardon G., Traka M., Botterman J., Vancanneyt G., Mithen R. (2004). Glucosinolate and amino acid biosynthesis in Arabidopsis. Plant Physiol**.** 135:828-839.

Forde B.G., Lea P.J. (2007). Glutamate in plants: metabolism, regulation, and signalling. J Exp Bot. 58:2339-2358.

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

Gilliland L.U., Magallanes-Lundback M., Hemming C., Supplee A., Koornneef M., Bentsink l.,

Dellapenna D. (2006). Genetic basis for natural variation in seed vitamin E levels in *Arabidopsis thaliana*. Proc Natl Acad Sci. USA 103:18834-18841.

Goffman F.D., Becker H.C. (2001). Diallel analysis for tocopherol contents in seeds of rapeseed. Crop Sci. 41:1072-1079.

Harborne J.B., Williams C.A. (2000). Advances in flavonoid research since 1992. Phytochemistry 55:481-504.

He H., Fingerling G., Schnitzler W.H. (2000). Glucosinolate contents and patterns in different organs of Chinese cabbages, Chinese kale (*Brassica alboglabra* bailey) and Choy sum (*Brassica campestris* L. ssp. *chinensis* var. *Utilis*). J Appl Bot. 74:21-25.

Heidel A.J., Clauss M.J., Kroymann J., Savolainen O., Mitchell-Olds T. (2006). Natural variation in MAM within and between populations of *Arabidopsis lyrata* determines glucosinolate phenotype. Genetics. 173:1629-1636.

Hirai M.Y., Sugiyama K., Sawada Y., Tohge T., Obayashi T., Suzuki A., Araki R., Sakurai N., Suzuki H., Aoki K., Goda H., Nishizawa O.I., Shibata D., Saito K. (2007). Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proc Natl Acad Sci. USA 104:6478-6483.

Hofgen R., Streber W., Pohlenz H. (1995). Antisense gene expression as a tool for evaluating molecular herbicide targets. Pest Sci. 43:175-177.

Koes R.E., Quattrocchio F., Mol J.N.M. (1993). The flavonoid biosynthetic pathway in plants: function and evolution. BioEssays. 16:123-132.

Kohlmeier L., Su L. (1997). Cruciferous vegetable consumption and colorectal cancer risk: meta-analysis of the epidemiological evidence. FASEB J. 11:369.

Koornneef M. (1990). Mutations affecting the testa colour in Arabidopsis. Arabidopsis Inf Serv. 27:1-4.

Lea P.J., Sodek L., Parry M.A.J., Shewry P.R., Halford N.G. (2007). Asparagine in plants. Ann Appl Biol. 150:1-26.

Lepiniec L., debeaujon I., Routaboul J-M., Baudry A., Pourcel L., Nesi N., Caboche M. (2006). Genetics and biochemistry of seed flavonoids. Rev Plant Biol. 57:405-430.

Lou P., Zhao J., He H., Hanhart C., Del Carpio D.P., Verkerk R., Custers J., Koornneef M., Bonnema G. (2008). Quantitative trait loci for glucosinolate accumulation in *Brassica rapa* leaves. New Phytol. 179:1017-1032.

Marwede V., Gul M.K., Becker H.C., Ecke W. (2005). Mapping of QTL controlling tocopherol content in winter oilseed rape. Plant Breed. 124:20-26.

Mithen R., Faulkner K., Magrath R., Rose P., Williamson G., Marquez J. (2003). Development of isothiocyanate-enriched broccoli and its enhanced ability to induce phase 2 detoxification enzymes in mammalian cells. Theor Appl Genet. 106:727-734.

Moco S., Bino R.J., De Vos R.C.H., Vervoort J. (2007). Metabolomics technologies and metabolite identification. Trac-Trend Anal Chem. 26:855-866.

Padilla G., Cartea M.E., Velasco P., de Haro A., Ordas A. (2007). Variation of glucosinolates in vegetable crops of *Brassica rapa*. Phytochemistry. 68:536-545.

Pichersky E., Gang D.R. (2000). Genetics and biochemistry of secondary meatbolites in plants: an evolutionary perspective. Trends Plant Sci. 5:439-445.

Porfirova S., Bergmüller E., Tropf S., Lemke R., Dörmann P. (2002). Isolation of an *Arabidopsis* mutant lacking vitamin E and identification of a cyclase essential for all tocopherol biosynthesis. Proc Natl Acad Sci. USA 99:12495-12500.

Samach A., Harevent D., Gutfinger T., Ken-Dror S., Lifschitz E. (1991). Biosynthetic threonine deaminase gene of tomato: Isolation, structure, and upregulation in floral organs. Proc Natl Acad Sci. USA. 88:

2678-2682.

Schijlen E.G.W.M., Ric de Vos C.H., van Tunen A.J., Bovy A.G. (2004). Modification of flavonoid biosynthesis in crop plants. Phytochemistry. 65:2631-48.

Schledz M., Seidler A., Beyer P., Neuhaus G. (2001). A novel phytyltransferase from Synechocystis sp. PCC 6803 involved in tocopherol biosynthesis. FEBS Lett. 499:15-20.

