The genetics of phytate content and morphological traits in *Brassica rapa*

Jianjun Zhao

Promotor: Prof. dr. ir. M. Koornneef
Persoonlijk hoogleraar bij het laboratorium voor Erfelijkheidsleer
Wageningen Universiteit

Co-promotor: Dr. ir. A.B. Bonnema
Universitair docent bij het Laboratorium voor Plantenveredeling
Wageningen Universiteit

Dr. D. Vreugdenhil
Universitair hoofd docent bij het Laboratorium voor Plantenfysiologie
Wageningen Universiteit

Promotiecommissie: Prof. dr. ir. E. Jacobsen, Wageningen Universiteit
Prof. dr. W.J. Stiekema, Wageningen Universiteit
Dr. ir. R.E. Voorrips, Wageningen Universiteit
Prof. dr. X. Wang, Chinese Academy of Agricultural Sciences,
Beijing, China

Dit onderzoek is uitgevoerd binnen de onderzoeksschool 'Experimental Plant Sciences'

Jianjun Zhao

The genetics of phytate content and morphological traits in *Brassica rapa*

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit
Prof. dr. M. J. Kropff
in het openbaar te verdedigen
op maandag 22 januari 2007
des namiddags te 13:30 in de Aula

The genetics of phytate content and morphological traits in *Brassica rapa*

Jianjun Zhao

PhD thesis Wageningen University, The Netherlands, 2007
-with references-with summary in English, Dutch and Chinese.

ISBN 90-8504-588-6

Contents

Chapter 1

General introduction

1 History and botanical description of *Brassica rapa*

Brassica rapa (syn. *Brasscia campestris*) belongs to the Brassicaceae (Cruciferae) family. The *Brassica* genus comprises a number of economically important species among which the three elementary diploid species are *B. rapa* (2n=20; genome composition AA), *B. nigra* (2n=16; genome composition BB) and *B. oleracea* (2n=18; genome composition CC), and the three amphidiploids *B. juncea* (2n=36; genome composition AABB), *B. napus* (2n=38; genome composition AACC) and *B. carinata* (2n=34; genome composition BBCC). The latter three species originated through interspecific hybridization between any two of the three diploid species. The relationship between the different genomes is clearly outlined in the well-known "Triangle of U" (U 1935). A hypothetic scheme of genomic relations of *Brassica* and related genera is summarized by Song et al. (1988a, 1990) suggesting that species with the chromosome number n=7 represent the most ancient groups which are probably derived from the prototype species with n=6 (Prakash and Hinata 1980). For comparison of synteny between members of the Brassicaceae an ancestral karyotype with n=8 was recently proposed (Schranz et al. 2006).

Being the first domesticated *Brassica* species, *B. rapa* has been cultivated for many centuries from the Western Mediterranean region to Central Asia, and is still present throughout this area (Gomez Campo 1999). During the long history and breeding for different traits abundant morphological and genetic variation has been formed and a large number of subspecies is recognized at present.

Within *B. rapa*, various schemes and names have been proposed for categorizing different morphotypes (Gomez Campo 1999; Diederichsen 2001; CAAS-IVF 2001; http://www.plantnames.unimelb.edu.au/Sorting/Brassica.html). However, many aspects of the phylogeny within the species are not fully understood and an internationally accepted nomenclature of the subspecies or cultivar groups seems necessary. In general the classification is based on morphological appearance, resulting of a division of the cultivated forms of *B. rapa* into three main groups: turnip, oil, and leafy types.

A number of studies based on morphology, geographic distribution, isozymes and molecular data indicate that *B. rapa* originates from two independent centers (Gomez Campo 1999). Europe is proposed as one primary center of origin for oil and turnip types, which were further developed in Russia, Central Asia and the Near East. East Asia is proposed as another primary center of origin for Indian oil types and Chinese leafy vegetables. Other cultivar groups of *B. rapa* most likely originated from different morphotypes within the two centers of origin and subsequently evolved separately.

Turnip (ssp. *rapa* or *rapifera*) is a biennial crop with an enlarged hypocotyl and taproot, which varies widely in shape and colour. It is a very old *B. rapa* sub-species and was probably directly domesticated from the wild progenitor in Europe (Reiner et al. 1995). In Europe, it has been cultivated since 2500-2000 BC and it spread to Asia after 1000 BC (De Candolle 1886). Many distinct types were known to the Romans at the beginning of the Christian era and some of those varieties bore Greek place names, indicating turnips were earlier cultured in the Roman Empire and Ancient Greece. Turnips were introduced to China and cultivated before Christ based on the Chinese book of poetry "Shih Ching" (Keng 1974). The names and cultivation methods are also mentioned in some old Chinese books, indicating that turnips were commonly consumed as vegetables and were cultivated during the Han dynasty (202 BC-220 AD) in China. Nowadays Chinese turnips are often replaced by other vegetables and its cultivation area is reduced. A main Italian group of cultivars (Broccoletto, Broccoli raab, Cima di rapa or ruvo) of which the young inflorescences are consumed is regarded as a turnip-tops form within this group.

Oleiferous *B. rapa* (spp. *oleifera*) is mainly cultivated in Europe, China, India, and Canada, and there is potential for the crop to be successfully grown in the United States, South America and Australia. It is believed that European forms and Asian types have different origins, involving the Mediterranean area and the region of Central Asia, Afghanistan and the adjoining Indian subcontinent. There is a lot of evidence that European oilseed *B. rapa* is genetically very close to the turnip type (Reiner et al. 1995). Domestication is believed to have occurred in the early middle ages in Europe. Three Indian oleiferous *B. rapa* ecotypes, viz. Brown Sarson, Toria and Yellow Sarson, have been developed probably in isolation from European and Chinese cultivars for many centuries (Gomez Campos 1999). In China, three main ecotypes, viz. spring, winter and semi-winter turnip rape, were developed in adaptation

to different climates, soil conditions, cultivation methods and farmer preferences (He et al. 2003), and have been grown as food oil and vegetable crop. The history of Chinese oleiferous *B. rapa* domestication in China needs to be further clarified although evolutional pathways have been proposed by some Chinese researchers, in which the common point is that it possibly derived from Chinese Pak choi (Liu 1984; Cao et al. 1997). This is also supported by our result (Zhao et al. 2005 and chapter 2 of this thesis). Recently, the cultivated oleiferous *B. rapa* in China was substantially replaced by the recently introduced *B. napus* cultivars, which have a higher yield and better adaptation. A group of winter oil types from Pakistan is characterized by self-incompatibility and dark seed, which is not directly related to either East Asia or European types (Zhao et al. 2005).

A large group of *B. rapa* is formed by the leafy vegetables differentiated into several subspecies or cultivar groups mainly from China and Japan. Within this group, Chinese Pak choi (ssp. *chinensis*) with green-white midrib is likely to be the most ancient form (Li 1981; Song et al. 1988b). Chinese cabbage (ssp. *pekinensis*) is native to China and is characterized by larger leaves and heads of different shape. Two main hypotheses regarding its origin exist in China: one is the hybridization hypothesis suggesting that Chinese cabbage originated from hybridization between turnip (or turnip rape) and Pak choi (Li 1981). The loose-leaved type is the ancestral form and gradually developed into the heading form, which was selected for as an adaptation to cool temperatures. The other evolutionary hypothesis was proposed by Tan (1979), who suggested that Chinese cabbage was formed during the introduction of Pak choi from southern to northern China. Chinese cabbage cultivated forms appear later than Pak choi in ancient Chinese literary records. During the Tang dynasty (AD 659), the loose-leaved Chinese cabbage was mentioned in the book of 'Xin Xiu Ben Cao'. The pictures of semi-heading types appeared in the book of 'Yin Shan Zheng Yao' in the Yuan dynasty (1330 AD). The heading Chinese cabbage originated between the Yuan and the Ming dynasty (1368-1644 AD), and became especially popular during the Ming dynasty. In the 15th century, Chinese cabbage was introduced to Korea as a staple vegetable for making kimchi. At present the Chinese cabbage is commonly found in markets throughout the world. Several other leafy types such as Wutacai (ssp. *narinosa*) with flat rosettes and dark-green leaves, Zicaitai (ssp. *chinensis* or *purpuraria*) with purple red stem, Taicai's (or Tai tsai's) with irregularly notched leaves and Caixin (or Caitai, ssp. *parachinensis*), an early flowering Pak choi are also

consumed locally in China. The origin of these types has been discussed but is unclear up to now (Cao et al. 1997). Another small group of Japanese leafy vegetables includes Mizuna with serrated leaves and Mibuna with long narrow leaves (*nipposinica* or *japonica* group), and Komatsuna consumed for young leaves, stalks and flower shoots (Japanese mustard spinach, *perviridis* group). Various Japanese vegetables are likely to be derived directly or indirectly from different types of Pak choi, but have diverged through geographic isolation and intensive selection (Song et al. 1988b, 1990). In our experiment, Chinese Shuicai accessions that resemble Mizuna form no clearly separate cluster and group in the Pak choi cluster (Zhao et al. 2005), while the Japanese Mizuna's form a distinct group.

2 Economic importance and breeding of *Brassica rapa*

2.1 Economic importance

Brassica species play an important role in agriculture and horticulture, as well as contributing both to the economy and health of populations around the world. *B. rapa* has worldwide importance in agriculture, providing many vegetables together with *B. oleracea* and oil products together with *B. napus*. Typically, the growing range of *B. rapa* extends to coastal lowlands, plateaus, hills, and mountain areas up to 2300 m (Warwick and Francis 1994).

Turnips are a nutritious root vegetable and well adapted to the northern parts of the United States, Europe and Canada, because it grows well in temperate climates and can be stored for several months after harvest. According to its utilization, turnip type includes turnips with enlarged root, turnip greens of which leaves are consumed and turnip tops cultivated for their numerous flowering stalks (Padilla et al. 2005). Turnip greens and turnip tops are used as vegetables for culinary use. Turnips are cultivated for vegetable use but have also traditionally been used as fodder and forage crops.