Schultz G. (1990). Biosynthesis of α-tocopherol in chloroplasts of higher plants. Fat Sci Technol. 92:86-90.

Textor S., de Kraker J., Hause B., Gershenzon J. Tokuhisa J.G. (2007). MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in *Arabidopsis thaliana*. Plant Physiol. 144:60-71.

Van Eenennaam A.L., Lincoln K., Durrett T.P., Valentin H.E., Shewmaker C.K., Thorne G.M., Jiang J., Baszis S.R., Levering C.K., Aasen E.D., Hao M., Stein J.C., Norris S.R., Last R.L. (2003). Engineering vitamin E content: From *Arabidopsis* mutant to soy oil. Plant Cell. 15:3007-3019.

Veit M., Pauli G.F. (1999). Major flavonoids from *Arabidopsis thaliana* leaves. J Nat Prod. 62:1301-1303.

Wink M. (1988). Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. Theor Appl Genet. 75: 225-233.

Zhou L., Jang J-C., Jones T.L., Sheen J. (1998). Glucose and ethylene signal transduction crosstalk revealed by an Arabidopsis glucose-insensitive mutant. Proc Natl Acad Sci. USA. 17:10294-10299.

Chapter 6

General discussion

Insertional mutagenesis using heterologous maize transposons or *Agrobacterium* mediated T-DNA insertions, has been a valuable tool for the identification and isolation of genes that display a mutant phenotype (Pereira, 2000). The phenotype associated with the altered expression of a gene is often the best clue to its functional role in the plant. Mutant phenotypes can be broadly defined at the morphological, biochemical or physiological levels and provide information on the interaction between different processes. The maize transposon systems *Ac–Ds* and *En-I (Spm/dSpm*) have been shown to transpose after being introduced into numerous heterologous hosts. These heterologous transposons can be modified *in vitro* and thus offer several advantages for transposon mutagenesis over the endogenous systems (Pereira, 1998), which are also not available in many plant species. Activation tagging provides an alternative to knockouts or RNA-silencing induced knockdowns, through the upregulation, rather than abolition, of native gene expression. Employing insertion sequences that carry a strong enhancer element near the border, activates the adjacent gene ectopically or just enhances its expression level resulting in dominant phenotypes (Pereira, 2000). A dominant mutant frequency of about 1% was obtained using *En/I* transposon based-activation tagging in Arabidopsis by Marsch-Martinez et al., (2002).

It would be very useful to have access to a high efficiency mutagenesis tool in *Brassica rapa,* which is receiving increasingly more interest for functional genomics studies, in view of its economic importance, its development as a model *Brassica* crop species (http://www.brassica.info/) and the genomic advantages as a close relative of the plant reference species *Arabidopsis thaliana*. The initial goal of this research project was therefore to develop an *En/I* transposon–based activation tagging population in *B. rapa*. This mutant population, together with an EMS-generated M2 population for tilling purposes as developed by Østergaard e.a. at the John Innes Center (Norwich, UK) (http://www.jic.ac.uk/) could provide valuable and complementary tools to study plant gene function in *B. rapa*.

The first question that had to be answered was if the *En/I* transposon system is active in *B. rapa*. In chapter 2, the activation of a heterologous *En/I* tagging construct was shown in *B. rapa* using *Agrobacterium rhizogenes* hairy root transformation as an easy and fast gene transfer method. Although in principle activation tagging could be achieved with a collection of root cultures containing independent activation tag insertions to screen for altered root phenotypes, we considered a collection of stable activation tag insertion lines a more widely applicable tool. Therefore we initiated *A. tumefaciens* mediated transformation of *B. rapa*. *B. rapa* is known as a species recalcitrant to transformation (Moon et al., 2007; Tsukazaki et al., 2002; Narasimhulu and Chopra 1988), so we expected low transformation frequencies, as we obtained upon transformation with the pCAMBIA2301 construct expressing 35S-GUS. Unfortunately we did not get any transformed shoots carrying the *En/I* activation tagging construct, not even in *B. napus*, which is much less recalcitrant to transformation. Since

tobacco could be easily transformed with the construct we concluded that one of the genes present on the construct poses particular problems upon expression in *B. rapa*. Previously this construct was transformed to Arabidopsis upon *A. tumefaciens* dip-inoculation, also at relatively low transformation frequencies (Marsch-Martinez et al., 2002). The transformants that were obtained all showed a dark-green, dwarfed phenotype, which could be attributed to sequences at the *En/I* construct, even after transposition of the *I*-element activation tag. This construct contains the *SU1* negative selection marker gene, encoding a protein that converts the pro-herbicide R7402 into the herbicide sulfonylurea. Since this gene is expressed under control of the promoter of the Rubisco small subunit, which is mainly active in green tissue, it also explains that production of transgenic roots or calli with this construct was not much of a problem, but that regeneration of photosynthesizing green shootlets turned out to be impossible.

It took a long time and many transformation experiments to find this out. Although the best way to proceed towards an activation tagging procedure was to delete the SU1 gene from the construct and restart transformation, time was too short to be able to achieve this and still obtain an activation tagged population which could be screened. Instead, the focus of the research was shifted towards analysis of genetic variation for transformation ability in *B. rapa*, since the transformation experiments with a few different genotypes showed striking differences in regeneration and in their response to *A. tumefaciens* and *A. rhizogenes*. Previously Zhang et al., (1998) observed a large variation in regeneration frequency in *B. rapa* when testing 123 genotypes. Also Sparrow et al. (2004) showed that a number of stages within the transformation process in *B. oleracea* were under strong genetic control, suggesting that altering the tissue culture conditions alone would have only a limited effect on transformation and regeneration efficiencies of recalcitrant material. This study by Sparrow et al. (2004) led to the finding of highly significant quantitative trait loci (QTL) associated with susceptibility to *A. tumefaciens*. Therefore we decided to test if there are any QTLs associated with susceptibility to *A. tumefaciens* in *B. rapa*. Since *A. tumefaciens*-mediated transformation is very time-consuming with the low transformation frequencies that are expected when using a transformation-recalcitrant species, we considered *A. rhizogenes*-mediated hairy root transformation as an alternative to test for transformation competence. Since the virulence genes of *Agrobacterium* are conserved between *A. tumefaciens* and *A. rhizogenes* and since other important factors needed for T-DNA transfer are common to both, the analysis of plant genes regulating transformation by *A. rhizogenes* was expected to also shed light on the response to *A. tumefaciens* (Cogan et al., 2002). In chapter 3, the generation of two reciprocal F2 populations of about 200 individuals each, derived from crossing genotypes L58 and R-o-18 was described. These genotypes were chosen based on their difference in susceptibility to *A. rhizogenes* hairy root transformation, but also based on

morphological characteristics, such as plant architecture (vegetable type × oilseed type), different geographical origin, self-compatibility and rapid-cycling, to allow the rapid development of a large segregating population and the prospect to carry this on towards the generation of a Recombinant Inbred Line population. An Amplified Fragment Length Polymorphism (AFLP) platform was used to genotype the F2 population. A great advantage of this technique is that it is quick, relatively cheap and usually produces many polymorphic bands using a single PCR reaction. A disadvantage is that these markers are mostly dominant, which particularly for an F2 population, in which many loci are still heterozygous, is inconvenient for genetic mapping. Therefore Simple Sequence Repeat (SSR) markers were added to provide semi-dominant markers and to be able to assign the identified linkage groups to their respective chromosomes as presented in the *B. rapa* reference linkage map (Choi et al., 2007).

The genetic map that was made for one of the two reciprocal F2 populations was sufficiently detailed to allow QTL analysis. Initially traits related to transformation and regeneration were studied, revealing two QTLs for adventitious and hairy root production. These QTLs can be used for future selection of *B. rapa* lines that are more efficient in regeneration and transformation. In addition to these traits, the population segregated for many more traits for which QTLs could be identified. A total of 13 QTLs for nine traits including flowering time, total plant height, plant height until first open flower, number of siliques, number of seeds per silique, seed oil content, seed size, seed weight and seed colour were found. All of these traits are interesting for breeding purposes, both for vegetable crop characters (as found in L58) as for oilseed properties (R-o-18).

Yellow seededness has often been promoted for oilseed *Brassicas* as it often corresponds with significantly higher seed oil contents (Rakow et al., 1999) and the meal of yellow-seeded varieties that remains after extracting oil has higher protein and lower crude fiber contents than that of black-seeded varieties and is of better quality for use as poultry and livestock feed (Tang et al., 1997). However, surprisingly there are no yellow-seeded modern commercial varieties in many parts of the world, due to a common believe among farmers that rape seed should be black (R.J. Snowdon, pers. communication). There may be some experimental ground for that, as we found. The F2 population segregated for a strong QTL explaining 52 % of the seed colour variation mapping to chromosome A9 (chapter 2). The black seed coat allele was fully dominant over the allele for yellow seed coat. The seed coat colour QTL co-localized with QTLs for seed size, seed weight, seed oil content, number of siliques and number of seeds per silique. Seed colour and seed size were controlled by the maternal plant genotype rather than by the seed genotype. Much to our surprise, no correlation was detected between seed coat colour and seed oil content when comparing all genotypes. We hypothesized that this absence of correlation is mainly due to the complete dominance of

black seededness over yellow seededness, while seed oil content shows additive allelic effects. However, a significant positive correlation was observed between seed coat colour and seed oil content when examining the light coloured classes only. This correlation contradicts the previously found negative correlation between dark seededness and oil content, as described by Badani et al. (2006) and Rakow et al. (1999). As also black seeded genotypes in general have higher oil seed contents than yellow-seeded genotypes it may not always be advantageous to select for yellow-seededness when breeding for high seed oil content in Brassicas. Our observation also questions the presence of one gene controlling seed colour and oil content at the QTL on chromosome A9. Although further detailed genetic analysis will be needed, it may well be that there are two closely linked genes at the A9 QTL, for which favourable alleles for low fibre content due to yellow seededness and for high oil content happen to be in coupling phase in most genotypes, but not in one or both of the L58 and R-o-18 parents used to make the population.