The attractiveness of oilseeds types is that they are reservoirs of oils and proteins. Oil is used for cooking, salad or margarine for human consumption, and meal with valuable protein is used for livestock feed. Some work has also explored the preparation of protein isolates and concentrates for human consumption. Canada is one of the four regions with the highest oilseed (*B. rapa* and *B. napus*) production. In the 1970s, 75% of rapeseed area in Canada was of spring *B. rapa* cultivars, later in 1990s this proportion decreased. Currently *B. napus* and *B. rapa* make up 90% and 10%, respectively of the oilseed rape (canola) grown in western Canada (Warwick et al. 2002). Although *B. rapa* varieties are somewhat lower in yield than

B. napus varieties they have shorter growing period and are more suited to the northern growing area in Western Canada. Similarly *B. napus* is the most important oilseed crop in Europe, although *B. rapa* is also grown in northern Europe. Oil from rapeseed is the basis of industrial applications for margarines and other edible products. Rapeseed oil is also used as fuel (Bbiodiesel) (http://www.brassica.info: Draft "White Paper"). Besides *B. juncea*, *B. rapa* is one of the two traditional oilseed crops in the Indian subcontinent. In China, *B. rapa* was the traditional species for oilseed production with the largest cultivation area before introduction of *B. napus* in the 1940s. However, the short growing period makes *B. rapa* still an optimal choice in some areas, accounting for about 15% acreage of oilseed *Brassica* in recent years (He et al. 2002). Chinese oleiferous *B. rapa* could make great contributions to the improvement of *B. napus* because of its abundant genetic resources and good agronomic traits like short life cycle, high oil content and self-incompatibility. Recently, efforts have been made to broaden the genetic basis of rapeseed (*B. napus*) by introgression of Chinese oil *B. rapa* (Qian et al. 2006).

In vegetable *Brassicas*, levels of useful nutritional components are notably high and contribute to a healthy human diet, being a valuable source of dietary fibre, vitamins (A, C and E), potassium and other health-enhancing factors such as anticarcinogenic compounds (some glucosinolates and folate) (http://www.brassica.info: Draft "White Paper"). *B. rapa* vegetables are one of the most important vegetables in eastern Asia, where Chinese cabbage is ranking fist in annual vegetable production in China, especially in the north. Since cabbages can be stored for extended periods, it is the main vegetable consumed in winter. In China, the annual cultivation area of heading Chinese cabbage and non-heading Pak choi is about 1.3 million hectare, accounting for around 30% of nationally supplied vegetables each year. In Korea, *Brassica* vegetables are used as major components of “kim-chi”, the traditional preserved recipe and salad. Chinese cabbage is therefore one the most important *Brassica* crops in Korea.

2.2 Breeding of B. rapa

In *Brassica* breeding systems, Doubled Haploid (DH) technology has been widely applied to generate inbred lines and self-incompatibility (SI) has been used successfully to produce F1 hybrids. Cytoplasmic male sterility (CMS) is increasingly used in hybrid production. DNA marker technology can speed up the traditional breeding programs, but its use is still limited.

In general breeding of *B. rapa* aims at increasing yield, improving agronomic characteristics and improving quality. In oil types, one important task in breeding programs is to increase seed oil content and seed yield, although it is difficult to achieve these two simultaneously. In *B. rapa* vegetables breeding programs have different objectives and priorities since each vegetable type is characterized by its own characteristics. The market demands are considered by breeders in designing the most desirable ideotype, like Chinese cabbage with ovate or cylindrical heads favored in different geographical regions. Bolting resistance is an important breeding aim to enable year round heading Chinese cabbage production. Disease resistant varieties are also very much needed. Clubroot, caused by *Plasmodiophora brassicae*, is one of the most damaging diseases in *Brassica* crops because the majority of commercial *B. rapa* cultivars is very susceptible, which implies that breeding for resistance has a high priority.

2.2.1 Nutritional quality

Recently improved nutritional quality of *B. rapa* products has become an important selection criterion to globally improve the living standard. Improvement of fatty acid composition and increases of tocopherol (vitamin E, an antioxidant) content of oilseeds has been a target for breeding. In addition to oil quality, improvement of meal (cake) quality is currently in focus. A high nutritional value of *Brassica* meal resulting from a high energy and protein content and a favourable amino acid composition is restricted by its content of glucosinolates, tannins, phenolic acids and phytate, which are referred to as anti-nutritional compounds. A wide range of *Brassica* species and varieties is also used as vegetables, and provide a useful resource for phosphate and other minerals such as calcium, magnesium, potassium, iron and zinc. It is desirable that breeders pay attention to the nutritional composition of vegetables to meet human consumption. However, in a largely vegetarian diet, the micronutrient content of vegetables and the content of minerals and of compounds such as phytate that absorb minerals determine to a large extend the amount of bio-available micronutrients.

2.2.2 Phytate

Till now, little information is available about the phytate content in studies of anti-nutritional elements in *Brassica*. Phytic acid, *myo*-Inositol-1,2,3,4,5,6-hexakisphosphate (IP6), is ubiquitous in eukaryotic species. This compound especially accumulates in seeds or grains, in which it can represent from 1.0 to several percent of seed dry weight and about 65-85% of seed total phosphorus (Raboy 2001, 2003). It might play a number of roles in cells, which

includes a storage function as a major storage form of phosphorus. Furthermore it is a major metabolic pool in the inositol phosphate and pyrophosphate pathway, it provides energy for ATP regeneration, RNA export, affects DNA repair, and can act as anti-oxidant (Raboy 2003). Phytic acid is considered to be an anti-nutritional substance because the highly negatively charged phosphates in IP6 form a complex (phytate) with cations (potassium, magnesium, iron and zinc) that are therefore not bioavailable resulting in micronutrient (iron and zinc) deficiencies in animals and humans. Moreover, a high level of phytate in plant tissue can cause phosphorus pollution of the environment. The high level of phytate in the modern "double low" (<2% erucic acid in the oil; <30 micromols/g glucosinolate in meal) varieties implies that the utility of seed meal for animal feed is limited (Peng et al. 2001). Furthermore, micronutrient deficiency is a serious problem in large parts of the world: nearly 20% of the population suffers from iron deficiency anemia in China (Du et al. 2000) and phytate in the diet of people impairs the bioavailability of iron and calcium (Ma et al. 2005).

In order to solve above-mentioned problems, breeding for low phytate accumulation but also for higher micronutrient content is a possible solution, which was only recently realized (Raboy 2001). Several low phytate mutants have been produced in cereal crops (Raboy 2001, 2003). For breeding of this trait it is relevant that a survey is made of the genetic variation present in the available germplasm. Phytate concentration in seeds or leaves varies widely between cultivars or accessions in *B. napus* (Mollers et al. 1999) and Arabidopsis (Bentsink et al. 2003). It is likely that considerable genetic variation for phytate also exists in *B. rapa*.

3 Genetic and genomic research in *B. rapa*

The genome size of *B. rapa* is about 529 Mb per haploid genome, which is smaller than the genomes of *B. oleracea* (696 Mb) and *B. nigra* (632 Mb). Several international research groups working with *Brassica* species are brought together under the banner "Multinational *Brassica* Genome Project" (http://www.brassica.info) in order to further develop our knowledge about the genomes of *Brassica* species.

3.1 Genetic maps

The development of genetic maps in *Brassica* is essential to understand the origin and relationship among the genomes of the diploid cultivated *Brassica* species and can be utilized in applied genetics and breeding of *Brassica* crops. Genetic maps have been generated for all *Brassica* species, in which most effort is recently being focused on *B. rapa*. More than twenty

maps (summarized in Table 1) have been developed for this species independently in different laboratories, often involving crosses between different cultivar groups. Most of the maps were constructed by means of RAPDs (Random Amplified Polymorphic DNA), AFLPs (Amplified Fragment Length Polymorphism) and RFLPs (Restriction Fragment Length Polymorphisms) markers. SSRs (also called Microsatellite Sequence Repeats) represent a valuable tool as anchor molecular markers linking different maps and as markers for characterizing germplasm in *Brassica* species due to their high polymorphism rate. Large investments have been made in the development of *Brassica* SSRs, a subset of which is available to the scientific community (http://www.brassica.info/ssr/SSRinfo.htm). More recently, SSRs have been used to establish genetic maps in *B. rapa* with linkage groups assigned to the internationally agreed chromosomal nomenclature of *B. rapa,* R1-R10. Lim's research group cabbage inbred lines, Chiifu and Kenshin (Choi et al. 2004). A genetic linkage map was constructed based on many DNA markers segregating in this population, where 644 markers were mapped on 10 linkage groups covering 1131 cM with an average distance of 1.8 cM (Choi et al. 2004). This map will serve as useful reference to undertake physical mapping and genome sequencing of *B. rapa* under the aegis of the Multinational *Brassica* Genome Project (http://www.brassica.info). A detailed *B. rapa* linkage map using sequenced EST clones derived from tissue-specific libraries of *B. rapa* containing 544 sequence tagged loci covering 1287 cM, with an average mapping interval of 2.4 cM, has been established (Kim et al. 2006). Suwabe et al. (2006) developed another SSR-based linkage map identifying genes controlling clubroot resistance. In the two maps, anchored SSR markers are applied to assign the linkage groups to the agreed chromosome nomenclature R01-R10.

3.2 Genetic mapping

One of the most important applications of genetic maps is to identify markers associated with important qualitative and quantitative agronomic traits, which may assist breeders to make more efficient selections in breeding programs. A number of molecular markers or chromosome regions linked to traits in *B. rapa* have been identified during the past years. Much attention has been paid to the content of fatty acids, disease resistance related traits, and important morphological and physiological traits. Information on genetic mapping of several traits in *B. rapa* is summarized in Table 2.

Table 1 Summary of genetic linkage maps in *B. rapa*

Population (parent)	Marker type	Marker number	Map distance (cM)	Number of linkage group	Reference
F2 (Chinese cabbage, Spring broccoli)	RFLP	280	1850	10	Song et al. 1991
F2 (Turnip, Pak choi)	RFLP Isozyme	49	262	8	McGrath and Quiros 1991
F2 (Sarson, Canola)	RFLP	360	1876	10	Chyi et al. 1992
F3 (Turnip rape, Yellow sarson)	RFLP	139	1785	11	Teutonico and Osborn 1994
F2 (Chinese cabbage, spring broccoli)	RFLP	220	1593	10	Song et al. 1995
F2 (Chinese cabbage, Chinese cabbage)	RAPD Isozyme	117	860	16	Ajisaka et al.1995
F2 (Turnip rape, Turnip rape)	RAPD AFLP	166	519	10	Tanhuanpää et al. 1996a
RI (Chinese cabbage, Chinese cabbage)	RFLP	83	1138	10	Novakova et al. 1996
RI (Turnip rape, Yellow sarson)	RFLP	126	821	12	Kole et al. 1996
RI (Turnip rape, Yellow sarson)	RFLP	144	890	12	Kole et al. 1997
F2 (Chinese cabbage, Mizuna)	RAPD	52	733	10	Noziki et al. 1997
F2 (Chinese cabbage, Chinese cabbage)	RFLP	63	735	10	Matsunomoto et al.1998
F2 (Turnip, Chinese cabbage)	RAPD	99	1632	13	Zhang et al. 2000
F2 (Turnip, Pak choi)	AFLP RAPD	131	1811	12	Lu et al. 2002
RI (Chinese cabbage, Chinese cabbage)	AFLP RAPD	352	2666	17	Yu et al. 2003
DH (Chinese cabbage, Chinese cabbage)	AFLP	255	884	10	Wang et al. 2004
DH (Chinese cabbage, Chinese cabbage)	SSR RFLP RAPD	262	1005	10	Suwabe et al. 2004
DH (Chinese cabbage, Chinese cabbage)	AFLP RFLP ESTP CAPs SSR	644	1131	10	Choi et al. 2004
F2 (Chinese cabbage, Chinese cabbage)	RAPD RFLP SSR	262	1005	10	Suwabe et al. 2006
F2 (Chinese cabbage, Chinese cabbage)	RFLP SSR	544	1287	10	Kim et al. 2006

In order to improve the oil quality and quantity in *B. rapa*, studies have been undertaken to generate markers linked to genes for various fatty acids such as linolenic, oleic and erucic acids. Among disease resistance traits, much research is focused on clubroot resistance and some related genes (*Crr1*, *Crr2*, *Crr3* and *Crr4*) are identified (Hirai et al. 2004; Suwabe et al. 2006). Another important agronomic trait in *Brassica* is the colour of the seed coat, since it

relates to oil and seed meal quality and quantity: yellow-seeded lines have low fibre, more protein and higher oil content. The seed coat in *B. rapa* varies from yellow, brown to black, and various genetic patterns have been proposed. Mapping of this trait has mainly focused on *B. napus* (Liu et al. 2005) and *B. juncea* (Mahmood et al. 2005; Padmaja et al. 2005) and revealed that it was controlled by one to three genes. In *B. rapa,* Teutonico and Osborn (1994) reported that yellow seed colour segregated as a maternally inherited recessive trait, and the locus controlling yellow seed colour (Yls) was mapped to LG5, which linkage group was not yet assigned to an R group. *Brassica* exhibits a self-incompatibility system and the recognition specificities of the pollen and the stigma are determined by a single polymorphic locus (S-locus). Considerable studies indicated that there are three highly polymorphic genes (*SRK, SP11/SCR*, *SLG*) at the S-locus (Shiba et al. 2004; Fujimoto et al. 2006). A series of loci affecting the microspore culture efficiency in *B. rapa* were mapped (Ajisaka et al. 1999; Zhang et al. 2003).

Quantitative Trait Loci (QTL) analysis has been used to identify genes related to a wide range of developmental and morphological traits in *B. rapa*. Flowering time is a very important developmental trait and wide variation exits among *B. rapa*. Many QTL for flowering-time related genes of *B. rapa* have been reported (Teutonico and Osborn 1995; Osborn et al. 1997), like three QTL *VFR1*, *VFR2* and *VFR3* related to vernalization response, and three QTL *FR1*, *FR2*, *FR3* not related to vernalization response. A major locus, '*VFR2*' has been shown to correspond to a *B. rapa* homolog of the Arabidopsis *FLC* (*Flowering Locus C*) gene, encoding a MADS domain and being a dosage-dependent repressor of flowering (Kole et al 2001). Schranz et al. (2002) further cloned four *B. rapa FLC* homologues (*BrFLC1, BrFLC2, BrFLC3, BrFLC5*). Three of the *FLC* homologues co-segregated with the flowering time QTL *FR1* (*BrFLC2*), *FR2* (*BrFLC5*) and a qualitative flowering time locus *VFR2* (*BrFLC1*) in populations derived by backcrossing late-flowering alleles from a biennial parent into an annual parent. The three *BrFLC* genes *BrFLC2*, *BrFLC3* and *BrFLC1* were mapped to position on R02, R03 and R10 (Schranz et al. 2002; Kim et al. 2006). In the later study an SSR marker derived from *MAF* (MADS *Affecting Flowering*) was mapped to the long arm of R02 and appears to correspond to another vernalization response QTL *VFR1* (Kim et al. 2006). Bolting time, a flowering associated trait, is agronomically important for vegetable types because it affects the yield and quality of products. Ten possible QTL located on six

linkage groups have been detected in a Chinese cabbage DH population (Nishioka et al. 2005). It can be concluded that multiple functional loci are involved in the variation of flowering time in *B. rapa*. For other complex morphological traits, few studies using QTL mapping are reported. A dwarf gene, *DWF2*, was mapped at the end of linkage group R06 (Muangprom and Osborn 2004). Forty-eight QTL determining twenty-eight phenotypic traits related to flowering (days to bud and flower, plant height), leaf (pubescence, length, lobes and petiole characteristics) and stem (stem length and index) traits were detected (Song et al. 1995). Yu et al (2003) identified 50 QTL based on a RI (Recombinant Inbred) population, including 5 for plant growth habit, 6 for plant height, 5 for plant diameter, 7 for leaf length, 4 for leaf width, 6 for leaf length/leaf width ratio, 7 for petiole length, 4 for petiole width and 6 for bolting character. Twenty-one polygenic traits including yield and morphological attributes (head weight, leaf blade width, head compactness and head length) were studied using QTL analysis in a South Korean Chinese cabbage DH population (Choi et al. 2004). Using a F2 population derived from a cross between Pak choi and turnip, 24 QTL affecting eight traits of agronomic characteristics in shoots were identified by Lu et al. (2002). Another interesting trait is turnip formation, which maybe controlled genetically and develops under suitable conditions (Cao et al. 1997, Chapter 5 of this thesis). One AFLP marker EM220 on LG 4 related to root swelling was determined through Bulked Segregant Analysis in a F2 population (turnip X Chinese turnip rape), with both dominant and additive effect, explaining 65.2% variation (Jiang 2001). In general the genetic control of many quantitative traits is unknown due to their complex inheritance patterns. In the studies mentioned above, there is no information, which allows the assignment of a chromosome number (R group), making it difficult to compare and integrate these results.

The linkages described above were identified in segregating populations involving crosses between two contrasting parental genotypes. Association mapping (AM) provides an additional opportunity to detect allele trait associations. Localization of QTL for complex traits in mapping populations is limited to those loci for which the crossed parents segregate, whereas association mapping can provide greater resolution for identifying loci controlling complex traits in a natural population based on linkage disequilibrium (LD) (Flint-Garcia et al. 2003). In crop species, marker-trait associations have been demonstrated (see review by Gupta et al. 2005). However, the use of AM in *Brassica* species has not yet been explored.

Table 2 Genetic mapping of several traits in *Brassica rapa*

Trait	Gene symbol	Population	Flanking markers or QTL	Marker types	Linkage group(LG)	Reference
Oleic acid content	fad2	F2	OPH-17	SCAR	LG6	Tanhuanpää et al. 1996b
Linolenic acid content	fad3	F2	OPS-01~OPJ-20 OPP-05~OPG-16	RAPD	LG3,LG9 LG10	Tanhuanpää and Schulman 2002
Erucic acid content	Eru	F3	tg1f8 GAP-b	RFLP	LG1	Teutonico and Osborn 1994
Clubroot resistance		DH	RA12-75A WE22B WE49B	RAPD		Yasuhisa et al. 1997
Clubroot resistance	CRa	F2	HC352b HC181	RFLP	LG3	Matsumoto et al. 1998
Clubroot resistance	CRb	F2	TCR09 TCR05	SCAR		Piao et al. 2004
Clubroot resistance	Crr1 Crr2	F2	BRMS-088 BRM-096	SSR		Suwabe et al. 2003
Clubroot resistance	Crr3	F3	OPC11-1S OPC11-2S	STS		Hirai et al. 2004
Clubroot resistance	Crr1 Crr2 Crr4	F2	BRMS297-BRMS088 BRMS100-BRMS096 BN288D-WE24-1	SSR	R08 R01 R06	Suwabe et al. 2006
Whiterust resistance	Aca1	RI	Wg6c1a, Pub1	RFLP	R04 R02	Kole et al. 2002
TuMV resistance		F2	CAG 150 CAC 150	AFLP		Han et al. 2004
Dwarfism	DWF2	F2, BC2	At2g 01810	RFLP	R06	Muangprom and Osborn 2004
Microspore embryogenic ability		F2	OPE 03-1600 OPA 13-1200 OPB 70-1400	RAPD	4LGs	Zhang et al. 2003
Seed coat color		BC3	B06-600	RAPD		Chen et al. 1997
Seed coat color	Yls	F3	M456b Ec3c8b	RFLP	LG5	Teutonico and Osborn 1994
Pubescence	Pub1	F3	Ec2b3 Ec2e12	RFLP	LG4	Teutonico and Osborn 1994
Bolting time		F2	BN007-1	RAPD		Ajisaka et al. 2001
Bolting time	bt1~bt10	DH	10QTL	AFLP	6LGs	Nishioka et al. 2005
Flowering time	VFR1-VFR3 FR1-FR3	F2, RI	6 QTL	RFLP	LG2 LG8	Teutonico and Osborn 1995; Osborn et al. 1997
Flowering time	VFR2	BC3S1	TG1G9	RFLP	R10	Kole et al. 2001
Flowering time		F2	pCOE2pAt12E1 pN121E2pN102B1		R02 R03	Axelsson et al. 2001
Flowering time	BrFLC2 BrFLC3 BrFLC1	BC3S1 BC1S1		RFLP SSR	R02 R03 R10	Schranz et al. 2002
Morphological traits		F2	48 QTL	RFLP	9 LGs	Song et al. 1995
Morphological traits		F2	24 QTL	RAPD AFLP	8 LGs	Lu et al. 2002
Morphological traits		RI	50 QTL	AFLP RAPD	14 LGs	Yu et al. 2003
Morphological traits		DH				Choi et al. 2004
Heat tolerance	ht-1~ht-5	RI	5 QTL	AFLP RAPD	LG3,LG8 LG9	Yu et al. 2003

3.3 Comparative mapping

The genome relationships of the three diploid *Brassica* species allow comparative analysis between the A, B and C genomes. The *Brassica* genus is also closely related to *Arabidopsis thaliana*, which separated around 14.5–20.4 million years ago from a common ancestor (Bowers et al. 2003). Comparative genome analysis between *A. thaliana* and *Brassica* species can be used to transfer information and resources from the widely studied model organism to this important group of crop plants.