The major drawback of using an F2 population is that it is not immortal, in the sense that F3 progeny will still segregate for many of the traits observed in the F2 and thus the population cannot be easily maintained as stable genotypes. For Brassicas this is normally achieved through the development of doubled haploid derived populations. However, these generally have the disadvantage of showing self-incompatibility and sterility hampering production of sufficient seeds, and the tissue culture phase appears to be a considerable selection bottleneck (Kole et al., 1997) often leading to strong skewedness towards the alleles of only one of the parents (Suwabe et al., 2004). Therefore instead a Recombinant Inbred Line (RIL) population was developed from this F2 population by single seed descent, as described in chapter 4. The Illumina® BeadXpress™ genotyping platform combined with the GoldenGate assay (van Orsouw et al. 2007) as recently developed by KeyGene N.V (Wageningen, The Netherlands) was used to genotype the RIL population in F7. This platform provided 86 SNP polymorphic markers that could be mapped. We subsequently added 6 publicly available SSRs also used to map the F2 population to compare both and to relate the linkage map to the *B. rapa* reference map (Choi et al., 2007). The RIL genetic map covers a total distance of 399.8 cM with an average resolution of 4.34 cM. Still 28% of the markers showed considerably distorted segregation but this is less comparing with *Brassica rapa* DH populations (more than 50%) (Wu et al., 2007; Lou et al., 2007; Suwabe et al., 2004).

A total of 26 QTLs were detected in the RIL population for 19 traits including adventitious and hairy root production, flowering time, total plant height, plant height until first open flower, leaf number until first open flower, total leaf number, cuticular wax appearance, branch number, number of siliques, silique length, silique beak length, silique carpel number, seed weight, seed colour, shattering and seed ripening. Two strong QTLs for the seed colour (SC1 and SC2) were found, explaining about 70% of the seed coat colour variation. Dark

alleles for SC1 derive from L58 and for SC2 from R-o-18. QTLs for relevant breeding traits such as carpel number, pod shattering, seed vivipary and cuticular wax were also detected in this RIL population.

Next to an important vegetable and oil seed crop, *Brassica rapa* is a valuable source of diverse health-promoting antioxidant metabolites. There appears to be a wide variation for these compounds in *B. rapa*, for which the genetic basis is largely unknown. In chapter 5 the genetic analysis of health related compounds, as well as other primary and secondary metabolites, is presented and discussed. QTLs for seed tocopherol (α-, γ-, δ-, α-/γ-, and total tocopherol) content were detected on A6 and A9 explaining a large fraction (up to 37%) of the observed phenotypic variance for seed tocopherol content. Previously Abdel-Farid et al. (2007) demonstrated the use of ^{1}H NMR spectroscopy to uncover qualitative and quantitative differences in metabolite concentrations in various cultivars of *B. rapa*, therefore this tool was used to detect elucidated metabolites, several organic and amino acids, carbohydrates, phenylpropanoids, flavonoids and glucosinolates in seedlings of the RIL population.

This yielded a large dataset of NMR spectra comprising 238 signals. The conversion of NMR signals to compounds is not easy, especially if no two-dimensional NMR spectra are available. So far, 19 compounds, distributed over organic/amino acid, sugar/glucosinolate and aromatic regions could be annotated. Out of the 238 signals, QTLs were determined for 148 signals. Among these are six glucosinolates (progoitrin and neoglucobrassicin), four essential amino acids and eight non-essential amino acids. Although this is a first screen for seedling metabolites, it shows the potential of the use of this population for further development of genetic markers suitable for breeding for improved nutritional quality in *B. rapa*. Further analysis is needed for the cloning of genes involved, which could be used for future genetic engineering of nutritional quality.

The developed RIL population is a self compatible and rapid cycling population and a useful source for genetic analysis and identification of genetic loci that can be used in future breeding programs. The results described in chapters 4 and 5 are mainly illustrating the potential of the population for genetic analysis. We did not focus in detail on traits particularly useful for vegetable or oilseed morphotypes, which will be segregating in this population considering the origin of both parents. Such traits could be shoot tenderness, resistance to particular pathogens, storability, adaptation to required growth conditions, etc. The population can also be used to directly select suitable lines as starting material for further breeding. For instance, the genotypes with better regeneration and transformation abilities could be used in future transformation experiments. Since transposon based activation mutagenesis is still an important goal to pursue in *B. rapa* they would be the genotypes of choice for transformation with improved activation tagging constructs. The genetic map of the RIL population can be further improved by adding the AFLP markers used for the F2 map.

This would enhance the resolution of the map and allows a much better comparison of traits and QTLs analysed in both populations. Further studies are therefore needed to refine the map positions and improve the correlation of the RIL genetic map to the *B. rapa* reference map. This reference map is expected to be integrated soon with the physical map based on genomic DNA sequence (http://www.brassica-rapa.org/BRGP/status.jsp). This will simplify identification of candidate genes, and facilitates the comparison to functional analyses for related genes performed in Arabidopsis, or other Brassicaceae species for which genome sequence and functional genetic information is available.

References

Abdel-Farid I.B., Kim H.K., Choi Y.H. Verpoorte R. (2007). Metabolic characterization of *Brassica rapa* leaves by NMR spectroscopy. J Agric Food Chem**.** 55:7936-7943.