Early comparative studies conducted at the level of genetic linkage maps revealed extensive duplication within *Brassica* genomes using a common set of RFLP probes (Lagercrantz and Lydiate 1996). Subsequent comparative analyses between *Brassica* and *Arabidopsis* genome discussed segmental duplications and extend of the genome rearrangements (Lagercrantz 1998; Lan and Paterson 2000; Lukens et al. 2003; Schranz et al. 2006). Furthermore, BAC contigs of *A. thaliana* genome were used to exploit homologous chromosomal regions in Brassicaceae species, revealing genome triplication within *Brassica* and related species (Lysak et al. 2005; Schranz et al. 2006). A linkage map of *B. napus* was constructed using RFLP probes derived from sequences of each of five Arabidopsis chromosomes, and a comparative genome analysis was conducted which identified 21 conserved Arabidopsis genomic blocks the majority of which could be aligned 3 times to both A (A1-A10) and C (N11-N19) genomes (Parkin et al. 2005). These findings strongly support that triplication in the genomes of both *B. rapa* and *B. oleracea* is involved. However, 3 segments of the *Arabidopsis thaliana* genome aligned to seven segments of the *B. napus* genome, and 5 other segments of the *Arabidopsis* genome aligned to only 4 or 5 segments of the *B. napus* genome indicating additional duplications or losses of DNA segments.

The *B. rapa* genome is about 4 to 5 times larger in size than that of *A. thaliana*. The synteny of the *B. rapa* chromosomes with the Arabidopsis chromosomes is currently exploited in several labs, as observed for genes controlling flowering time (Osborn et al. 1997; Kole et al. 2001; Schranz et al. 2002; Yang et al. 2006), the dwarf gene *DWF2* (Muangprom and Osborn 2004), and resistance to white rust genes (Kole et al. 2002). This genetic information will generate knowledge on the position of genes affecting plant morphology and growth characteristics, and will also provide markers for genetic improvement of *B. rapa* by breeding.

4 Scope of the thesis

The main objective is to investigate the genetic variation and regulation of phytate and other important agronomic traits in *B. rapa*. From an extensive screen of the available germplasm using AFLP fingerprinting, the population structure and marker-trait association in *B. rapa* will be analysed. Furthermore, parental genotypes will be selected for the development of F2 and DH mapping populations. This will allow quantitative trait locus (QTL) mapping of phytate and phosphate accumulation, and other interesting agronomic traits. This mapping will result in the identification of loci affecting these traits and will help to develop molecular markers that will facilitate breeding for these traits both in vegetables and oil seed *B. rapa*.

In **Chapter 2** we investigate the genetic variation in two sets of diverse accessions of *B. rapa* representing different morphotypes and geographic origins. The relationship among the accessions is evaluated using AFLP technology. European and Asian accessions are compared because they have a long and independent domestication and breeding history in both regions.

In **Chapter 3** association mapping is applied to investigate the genetic basis of variation within *B. rapa*. The high amount of variation at DNA and phenotype levels observed in chapter 2 prompted us to investigate the association between markers and traits. The possible population structure of a set of *B. rapa* accessions used in chapter 2 is further discussed. Furthermore, we exploit the variation for phytate and phosphate in seeds and leaves and tested if association mapping could be used to identify genomic regions controlling these traits.

In **Chapter 4** we describe QTL analysis using 5 segregating populations in order to unravel the genetics of phytate and phosphate accumulation in *B. rapa*. The localization of major QTL on the genetic linkage maps of *B. rapa* is presented. Consecutively, the difference between phytate and phosphate QTL, and also between leaf and seed QTL in the same or different populations is discussed. Furthermore, the correlations between phosphate and phytate in leaves and seeds are described.

In **Chapter 5** multiple populations derived from 5 morphologically distinct genotypes are used to identify genetic loci involved in several important morphological/agronomic traits. QTL are identified for flowering time, leaf traits, seed traits and turnip traits. The possible genetic correlation among these traits is discussed. Furthermore, we compare QTL positions in genetic maps of Arabidopsis and related *Brassica* species.

In **Chapter 6** the combined findings of chapter 2 to 5 are discussed and suggestions for further research are formulated.

Chapter 2

Genetic relationships within *Brassica rapa* as inferred from AFLP fingerprints

Jianjun Zhao, Xiaowu Wang, Bo Deng, Ping Lou, Jian Wu, Rifei Sun, Zeyong Xu, Jaap Vromans, Maarten Koornneef, Guusje Bonnema

Abstract

Amplified fragment length polymorphism (AFLP) markers were employed to assess the genetic diversity amongst two large collections of *Brassica rapa* accessions. Collection A consisted of 161 *B. rapa* accessions representing different morphotypes among the cultivated *B. rapa*, including traditional and modern cultivars, and breeding materials from geographical locations from all over the world and two *Brassica napus* accessions. Collection B consists of 96 accessions, representing mainly leafy vegetable types cultivated in China. On the basis of the AFLP data obtained, we constructed phenetic trees using MEGA 2.1 software. The level of polymorphism was very high, and it was evident that the amount of genetic variation present within the groups was often comparable to the variation between the different cultivar groups. Cluster analysis revealed groups, often with low bootstrap values, which coincided with cultivar groups. The most interesting information revealed by the phenetic trees was that different morphotypes are often more related to other morphotypes from the same region (East Asia versus Europe) than to similar morphotypes from different regions, suggesting either an independent origin and/or a long and separate domestication and breeding history in both regions.

This chapter has been published in Theoretical and Applied Genetics (2005) 110: 1301-1314

Introduction

The *Brassica* genus comprises six crop species, each with considerable morphological variation. Through interspecific hybridizations in all possible combinations, three basic diploid plant species *B. rapa* (A genome, n=10), *B. oleracea* (C genome, n=9) and *B. nigra* (B genome, n=8) gave rise to three amphidiploid species *B. napus* (AC genome, n=19), *B. juncea* (AB genome, n=18) and *B. carinata* (BC genome, n=17) (U 1935). Fingerprinting using restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) has generated information on the evolution of the amphidiploid species, the origins of the dipoid species and the relationship between different morphotypes or cultivar groups (Mizushima 1980; Prakash and Hinata 1980; Song et al. 1988a, b, 1990; Demeke et al. 1992; Jain et al. 1994; Thormann et al. 1994; Das et al. 1999; Chen et al. 2000; Guo et al. 2002; He et al. 2002, 2003). Although sufficient proof of the origin of cultivated *B. rapa* is lacking, the most likely explanation is that the wide variation within cultivated *B. rapa* arose independently at different places in the world from wild *B. rapa*. Only a few studies using small numbers of accessions and a limited number of RFLPs, RAPDs and AFLPs have been published (Vaughan 1977; Song et al. 1988b; Chen et al. 2000; Guo et al. 2002; He et al. 2003). The results of these studies suggest that cultivated subspecies of *B. rapa* most likely originated independently in two different centers-Europe and Asia. Turnip and turnip rape (*oleiferous* forms) are the dominating forms in the European center (Reiner et al. 1995; Gomez Campo 1999). In East Asia, leafy vegetables such as Chinese cabbage, Pak choi and Narinosa may have been domesticated first in China. China is also the center of origin of Chinese turnip rape (ssp. *oleifera*) (Li 1981), which is a unique turnip rape (oil type). Other accessions of *B. rapa* most likely derived from different morphotypes in the two centers of origin and subsequently evolved separately.

B. rapa is an important vegetable crop and to a minor extend also an oil seed crop. *B. rapa* vegetables are consumed worldwide and provide a large proportion of the daily food intake in many regions of the world. It is of interest that there is variation in the plant organs that are consumed, which has resulted in the selection of different morphotypes depending on local preferences. Because *B. rapa* has been cultivated for many centuries in different parts of the world, the variation within the species has increased as a result of ongoing breeding. Based primarily on the organs used and secondly on their morphological appearance, a number of

major cultivar groups, which have been given sub-species names in the past, can be distinguished (Diederichsen 2001).

The oil seed types (ssp. *oleifera*) fall into different subgroups based on their growth habit (spring and winter types). The Chinese turnip rape is possibly developed from Pak choi in south China (Li 1981; Liu 1984) and shows strong branching. The separate breeding tradition in India led to the development of the Sarson types which are very early, self-compatible and often yellow-seeded (Gomez Campo 1999).

A group of cultivars grown for their swollen stem basis are the turnip types (ssp. *rapa*), which can be subdivided in vegetable and fodder turnips. This group probably represents one of the oldest groups of cultivated *B. rapa* types (Siemonsma and Piluek 1993). Manifold shapes and colors are typical characteristics of turnips, especially vegetable turnip.

A large and diverse group of *B. rapa* cultivars are cultivated for their leaves. In these leafy vegetables several subgroups can be clearly distinguished. The Chinese cabbage group (ssp. *pekinensis*) is characterized by large leaves with a wrinkled surface, a pale-green color, large white midribs and heads of different shapes. Pak choi (ssp. *chinensis*) does not form a head, has darker green and smooth leaves with a pronounced white midrib. Wutacai (ssp. *narinosa*) forms a subgroup of Pak choi-like cultivars that differ from typical Pak choi types by their flat appearance and many dark leaves. Taicai's (or Tai tsai's) (ssp. *chinensis*) are non-heading cabbage cultivars with irregularly notched leaves of different blade shapes. The tender leaves, stems, and even the conical-shaped succulent taproots are edible. These types are mainly distributed throughout eastern China and are widely cultivated in the Shandong and Jiangsu provinces (Cao et al. 1997; Zhu and Zheng 2001).