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

Choi Y.H., Kim H.K., Linthorst H.J.M., Hollander J.G., Lefeber A.W.M., Erkelens C., Nuzillard J.M., Verpoorte R. (2006). NMR Metabolomics to revisit the Tobacco Mosaic virus Infection in *Nicotiana tabacum* Leaves. J Nat Prod. 69:742-748.

Cogan N.O.I., Lynn J.R., King G.J.,Kearsey M.J., Newbury H.J.,Puddephat I.J. (2002). Identification of genetic factors controlling the efficiency of *Agrobacterium rhizogenes*-mediated transformation in *Brassica oleracea* by QTL analysis. Theor Appl Genet. 105:568-576.

Kole C., Kole P., Vogelzang R., Osborn T.C. (1997). Genetic linkage map of a *Brassica rapa* recombinant inbred population. J Hered. 88:553-557.

Lou P., Zhao J., Kim J., Shen S., Pino Del Carpio D., Song X., 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.

Marsch-Martinez N., Greco R., Van Arkel G., Herrera-Estrella L., Pereira A. (2002). Activation tagging using the *En-I* maize transposon system in Arabidopsis. Plant Physiol.129:1544-56.

Moon H.S., Halfhill M.D., Good L.L., Raymer P.L., Neal Stewart C.Jr. (2007). Characterization of directly transformed weedy *Brassica rapa* and introgressed *B. rapa* with Bt cry1Ac and gfp genes. Plant Cell Rep. 26:1001-10.

Narasimhulu S.B., Chopra V.L. (1988). Species specific shoot regeneration response of cotyledonary explants of Brassicas. Plant Cell Rep. 7:104-106.

Pereira A. (1998). Heterologous transposon tagging systems. In: Lindsey K (ed). Transgenic Plant Research. Harwood Academic Publishers, UK. pp:91-108.

Pereira A. (2000). A transgenic perspective on plant functional genomics. Transgenic research. 9:245- 260.

Rakow G., Relf-Eckstein J., Raney P., Gugel R. (1999). Development of high yielding, disease resistant, yellow-seeded *Brassica napus*. 10th International Rapeseed Congress. pp:67.

Sparrow P.A., Townsend T.M., Arthur A.E., Dale P.J., Irwin J.A. (2004). Genetic analysis of *Agrobacterium tumefaciens* susceptibility in *Brassica oleracea*. Theor Appl Genet. 108:644-50.

Suwabe K., Iketani H., Nunome T., Ohyama A., Hirai M., Fukuoka H. (2004). Characteristics of microsatellites in *Brassica rapa* genome and their potential utilisation for comparative genomics in cruciferae. Breed Res. 54:85-90.

Tang Z.L., Li J.N., Zhang X.K., Chen L., Wang R. (1997). Genetic variation of yellow-seeded rapeseed lines (*Brassica napus* L.) from different genetic sources. Plant Breeding. 116:471-474.

Tsukazaki H., Kuginuki Y., Aida R.,Suzuki T. (2002). Agrobacterium-mediated transformation of a doubled haploid line of cabbage. Plant Cell Rep. 21:257-262.

van Orsouw N.J., Hogers R.C.J., Janssen A., Yalcin F., Snoeijers S., Verstege E., Schneiders H., van der Poel H., van Oeveren J., Verstegen H. van Eijk M.J.T. (2007). Complexity reduction of polymorphic sequences (CRoPSTM): a novel approach for large-scale polymorphism discovery in complex genomes. PLOS One. 11:e1172.

Wu J., Schat H., Sun R., Koornneef M., Wang X., Aarts M.G.M. (2007). Characterization of natural variation

for zinc, iron and manganese accumulation and zinc exposure response in *Brassica rapa* L. Plant Soil. 291:167-180.

Zhang F.L., Takahata Y., Xu J.-B. (1998). Medium and genotype factors influencing shoot regeneration from cotyledonary explants of Chinese cabbage (*Brassica campestris* L. ssp. pekinensis). Plant Cell Rep. 17:780-786.

Summary

Brassica rapa is an important crop with a variety of forms, and a wide distribution in the world. It is used as oil seed and vegetable crop and a valuable source of diverse health-promoting metabolites. It also can serve as a model for genetic and molecular analysis in the *Brassica* genus, to which all rapes, kales and cabbages belong, as it has the smallest genome size and some genotypes with a rapid life cycle.
Insertional mutagenesis using heterologous maize transposons has been a valuable tool for the identification and isolation of genes in Arabidopsis. Transposon-based activation tagging systems use a construct with constitutive enhancer elements that can cause transcriptional activation of flanking plant genes, which can result in dominant mutant phenotypes and subsequent isolation of the genes involved. Chapter 2 describes the action of an *En/I* activation tagging construct in *B. rapa* through *Agrobacterium rhizogenes*–mediated hairy root transformation. Successful transformation of this construct to *B. rapa* ssp. by *A. tumefaciens* was not achieved, probably due to the combination of an inefficient plant transformation and regeneration system, the length of the construct and most importantly the presence of the *SU1* gene in the construct that appears to inhibit the regeneration of transformed shoots.

As an alternative to the insertional mutagenesis approach to identify genetic loci that impact traits, there is a genetic approach based on quantitative trait locus (QTL) analysis. Segregating populations are needed to map QTLs for traits of interest. Chapter 3 describes the analysis of an F2 population derived from a cross between two distinct, but early flowering and self compatible, *B. rapa* genotypes, L58 and R-o-18. Amplified fragment length polymorphism (AFLP) markers together with simple sequence repeat (SSR) markers were used to genotype this F2 population and anchor the linkage map to the reference genetic map of *B. rapa*. Highly significant QTLs associated with the production of adventitious roots and the transformation competence to *A. rhizogenes* were detected, which will allow the selection of lines that are more efficient in transformation experiments. The analysis detected a strong QTL associated with seed coat color as well as QTLs for various morphological traits.

To fix the recombination events as much as possible and to obtain an “immortal” mapping population, a recombinant inbred line (RIL) population was developed from this F2 population. Chapter 4 describes development of this RIL population, for which a genetic linkage map was constructed using the Illumina® BeadXpress^{TM} genotyping platform of Keygene NV and additional SSR markers. Analysis revealed an additional QTL for seed coat

colour as well QTL for pod shattering, carpel number, cuticular wax and seed vivipary. Chapter 5 describes the detection of QTLs related to primary and secondary metabolites in this RIL population. The two parental lines show clear differences in metabolite profile, which allowed the finding of QTLs for glucosinolates, phenylpropanoids, glucose, glutamate and amino acids after analysis with H^1- NMR. HPLC analysis of tocopherols revealed four QTLs controlling the levels of this important antioxidant.

The information on the genetic control of health related compounds indicates the potential to improve nutritional quality in classical crop breeding programs.

Samenvatting

Brassica rapa is een belangrijk gewas met een verscheidenheid aan planttypes en een brede wereldwijde verspreiding. Het gewas wordt gebruikt voor olie- en groenteproductie en is een waardevolle bron van verschillende gezondheidsbevorderende metabolieten. Het is ook een modelsoort voor genetische en moleculaire analyse in het genus *Brassica* (met daarin alle kolen en rapen vertegenwoordigd) met het kleinste genoom en sommige genotypes met een korte generatieduur.
Insertiemutagenese met behulp van maïstransposons is een waardevol hulpmiddel gebleken voor de identificatie en isolatie van genen uit Arabidopsis. "Activation-tagging" systemen gebaseerd op transposons maken gebruik van een construct met constitutieve expressie inducerende "enhancer" elementen die de transcriptie van flankerende plant genen kunnen activeren. Dit kan resulteren in dominante mutante fenotypes, op basis waarvan nieuwe genen geïdentificeerd kunnen worden. Hoofdstuk 2 beschrijft de werking van een *En/I* activation-tagging construct in *B. rapa* na worteltransformaties met *Agrobacterium rhizogenes*. Dit construct kon niet met behulp van *A. tumefaciens* naar *B. rapa* overgebracht worden om stabiele transgene planten te krijgen. Waarschijnlijk kwam dit door de combinatie van een inefficiënt plant transformatie- en regeneratieprotocol; de lengte van het construct; en vooral door de aanwezigheid van het *SUI* gen op het construct. De aanwezigheid van dit gen lijkt de regeneratie van transgene scheuten sterk te remmen.

Een alternatief voor insertiemutagenese om genetische factoren die een effect op planteigenschappen hebben te bepalen, is een genetische aanpak gebaseerd op de analyse van genetische loci voor kwantitatieve eigenschappen ("Quantitative Trait Loci – QTLs"). Uitsplitsende populaties zijn nodig om QTLs in kaart te brengen. Hoofdstuk 3 beschrijft de analyse van een F2 populatie afkomstig van een kruising tussen twee verschillende, maar vroegbloeiende en zelfcompatibele, *B. rapa* genotypes, L58 en R-o-18. "Amplified Fragment Length Polymorphism" (AFLP) merkers en "Simple Sequence Repeat" (SSR) merkers zijn gebruikt om deze F2 populatie te genotyperen en om de genetische koppelingskaart van de populatie te verankeren aan de genetische referentiekaart van *B. rapa*. Na analyse zijn enkele QTLs gevonden met een hoge statistische betrouwbaarheid voor associatie met de productie van adventiefwortels en transformatiecompetentie voor *A. rhizogenes*. Hiermee kunnen in de toekomst lijnen geselecteerd worden die efficiënter zijn in transformatie-experimenten. Verder is een sterke QTL gevonden voor zaadhuidkleur en QTLs voor diverse morfologische kenmerken.