Another group of cultivars also cultivated for their leaves are characterized by many, often narrow leaves that are either serrated or not serrated. These cultivars belong to the *perviridis* group, which includes neep greens from Europe, the Japanese cultivar group Komatsuna, and the *nipposinica* group, including Mizuna and Mibuna and leaf potherb mustards. The Shuicai cultivars from China resemble Mizuna or Mibuna, and Chinese Fennie (tilling) vegetable with strong stooling leaves also belongs to the *japonica* group (Cao et al. 1997).

Another use of *B. rapa* are the stems in red purple Zicaitai (ssp. *chinensis*) from southern China. This flowering purple-stemmed Chinese cabbage has tender early inflorescences, stems and shoots which are edible.

The inflorescences of flowering cabbages, such as the Broccoletto or Cima di rapa types found in Italy, are yet another plant organ of *B. rapa* that is consumed. In China flowering cabbages are called Caixin or Caitai. These have a growth habit similar to that of Broccoletto and probably have evolved independently. Caixin and Broccoletto have a rather different taste (http://www.plantnames.unimelb.edu.au/Sorting/Brassica.html), which also indicates their different origin.

The development of AFLP technology has been useful for analyzing genetic diversity in many plant species and has considerable potential for generating a large number of polymorphic loci (Vos et al. 1995; Mackill et al. 1996; Powell et al. 1996; Koopman et al. 2001; Srivastava et al. 2001; Huang et al. 2002; Negi et al. 2004). In the investigation reported here, we used AFLP technology to analyze the relationships among 257 *B. rapa* accessions derived from different parts of the world. Special emphasis was placed on comparing European and Asian accessions, which have a long independent breeding history.

Materials and methods

Plant materials

We used the nomenclature developed by Diederichsen (2001) to describe the different cultivar groups as subspecies. In experiment A, 163 *B. rapa* accessions, including various morphological types and two *B. napus* species were selected out of 230 accessions. The accessions were obtained from the Dutch Crop Genetic Resources Center (CGN) in Wageningen, the Chinese Academy of Agricultural Sciences (CAAS)-Institute for Vegetable and Flowers (IVF) and the Oil Crop Research Institute (OCRI) and from Dr. Osborn (University of Wisconsin, Madison, Wis., USA), who provided three parental lines of mapping populations. The collection includes traditional cultivars, breeding material and modern cultivars originating from different geographical locations. All of the accessions used in the study and their origins are listed in Table 1. In an independent experiment, experiment B, 96 *B. rapa* accessions from the CAAS-IVF were studied. This experiment included mainly leafy types from different provinces or regions in China, although a few came from outside of China. These accessions represent the various morphotypes cultivated in China and their origins are listed in Table 2.

The accessions listed in Table 1 were grown in the greenhouse and evaluated for leaf characteristics (4 weeks after sowing), flowering time, seed color and self-compatibility (see Table 4). Inflorescences were covered with plastic bags to prevent cross-pollination. Plants that set seeds on these bagged inflorescences were considered to be self-compatible.

Table 1 List of accessions used in experiment A

Genotyp[a]	Cultivar name[b]	Accession no.	Origin (country)	Genotype[a]	Cultivar name	Accession no.	Origin (country)
Chinese cabbage (ssp. *pekinensis*)				BRO-030	Sessantina	CGN06829	Italy
CC-057		CGN07182	China	BRO-127	Edible Flower	CGN17278	Japan
CC-148	Bao Tou Qing	VO2A0006	China	Turnip (ssp. *rapa*)			
CC-062		CGN07189	Germany	VT-116	Nagasaki Aka	CGN15200	Japan
CC-112	Bao Tou Qing	CGN15194	China	VT-117	Toya	CGN15201	Japan
CC-160	Qing Kou Bai Cai	VO2A0044	China	VT-115	Kairyou Hakata	CGN15199	Japan
CC-167	Luo Yang Large Bai Cai	VO2A0062	China	VT-124	Jinengu-Kabu	CGN15221	Japan
CC-147	Si Ji Qin Bao tou bai	VO2A0005	China	VT-123	Terauchi-Kabu	CGN15220	Japan
CC-142	Matsushima Jun Sang	CGN21732	Japan	VT-012	Ronde Rode Heelblad-Yurugu Red	CGN06720	Japan
CC-152	Huang Yang bai	VO2A0016	China	VT-013	Ronde Rode Heelblad-Scarlet Ball	CGN06721	Japan
CC-048		CGN06867	Soviet Union	VT-007	Maiskaja	CGN06710	Soviet Union
CC-049	Granaat	CGN07143	Netherlands	VT-009	Ronde Rode -Tsutsui	CGN06717	Japan
CC-153	Bao Tou Bai Cai	VO2A0020	China	FT-088	Blauwkop Heelblad-Oliekannetjes	CGN10985	Netherlands
CC-163	Tian jing Bai Cai	VO2A0049	China	VT-053	Teltower Kleine	CGN07167	Germany
CC-162	Luo Yang Bai	VO2A0048	China	VT-010	Platte Ronde Blauwkop Ingesneden Blad- Lila Ker	CGN06718	Hungary
CC-168	Luo Yang Da Bai Cai	VO2A0068	China	VT-044	Soloveckaja	CGN06859	Soviet Union
CC-060		CGN07185	China	VT-015	Bianca Lodigiana; Italiaanse Witte	CGN06724	Italy
CC-113	Bei jing 106	CGN15195	China	VT-017	Platte Witte Meirapen	CGN06732	Netherlands
CC-093		CGN11002	China	FT-001	Halflange Witte Blauwkop Ingesneden Blad-Barenza	CGN06669	Netherlands
CC-150	Yu Quan Bao Tou Qing	VO2A0012	China	FT-097	Buko; Bladraap	CGN11010	Germany
CC-169	Huang Yang Bai	VO2A0069	China	VT-018	Goudbal; Golden Ball	CGN06774	Netherlands
CC-158	Gao Zhuang Huang Yang Bai	VO2A0034	China	VT-008	Pusa Chandrina	CGN06711	India
CC-154	Luo Yang Da Bai Cai	VO2A0023	China	VT-120	Platte Gele Boterknol	CGN15210	Netherlands
CC-155	Huang Yang Bai	VO2A0029	China	VT-014	Platte Witte Blauwkop Heelblad-Milan	CGN06722	Italy
CC-166	Huang Yang Bai	VO2A0056	China	VT-045	Milanskaja; Italiaanse Witte	CGN06860	Italy
CC-156	Huang Yang Bai	VO2A0030	China	VT-092	Amerikaanse Witte Roodkop Heelblad	CGN11000	Netherlands
CC-071	BRA 211/69	CGN07200	Japan	VT-011	Platte Witte Blauwkop Ingesneden Blad-Siniaja	CGN06719	Soviet Union
CC-073	BRA 127/67	CGN07202	China	FT-005	Ochsenhorner	CGN06688	Germany
CC-125		CGN15222	Korea	VT-091	Snowball; Blanc Rond de Jersey	CGN10999	United Kingdom
CC-068		CGN07196	Bulgaria	VT-089	D`Auvergne Hative	CGN10995	France
CC-069		CGN07198	USA	FT-004	Lange Gele Bortfelder	CGN06678	Denmark
CC-067		CGN07195	Japan	VT-006	Pusa Chandrina	CGN06709	India
CC-114	Xiao Qing Kou	CGN15196	China	VT-137		CGN20735	Uzbekistan
CC-159	Gao Zhuang Da Bai Cai	VO2A0039	China	VT-052	Hilversumse; Marteau	CGN07166	Netherlands
CC-072	BRA 207/70	CGN07201	China	VT-090	De Croissy	CGN10996	France
CC-095		CGN11005	China	VT-119	Roodkop-Pfalzer	CGN15209	Netherlands
CC-058		CGN07183	Czech Republic	FT-047	Moskovskij	CGN06866	Soviet Union
CC-070	BRA 47/22	CGN07199	Korea	FT-002	Grote Ronde Witte Roodkop-Norfolk De Norfolk a Collet Rouge	CGN06673	United Kingdom
CC-165	Tian jing Bai Cai	VO2A0054	China	FT-003	Lange Witte Roodkop	CGN06675	Netherlands
CC-059		CGN07184	Korea	FT-051	Krasnaja	CGN07164	Soviet Union
CC-141	Kyoto Sang	CGN21731	Japan	FT-056	Daisy; Bladraap	CGN07179	France
CC-140	Kashin	CGN20771	Japan	FT-086		CGN07223	Pakistan
CC-157	Huang Yang Bai	VO2A0031	China	Neep greens (ssp. *perviridis*)			
CC-164	Tian jing Bai Cai	VO2A0053	China	KOM-041		CGN06843	Japan
CC-061		CGN07188	Yugoslavia	KOM-118	Komatsuna	CGN15202	Japan
CC-146	Long Kang er Gao Zhuang	VO2A0001	China	TG-129	Vitamin Na	CGN17280	Japan
CC-094		CGN11003	Japan	TG-131	Maruba Santo Sai	CGN17282	Japan

Table 1 (continued)

Genotyp[a]	Cultivar name[b]	Accession no.	Origin (country)	Genotype[a]	Cultivar name	Accession no.	Origin (country)
CC-161	Huang Yang Bai	VO2A0046	China	Mizuna(ssp. *nipposinica*)			
Pak choi (ssp. *chinensis*)				MIZ-019	Bladmoes	CGN06790	Netherlands
PC-099	Chinese Bai Cai	CGN13924	China	MIZ-079		CGN07213	Japan
PC-172	No 17 Bai Cai	VO2B0207	China	MIZ-128	Round Leaved Mibuna	CGN17279	Japan
PC-173	Kui Shan Li Ye Bai Cai	VO2B0223	China	Turnip rape (ssp. *oleifera*)			
PC-176	Ai Jiao Hei Ye Bai Cai	VO2B0232	China	OR-211	Yi Chang Xiao You Cai	OCRI1771	China
PC-107	Dwarf	CGN15184	Hong Kong	OR-210	Luo Tian You Bai Cai	OCRI1757	China
PC-175	HKG Nai Bai Cai	VO2B0226	China	OR-213	Huang Po Tian You Cai	OCRI0235	China
PC-189	Ai Hei Ye Kui Shan Bai Cai	VO2B0715	China	OR-216	Xi Qiu Bai Cai	OCRI3742	China
PC-187	Ai Hei Ye Kui Shan Bai Cai	VO2B0695	China	OR-214	Chang De Nanjing Zi	OCRI1789	China
PC-180	Jiang Mei Xiao Bai Cai	VO2B0612	China	OR-212	Xing Shan You Cai	OCRI1776	China
PC-186	D94 Bai Cai	VO2B0694	China	OR-218	Gao Zhi Huang You Cai	OCRI3764	China
PC-177	Ai Jiao Huang	VO2B0396	China	OR-219	Ping Ba Bai You Cai	OCRI3801	China
PC-171	B139 Xiao Bai Cai	VO2B0206	China	OR-209	Huang Gang Bai You Cai	OCRI1752	China
PC-195	Kuang Hei Fu Bing CC6	VO2B1299	China	OR-217	Cha Yuan Bai You Cai	OCRI3752	China
PC-185	Qing Ken Bai Cai	VO2B0691	China	SO-031		CGN06832	USA
PC-191	Wuhan Ai Jiao Huang	VO2B0988	China	SO-032	Pusa Kalyani	CGN06834	India
PC-193	CII	VO2B1263	China	SO-034	Australian RARS	CGN06836	Bangladesh
PC-182	Nan Jiang Bai	VO2B0620	China	SO-035	Somali Sarisa	CGN06837	Bangladesh
PC-192	Wang Yue Man	VO2B1223	China	SO-037	Kalyania	CGN06839	Bangladesh
PC-194	Qing Ken Bai Cai	VO2B1297	China	SO-038		CGN06840	Germany
PC-174	Bai Cai VS-2	VO2B0225	China	SO-039	Sampad	CGN06841	Bangladesh
PC-178	Ai Jiao Huang You Cai	VO2B0487	China	SO-040	Candle	CGN06842	Canada
PC-023	Si Yue Man	CGN06817	China	WO-024	Svalof 0308	CGN06818	Sweden
PC-188	Tai Wan Chi Ye Bai Cai	VO2B0697	China	WO-080		CGN07216	Pakistan
PC-022		CGN06816	Netherlands	WO-081		CGN07217	Pakistan
PC-076		CGN07205	China	WO-083		CGN07220	Pakistan
PC-100	Cabbage Tientsin	CGN13925	China	WO-084		CGN07221	Pakistan
PC-101	Tientsin; Celery,Shantung,Peking	CGN13926	China	WO-085		CGN07222	Pakistan
PC-183	Ai Kuang Qing	VO2B0655	China	WO-087		CGN07226	Pakistan
PC-184	Ai Jiao Bai	VO2B0656	China	WO-145	Per	KT18	USA
Caixin (ssp. *parachinensis*)				RC-144	Rapid cycling	FIL501	USA
BRO-103	Tsja Sin; No.P1R5T5	CGN15158	Indonesia	Yellow Sarson (ssp. *tricolaris*)			
PC-078	Choy Sam	CGN07211	Netherlands	YS-033	Dys 1	CGN06835	Germany
Broccoletto (ssp. *broccoletto*)				YS-143	R500	FIL500	USA
BRO-027	Quarantina	CGN06825	Italy	Wutacai (ssp. *narinosa*)			
BRO-029	Norantino	CGN06828	Italy	PC-105	BRA 77/72	CGN15171	China
BRO-026		CGN06824	Italy	*B. Napus*			
BRO-028	Tardivo	CGN06827	Italy	BN-222		OCRI0027	China
BRO-025	Natalino	CGN06823	Italy	BN-226		OCRI0046	China

[a] CC, Chinese cabbage; PC, Pak choi; BRO, Broccoletto; VT, vegetable turnip; FT, fodder turnip; KOM, Komatsuna; TG, turnip green; MIZ, Mizuna; OR, Chinese turnip rape; SO, spring turnip rape; WO, winter turnip rape; RC, rapid cycling; YS, Yellow Sarson; BN, *Brassica napus*.

[b] Bai, white; Cai, cabbage; Da, large; Huang, yellow; Hei, black; Kang, resistance; Tou, head; Xin, center; Xiao, small; Yang , seedling; Yuan, round; You Cai, oilseed rape.

Table 2 List of accessions used in experiment B

Genotype[a]	Cultivar name[b]	Accession no.	Origin[c]	Genotype[a]	Cultivar Name[b]	Accession no.	Origin[c]
Chinese cabbage (ssp. p*ekinensis*)				cPC83	Kang re 605 Qing Cai		Shang hai
cCC94	Huang Yang Bai	V02A0046	Si chuan	cPC86	Jing Guan Wang Wing Geng Cai Zhong		Shan tou
cCC98	Tai GuGu Diu	V02A1003	Tian jing	cPC88	Wu yue man		Bei jing
cCC120	Da Bang	V02A1489	Shang dong	cPC136	Shao Yang Tiao Geng Bai	V02B1236	Hu nan
cCC101	Xue Li Bai Xin Cai	V02A1096	Yun nan	cPC139	Taicai	V02B0445	Jiang su
cCC110	Ji Tui Bai	V02A0788	Nei meng	cPC142	Bai Bang You Cai	V02B0544	Tian jing
cCC102	Si Ji Bao Tou Qing	V02A0005	Bei jing	cPC143	Yu Yao Xiao You Cai	V02B1278	Zhe jiang
cCC116	Xiao Qing Kou		Shan xi	cPC145	Lv Bian Jv Hua Xin	V02B0002	An hui
cCC117	Huang Yang Bai		Yun nan	cPC154	Hou Ma You Cai	V02B0503	Shan xi
cCC119	Caul		North Korea	cPC155	Chun Taicai	V02B0097	Shan xi
cCC107	Niu Tui Bang	V02A0747	Qing hai	cPC159	Wu han ai jiao huang	V02B0481	Hu bei
cCC124	Fu Shan Bao Tou	V02A1382	Hu bei	cPC165	Duan Hei Ye Kui Shan Bai Cai	V02B0988	Hong kong
cCC121	Kao Zhuang Bai	V02A1358	Si chuan	cPC168	Piao er Cai	V02B0893	Si chuan
cCC108	He Tao Wen	V02A0574	Liao ning	cPC196	Ya Li Shan Jiao Nai Bai Cai		Guang dong
cCC128	Zhu Long Cai	V02A1499	Shan xi	cPC198	Da Tou Qing Jiang Bai Cai		Guang dong
cCC99	Xiao Shi Zi Tou	V02A0555	Jiang su	cPC208	Hai Lv You Cai		Tian jing
cCC111	Xin Hua er Bao Tou	V02A0710	Nei meng	cPC209	Chang Geng Bai Cai		Guang zhou
cCC118	Bleak Leaf 30 Days	V02A1564	Asia vegetables center	cPC211	Xia Qing		Shang hai
cCC122	Jia bai 2 hao		Hei long jiang	cRC216	Rapid cycling		USA
cCC125	Xing Cheng Xiao Cuo Cai	V02A1396	Ji lin	Cai xin (ssp. p*arachinensis*)			
cCC100	He Ze Da Bao Tou		Shan dong	C54	Xiang Gang Caixin		Shan tou
cCC103	Zhang Zhou Zhang Pu Lei	V02A0133	Fu jian	C58	Si Jiu Cai Xin		Bei jing
cCC105	Xiao Qing Kou	V02A0727	Ning xia	C60	Er Yue Caitai		Shang hai
cCC114	Gan Zhou Bai Cai	V02A1212	Jiang xi	C91	Si Ji Duan Ye You Qing Caitai		Guang dong
cCC112	Da Tai Zhong Qing Ma Ye	V02A0704	Nei meng	C190	Chang Sha Chi Hong Cai		Hu nan
cCC130	Xiao Gen	V02A0582	Lliao ning	C194	Chihua Cu Jing Te Qing Caitai		Guang dong
cCC123	77-KR		Bengal	C195	45 Days Caixin		Guang dong
cCC127	Ji Nan Da Ken		Shang dong	C210	48-19 Caixin		Guang dong
cCC133	Xin Jiang Da Bao Tou	V02A1022	Xin jiang	Zi Caitai (ssp. *chinensis* var.*purpurea* Bailey)			
cCC93	Cao Zhou Gao Zhuang	V02A0359	He nan	C62	Zicaitai		Bei jing
cCC134	Shou Guang Xiao Gen		Shan dong	Turnip (ssp. *rapa*)			
cCC104	Da Mao Bian	V02A0961	Shan xi	T172	Ka Ma Gu	V01C0082	Xin jiang
cCC109	Cheng Du Bai		Si chuan	T173	Bai Yuan Ken	V01C0054	Si chuan
cCC131	Cheng De Fan Xin Bai	V02A0200	He bei	T174	Yuan Man Qing	V01C0008	He bei
cCC115	Shi Zi Tou Da Bai Cai	V02A0002	An hui	T175	Man Qing	V01C0030	Shan dong
cCC132	Xiao Qing Kou		Gui zhou	T176	Hua Ye Hong Pi	V01C0036	Shan xi
cCC106	Ci Xi Huang Ya Cai		Zhe jiang	T178	Yuan Xing Wu Jing	V01C0001	An hui
cCC129	Yao Huang Zhong Huang Ya Cai		Zhe jiang	T179	Da Ying Pan Cai	V01C0044	Zhe jiang
cCC95	Wu Ping Zhai Ye Da Bai Cai	V02A0129	Fu jian	T180	Ke Bu er Man Qing	V01C0020	Nei meng
cCC96	Fen Kou Bai	V02A0172	Gui zhou	T181	Ji Xian Xian Sui Man Qing	V01C0067	He nan
cCC64	Zao Shu Wu Hao		Hang zhou	Wutacai (ssp. *Narinosa)*			
cCC82	Chun Xia Wang Bai Cai		North Korea	W56	Zhong Ba Ye Wutacai		Bei jing
cCCB70	Wan Quan		Tai wan	W87	Wutacai		Shang hai
Pak choi (ssp. *chinensis*)				W204	Hei You Bai Cai		Hu bei
cPC61	Jing Lv 7 Hao		Bei jing	W205	Hei Ta Cai		
cPC66	Si Yue Man		Nan jing	Mizuna (ssp. *Nipposinica*)			
cPC67	Bi Yu		Nan jing	S57	Bai Geng Qian Jin Jing Shui Cai		Bei jing
cPC71	Shang Hai Qing		Shang hai	S84	Dong Fang Ren Sheng Cai		Bei jing
cPC72	Su Zhou Qing		Su zhou	S203	Qian Jing Shui Jin Cai		Hu bei
cPC75	Shang Hai Qing		Bei jing	Taicai (ssp. *Chinensis var.tai-tsai Lin)*			
cPC78	Jing Guan		Bei jing	TC182	Yuan Ye Taicai	V02C0008	Shan dong
cPC80	Gao Hua Qing Geng Bai Cai		Hong kong	TC183	Hua Ye Taicai	V02C0012	Shan dong