Om de verschillende recombinatiegebeurtenissen in deze F2 populatie zo veel mogelijk vast te leggen en om zodoende een “onsterfelijke” populatie voor genetische analyse te krijgen, is een zogenaamde recombinante inteeltlijn (RIL) populatie ontwikkeld. Hoofdstuk 4 beschrijft deze ontwikkeling. Voor de constructie van een genetische kaart van deze populatie is gebruik gemaakt is van het Illumina® BeadXpress™ genotyperingsplatform van Keygene NV, aangevuld met SSR-merkers. Na analyse van de gegevens is een nieuwe QTL voor zaadhuidkleur gevonden, evenals QTLs voor voortijdige zaadval, het aantal vruchtbladen, de bladwaslaagkleur en voortijdige zaadkieming. Hoofdstuk 5 beschrijft de vondst van QTLs gerelateerd aan primaire en secundaire metabolieten in deze RIL populatie. De twee ouderlijnen vertonen duidelijke verschillen in metabolietprofiel, waardoor het mogelijk was om met behulp van H^1- NMR QTLs te vinden voor de gehaltes aan glucosinolaten, fenylpropanoiden, glucose, glutamaat en aminozuren. HPLC-analyse van tocoferolen leverde vier QTLs op betrokken bij de productie van deze belangrijke antioxidant.

De informatie over de genetische controle van de gehaltes aan gezondheidsbevorderende stoffen is illustratief voor de mogelijkheden om met behulp van klassieke veredelingsprogramma’s de voedingskwaliteit van gewassen te verbeteren.

Acknowledgements

I am glad to acknowledge many people who helped me to reach this step here. Without their help, support and advices it would have been impossible to finish this work.
First of all I would like to thank my promoter, Prof. dr. ir. Maarten Koornneef for his support, valuable advices and critical reading of my manuscripts.
I would like to appreciate my supervisour, Dr. Mark Aarts for his guidance, the patient supervision, useful discussions, and critical reading of the manuscript. Dear Mark: I learned a lot form you. I enjoyed the work discussions we had weekly in your research group. I would like to thank your family also for the kind hospitality.
I would like to thank Prof. Andy Pereira for his supervision in transposon tagging and providing the activation tagging constructs.
Special thanks to Corrie Hanhart for her valuable help with population growing, harvesting, phenotyping and specially with translating the Dutch letters.
I would like to thank all the members of the Genetic laboratory, memebers who left already: Sangita Talukdar, Judith van de Mortel, Artak Ghandilyan, Ana Assuncao, Wessel van Leeuwen, Jian Wu, Juan Du, Wu Huilan, Vanessa Stein, and the current members: Zeshan Hassan, Ya-Fen Lin, Mohamed El Soda, Ana Carolina Atala Lombelo Campos, Charles Neris Moreira, Inge van Oorschot, Padraic Flood, Hetty Blankestijn and Frank Becker. Dear Wessel and Artak, thanks for your valuable advices throughout my study.
I also had the pleasure to be in the same office with Diana Zuluaga Agudelo, Tjoa Aiyen, and Duy Nguyen in new Radix building. I am glad that we could interact and help each other nicely.
I would like to thank the memebers of Plant Phisiology group: Dr. Dick Vreugdenhil, Dr. Joost Keurentjes (especially for his useful ideas and discussions), Dr. Wilco Ligterink, Diaan Jamar, Jochem Eigenhuijsen, Lidiya Sergeeva, Myriam Olortegui Guzman, Rashid Kazemi, Noorullah Khan, and Ronny Joosen.
I would like to thank Casper Pillen and Taede Stoker, especially Casper for watering and taking care of my plants in greenhouse.
I would like to thank people in Plant Breeding department: Dr. Guusje Bonnema, Dunia Pino Del Carpio, Jianjun Zhao, and specially Johan Bucher for his help with LI-COR system to run molecular markers.
I would like to extend my gratitude to secretaries: Aafke van der Kooi, Corrie Eekelder and Wytske Nijenhuis for their administrative help.
Having some Iranian friends in abroad is a valuable opportunity. I would like to thank the Iranian community in Wageningen for creating a friendly atmosphere here. My special thanks to Reza, Mohammad and Benyamin and their families for nice talks, discussions and trips we had together.

I would like to thank the Ministry of Science, Research and Technology of Iran and the IOP Genomics project IGE050010 on Brassica Vegetable Nutrigenomics for the financial support.

My sincere thanks to my parents, father and mother-in-law, brothers and sister for their continuous emotional supports. My deepest gratitude is to my mother and father.
Finally, my special gratitude to my wife Taeibah for her helps and supports and my daughter Mozhdeh for creating precious moments in my life.

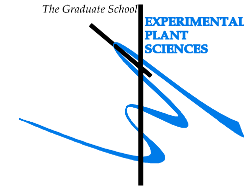
Hedayat Bagheri
Wageningen, December 2009

About the author

Hedayat bagheri was born March 26, 1973 in Meybod-Yazd, Iran. After completing his high school, he studied Plant Agronomy in Isfahan University of Technology (Isfahan-Iran). He continued his studies at Tarbiat Modares University (Tehran-Iran) where he obtained his MSc in the field of Plant Breeding. In 2001 he was employed at BU Ali Sina University in Hamadan-Iran. He was awarded a scholarship from the Ministry of Science, Research and Technology of Iran in 2005 to do a PhD abroad. In June 2005 he started his PhD program at Wageningen University, Lab of Genetics. This dissertation presents the results of his PhD study on genetic analysis of breeding-related traits in *Brassica rapa*.