[a] cCC, Chinese cabbage; cPC, Pak choi; C, Caixin or Caitai; T, turnip; W, Wutacai; S, Shui cai (Mizuna); TC, Taicai; cRC, Rapid cycling; [b] Bai, white; Cai, cabbage; Da, large; Huang, yellow; Hei, black; Kang, resistance; Tou, head; Xin, center; Xiao, small; Yang, seedling; Yuan, round; You Cai, oilseed rape; [c] Origin refers either to country, or to province within China

DNA isolation and AFLP analysis

In experiment A, total DNA was extracted from lyophilized young leaves or flower buds as described by Van der Beek et al. (1992). Lyophilized plant material was ground by shaking tubes containing plant material and iron bullets in a Retsch shaker.

The AFLP procedure was performed as described by Vos et al. (1995), with minor modifications according to Bai et al. (2003). The restriction enzymes, adapters and primers used are listed in Table 3. Total genomic DNA (250 ng) was digested using two restriction enzymes, *Pst* I and *Mse* I and ligated to adaptors. Pre-amplifications were performed in a 20 µl volumes of 1 x PCR buffer, 0.2 mM dNTPs, 30 ng P_{00} and M_{00} + C, 0.4 U *Taq* polymerase and 5 µl of a 10x diluted restriction ligation mix, using 24 cycles of 94°C for 30 s, 56°C for 30 s and 72°C for 60 s. Five-microliter aliquots of the diluted (1:20) pre-amplification product were used as template for the selective amplification with four primer combinations (P14M51, P21M47, P13M48 and P23M50). Only *Pst*I primers were labeled with IRD-700 or IRD-800 at 5´ end for the selective amplification. The selective amplification was carried out using the following cycling parameters: 12 cycles of 30 s at 94°C, 30 s at 65°C-56°C (with a 0.7°C-decrease each cycle), and 60 s at 72°C, followed by 24 cycles of 30 s at 94°C, 30 s at 56°C, and 60 s at 72°C.

Following the selective amplification, the reaction products were mixed with an equal volume of formamide-loading buffer (98% formamide, 10 mM EDTA pH 8.0 and 0.1% Bromo Phenol Blue). The samples were denatured for 5 minutes at 94°C, cooled on ice and run on a 5.5% denaturing polyacrylamide gel with a LI-COR (Lincoln, Neb.) 4200 DNA sequencer (Myburg and Remington 2000).

In experiment B, *Eco*RI/*Mse*I were selected as the restriction enzymes, and the primer and adapter sequences are listed in Table 3. The AFLP procedure is as described for experiment A with minor modifications. The selective amplification was carried out using 12 primer combinations (E33M61, E36M47, E38M48, E32M60, E42M50, E37M60, E37M59, E32M49, E41M49, E38M62, E39M51 and E33M48).

Data analysis

In experiment A, the AFLP gel images were analyzed with the software package AFLP-Quantar® Pro. All AFLP bands were treated as dominant markers and scored as either present (1) or absent (0). Clearly distinguishable bands ranging from 50 bp to 500 bp were used in the

data matrix and genetic analysis. Phenetic trees were constructed using MEGA 2.1 software (Kumar et al. 2001). Similarity was calculated as the proportion of AFLP markers at which the two accessions compared had the same score ($SM_{xy} = (n_{11} + n_{00})/n$; where n is the number of markers scored). The distance is 1-SM. Cluster analysis was performed using the unweighted pair group method with arithmetic averages (UPGMA). Bootstrap values were calculated in 1,000 permutations and presented in percentages.

In experiment B, the AFLP gel images were scored by eye. Clearly distinguishable polymorphic bands ranging from of 50 bp to 50 bp were scored as present (1) or absent (0). All weak and poor bands were not recorded. The data were analyzed as in experiment A.

Table 3 AFLP primers used in AFLP analyses

Primers	Sequences
M00	5′-GATGAGTCCTGAGTAA-3′
M02	M00 + C
M47	M00 + CAA
M48	M00 + CAC
M50	M00 + CAT
M51	M00 + CCA
M61	M00 + CTG
M60	M00 + CTC
M59	M00 + CTA
M49	M00 + CAG
M46	M00 + CTT
P00	5′-GACTGCGTACATGCAG-3′
P13	P00 + AG
P14	P00 + AT
P21	P00 + GG
P23	P00 + TA
E00	5′-GACTGCGTACCAATTC-3′
E33	E00 + AAG
E41	E00 +AGG
E37	E00 +ACG
E39	E00+ AGA
E42	E00 + AGT
E36	E00 + ACC
E38	E00 + ACT
E32	E00 + AAC

Results

Genetic variation

In experiment A, a set of 15 accessions representing different morphotypes was screened with 16 *Eco*RI/*Mse*I and 16 *Pst/Mse*I primer combinations. Four pairs of *Pst/Mse*I primers that gave clear banding patterns with sufficient polymorphism were used to fingerprint 161 *B. rapa* and two *B. napus* accessions. The AFLP patterns between *B. rapa* accessions were very polymorphic. In total, 524 scorable amplification products ranging from 50 bp to 500 bp were generated, 476 of which were polymorphic, with an average of 119 polymorphic bands per primer combination. The level of polymorphism was more than 90%. Two *B. napus* accessions (representing an outgroup) and the *B. rapa* lines MIZ079 and PC105 displayed several mono-morphic bands that contributed considerably to the polymorphism rate. If these mono-morphic bands were excluded from the analysis, the degree of polymorphism was still more than 80%.

A typical AFLP image is illustrated in Fig. 1a and shows that the Broccoletto group is clearly distinguishable by a specific set of AFLP bands. The polymorphism rates were calculated for the different cultivar groups as listed in Table 1. For the larger groups, these rates were very similar: Chinese cabbage, 77%; Pak choi, 75%; winter and spring turnip rape, 77%; turnips, 82%. Two Yellow Sarson and two Mizuna accessions had remarkably similar AFLP profiles.

For experiment B, a set of 96 lines representing different morphotypes and geographical origin was screened with some *Eco*RI/*Mse*I primer combinations (48 samples are depicted in Fig. 1b). Based on the screens of experiment A and experiment B, 12 pairs of *Eco*RI/*Mse*I primers that gave clear banding patterns with sufficient polymorphisms were used to fingerprint the 96 *B. rapa* accessions. In total, 332 scorable amplification products were generated, 137 of which were polymorphic, with an average of 11.5 polymorphic bands per primer combination. The level of polymorphism was 41%. The polymorphism rate for the two large groups of Chinese cabbage and Pak choi was 48% and 52%, respectively. In experiment A, the polymorphism rate was more than 70% if only Pak choi and Chinese cabbages were taken into account.

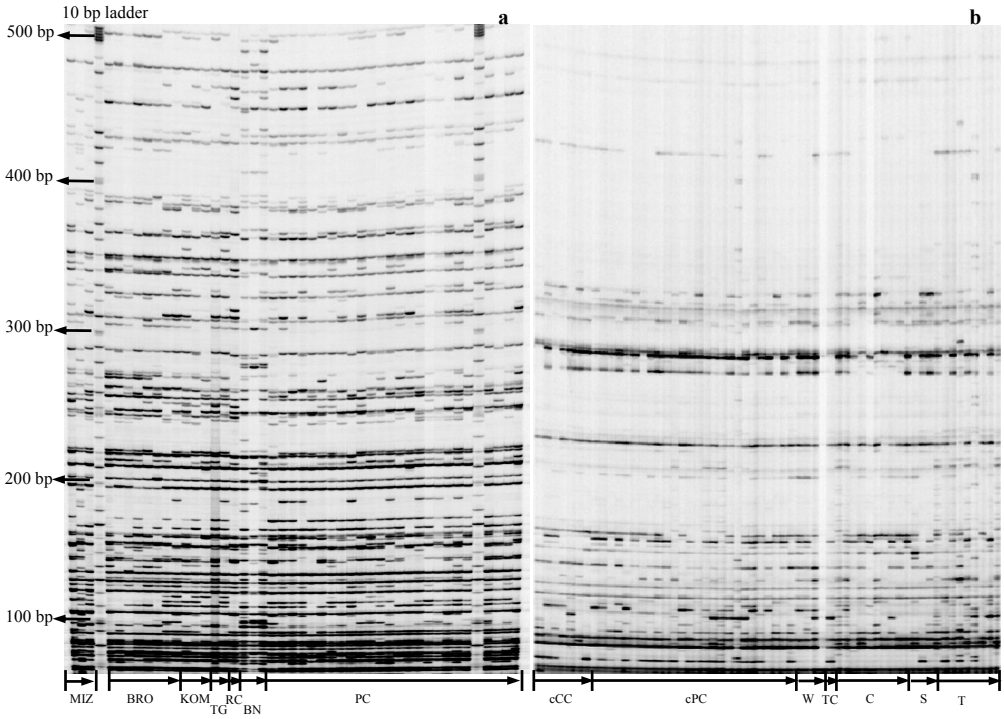

Fig. 1 An AFLP image of some *Brassica rapa* accessions using primer combination *Pst*I AG-*Mse*I CAC in experiment A and *EcoR*I AAC-*Mse*I CAG in experiment B.
See Table 1 for definition of abbreviations.