Education Statement of the Graduate School

Experimental Plant Sciences



Issued to: **Hedayat Bagheri**
Date: **15 December 2009**
Group: **Laboratory of Genetics, Wageningen University**

1) Start-up phase	date
► **First presentation of your project**	May 15, 2006
Construction of an activation tagging system in *B. rapa*	
► **Writing or rewriting a project proposal**	Mar 2006
Development of a transposon-based activation tagging population of *Brassica rapa*	
► **MSc courses**	
Gene Technology MOB-20306	Aug 2005
GATC GEN-30306	Dec 2005
► **Laboratory use of isotopes**	
Safe handling of radioactive material (level 5B)	Oct 03-05, 2006
Subtotal Start-up Phase	*12.0 credits**

2) Scientific Exposure	date
► **EPS PhD student days**	
EPS PhD student day, Wageningen University	Sep 19, 2006
EPS PhD student day, Wageningen University	Sep 13, 2007
PhD Student Retreat (EPS, SDV & IMPRS), Wageningen	Oct 02 & 03, 2008
► **EPS theme symposia**	
Theme 3 symposium 'Metabolism and Adaptation' , University of Amsterdam	Nov 10, 2006
Theme 3 symposium 'Metabolism and Adaptation' , University of Amsterdam	Feb 18, 2009
► **NWO Lunteren days and other National Platforms**	
NWO-ALW Experimental Plant Sciences, Lunteren	Apr 03-04, 2006
NWO-ALW Experimental Plant Sciences, Lunteren	Apr 02-03, 2007
NWO-ALW Experimental Plant Sciences, Lunteren	Apr, 2008
NWO-ALW Experimental Plant Sciences, Lunteren	Apr 06-02, 2009
NVBMB Spring Symposium, Wageningen	May 18, 2007
► **Seminars (series), workshops and symposia**	
6 seminars	2005-2006
Diversification of small RNA pathways in plants, prof.J.Carrington	Mar 26, 2007
physical cell biology	Feb 09, 2007
comparative genomics of barley and rice:, Prof:A.Graner	May 10, 2007
Building plant walls, F.Assaad	Feb 16, 2007
Studing microtubule organization, Dr. Marcel Janson	Aug 30, 2007
From molecular to morphology, Dr. Anja Geitmann	Sep 21, 2007
Regulation of Phase change in plants by MiRNAs, Prof. Scott Poethig	Sep 24, 2007
Coordination of cell polarity formation in plant tissues, Prof. Zhenbiao Yang	June 23,2008
Small RNAs in abiotic stress, Jian-Kang Zho	Nov 03, 2008
Seeds, microRNA and Darwin?, Dr. Hiro Nonogaki	Sep 17, 2009
► **Seminar plus**	
► **International symposia and congresses**	
15th Crucifer Genetics Workshop, Wageningen (NL)	Sep 30-Oct 04, 2006
Molecular mapping and marker assisted selection in plants, Austria	Feb 03-06, 2008
► **Presentations**	
Poster: 15th Crucifer Genetics Workshop, Wageningen (NL)	Sep 30-Oct 04, 2006
Poster: Lunteren days	Apr 02-03, 2007
Oral presentation: Molecular mapping and marker assisted selection in plants,Austria	Feb 03-06, 2008
Poster: Natural Variation in Plants, Wageningen	Aug 26-29, 2008
Poster: PAGXVII, San Diego	Jan 10-14, 2009
► **IAB interview**	Sep 14, 2007
► **Excursions**	
Subtotal Scientific Exposure	*12.3 credits**

3) In-Depth Studies	date
► **EPS courses or other PhD courses**	
Gateway to Gateway Technology	Nov 20-24, 2006
System biology course: statistical analysis of ~omics data	Dec 11-14, 2006
Statistics for the life science	May 30-Jun 05, 2007
Natural Variation in Plants, Wageningen	Aug 26-29, 2008
► **Journal club**	
Member of a literature discussion group at Genetics	2004-2009
► **Individual research training**	
Subtotal In-Depth Studies	*8.1 credits**

4) Personal development	date
► **Skill training courses**	
Academic Writing II	Sep-Nov 2005
WGS - PhD competence assessment	Feb 2006
WGS Project and Time Management	Oct-Dec 2006
► **Organisation of PhD students day, course or conference**	
► **Membership of Board, Committee or PhD council**	
Subtotal Personal Development	*4.2 credits**

TOTAL NUMBER OF CREDIT POINTS*	**36.6**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 credits

** A credit represents a normative study load of 28 hours of study*

The research described in this thesis was performed in the Lab of Genetics of Wageningen University and financially was supported by the Ministry of science, Research, and Technology of Iran and the IOP Genomics project IGE050010 on Brassica Vegetable Nutrigenomics.

Thesis layout and cover design: by the author
Front page: zero in on the *B. rapa*. *B. rapa* is a valuable source of diverse metabolites which can be studied by marker analysis to find related genes.
Back page: seed coat color in *B. rapa*. seed coat color is controlled by the maternal plant genotype and its variation in F2 and RIL can be mapped with QTL analysis.

Printed by: Wohrman Print service, Postbus 92, 7200 AB, Zutphen, The Netherlands./380