Phenetic relationships

A dendogram was established using the AFLP fingerprints (see Fig. 2). It was evident that the amount of genetic variation present within the groups was often comparable to the variation between the different sub-groups. Most accessions fell into a number of subgroups that had non-significant bootstrap values as groups, but these subgroups did represent the different morphotypes and were arranged into two main sets according to the origin of the accessions.

In experiment A, the tree formed two main groups. Group 1 consists of accessions of Asian origin, and can be subdivided in a group of Chinese cabbage cultivars (CC), one group consisting solely of Pak choi (PC1) and another group with both Pak choi and Chinese turnip rape (PC2). It also includes a small mixed group with accessions from mainly China and Japan (with two exceptions from Europe), a turnip group (T1) with accessions from Japan and a winter oilseed group (Oil1) group with accessions from Pakistan. The second group encompasses a turnip group (T2) with accessions from mainly European countries (two from

India and one from USA) and the Broccoletto group (Bro) with accessions from Italy. Furthermore, two distinct subgroups are formed by two Mizuna cultivars (Miz) from Japan and a spring oilseed and Yellow Sarson group (Oil2) with accessions from Bangladesh, USA and Germany.

The Chinese cabbage group (CC) consists of two clusters comprising solely Chinese cabbage and a less well-defined group consisting of Chinese cabbage accessions, one Pak choi type (PC-101) and one fodder turnip accession FT056. The Pak choi (PC) group is close to the CC group and is divided into PC1 and PC2. Most of Pak choi accessions are clustered in PC1 together with two Caixin accessions (BRO-103, PC-78). BRO-103 is not a Broccolleto-type cultivar and should be renamed to Caixin. PC2 is a mixed cluster, containing Pak choi and Chinese turnip rape (OR) accessions. A small group of different morphotypes of oriental origin (mainly Japan and China) can be found between the PC1 and T1 groups, assuming that PC-22 from the Netherlands also has an oriental origin. This latter group showed no clear structure.

The two turnip subgroups (T1 and T2) containing both vegetable and fodder turnip and the oil types originated from different geographical regions. T1 group accessions are all from Japan (except for VT-007 from Russia), while T2 accessions are from the western hemisphere, namely Europe, former Soviet Union and USA (except for two accessions from India and one from Japan). In group 2, all Broccoletto accessions (Bro group) are clearly distinguishable as a separate subgroup with a high bootstrap value of 86.

In addition to the groups described above a number of less related and small outgroups could be identified.

One group consists of two Mizuna types (ssp. *nipposinica*). Another group in which high bootstrap values indicated a clear distinction is the Oil2 group, with the early yellow-seeded oil types from India and the rapid-cycling lines developed by Dr. Paul Williams (Williams and Hill 1986), probably with Yellow Sarson types in their pedigree.

Two accessions, namely a Wutacai type (PC-105) and a Mizuna type (MIZ-079), have a separate position based on a relatively large number of unique AFLP bands. Additionally one Wutacai accession was collected and AFLP analysis was performed with three pairs of the four primer combinations; the results indicated that this accession clusters between CC2 and PC1. The two *B. napus* lines were completely different from *B. rapa* species and form an outgroup with very high bootstrap value (99).

Fig. 2 UPGMA phenogram (experiments A and B) of *B. rapa* accessions based on the AFLP data obtained. Numbers on branches are bootstrap values (values smaller than 30 are not indicated). Abbreviations of the different morphotypes are as given in Tables 1 and 2. The five common accessions, CC147/cCC102, CC161/cCC94, PC189/cPC165, PC191/cPC159 and RC144/cRC216, between experiments A and B are indicated by various symbols.

In the analysis of experiment B with IVF accessions, a similar pattern appeared. Different subgroups were formed, with again low bootstrap values. It was obvious that less commonly grown but morphological distinct types form no distinct subgroup, but are dispersed within the main subgroups, although the Chinese cabbage groups are rather pure.

The Chinese cabbage cultivars (heading cabbage) are again subdivided into two groups CCa, CCb. The CCb group also includes a separate cluster with one Caitai accession C190, three turnip types (T174, T172 and T176), one Taicai TC183 and one Pak choi, cPC136. The Pak choi types are subdivided into two groups (PCa and PCb). Most of Pak choi, Shuicai and Caixin accessions are clustered in PCa together with one Taicai accession TC182 and one turnip accession, T173. PCb is also a mixed cluster, containing Pak choi, Wutacai and turnip accessions. One accession (cPC168) is close to T178 and actually is a turnip according to its phenotype; it should be renamed. Zicaitai C62 is not grouped into Caitai, but close to Wutacai in PCb.

There are five common accessions (CC-147/cCC102, CC-161/cCC94, PC-189/cPC165, PC-191/cPC159 and RC-144/cRC216) between experiment A and B. The two Pak choi accessions (PC-189/cPC165, PC-191/cPC159) group similarly in both experiments; in experiment A and B they are organized in two different PC clusters. The rapid-cycling line RC-144 (cRC216) forms a distinct group in experiment A, and also in experiment B it is very distinct between CCa and PCb. Common Chinese cabbage accessions group differently in both experiments. In experiment B, CC-147/cCC102 is in CCb close to CC-161/cCC94. In experiment A, CC-147/cCC102 groups in CC, but CC-161/cCC94 groups in a separate branch of a diverse little cluster between PC2 and T1.

Phenotypic variation

The *B. rapa* genus is morphologically very diverse. As illustrated above, phenetic groups follow morphological groups with respect to classification. In Table 4, ten phenotypic traits are listed for the different subgroups (CC, PC1, PC2, T1, T2, Oil1, Oil2, Bro and Miz).

Most of the variation for leaf color was found in the CC and PC groups. Chinese cabbages in CC are mostly yellow-green or light-green, while Pak choi types in PC1 and PC2 have darker green leaves [the light-green accessions in PC2 represent most of the oilseed rapes (OR)]. The very dark green cultivars are the Wutacai types and two Pak choi accessions, PC-107 and PC-

175 found in PC1. Whitish petioles are characteristic for the CC and PC groups. A few accessions in these groups have light-green or green petioles.

Table 4 Phenotypic characteristics for all accessions of the different morphological groups from experiment A

Clusters		CC	PC1	PC2	T1	T2	Oil1	Oil2	Bro	Miz
Leaf surface	Smooth	1	15	11	0	0	0	0	0	2
	Wrinkled; intermediate	47	0	10	8	30	7	8	6	0
Leaf edge	Entire	1	9	15	1	0	0	0	0	0
	Slightly serrated	46	6	3	5	1	0	0	0	1
	Serrated	1	0	3	2	29	7	8	6	1
Leaf color	Yellow-green	12	0	1	0	0	0	0	0	1
	Light green	32	5	16	4	21	7	8	4	1
	Green	4	8	3	4	8	0	0	2	0
	Dark green	0	2	1	0	1	0	0	0	0
Leaf shape	Round; oval	48	15	20	7	3	0	0	6	0
	Elongate	0	0	1	1	27	7	8	0	2
Leaf firmness	Strong	0	15	21	8	24	7	6	4	0
	Intermediate; weak	48	0	0	0	6	0	2	2	2
Petiole color	White	40	13	9	0	0	0	0	0	0
	Light green; green	8	2	12	6	30	7	8	6	2
	Red	0	0	0	2	0	0	0	0	0
Trichomes	No	17	15	17	5	6	0	6	5	2
	Few	22	0	2	1	5	1	1	1	0
	Many	9	0	2	2	19	6	1	0	0
Flowering time[a]	Early	0	10	0	0	0	0	8	0	0
	Middle	0	3	2	0	0	7	0	6	0
	Late	48	2	19	8	30	0	0	0	2
Self-compatibility	Compatible	18	14	1	1	nt	0	6	0	0
	Compatible	30	1	20	7	nt	7	2	6	2
Seed color	Yellow	0	0	0	0	0	0	5	0	0
	Black	5	11	11	1	1	5	0	0	0
	Brown or pale brown	43	4	10	7	29	2	3	6	2

[a] Early flowering time, fewer than 60 days after sowing; middle flowering time, fewer than 90 days after sowing; late flowering time, more then 90 days after sowing; nt, not tested.

Smooth leaves are exclusively found in the two PC groups and the MIZ group, while leaves of accessions of all other groups are wrinkled. Turnips, oil types and Mizunas all have characteristically elongated leaves.

Leaf serration is a character that is associated strongly with the UPGMA grouping in the tree. No serrated leaves or mildly serrated leaves are typical of the CC, PC and the Japanese Turnip 1 group. All oil types, the European Turnip group 2 (except VT-014) and the Brocoletto's have dissected leaves. Two Mizuna lines, MIZ-128 and MIZ-079 have distinct dissected leaves, while a different Mizuna type (MIZ-019) has slightly serrated leaves. In experiment B, Chinese cabbage, most of Pak choi's (except cPC193, cPC154) and the Caitai and Wutacai accessions have no or mildly serrated leaves, while other Chinese types (Chinese turnips, Taicai) have dissected leaves.

The presence of trichomes (leaf hairs) is variable within all groups except in Oil1, where all genotypes have trichomes, and in the PC1 and Miz group, where hairs are absent. In PC2, the four accessions with trichomes are the Chinese oil types. All Pak choi, and the Caixin (Bro-103, PC-078), Wutacai (PC-105) and Mizuna accessions have no trichomes.

Yellow seeds are typical for the Yellow Sarson genotypes in the Oil2 group. Black seeds dominate in the PC groups, since especially all Chinese oil types within PC2 have dark-colored seeds.

Flowering time varies greatly among the accessions. Very early-flowering types include the Oil2 types, the Bro group and a number of PC types, including the Caixin cultivars. Late-flowering types are the Chinese turnip rape accessions in PC2 and the Oil1 types. Very late flowering types include all of the Turnip 2 accessions, which also cannot be vernalized at the seedling stage, a treatment that does accelerate flowering in the middle-to-late accessions.

The degree of self-incompatibility showed an interesting distribution. All of the PC1 and Oil2 types are self-compatible, while most PC2, T1, Oil1, Bro, Mizuna and Komatsuna genotypes are self-incompatible. Because of their late flowering, the T2 types could not be classified for this trait. Incompatibility clearly differentiates subgroups that were found within cultivar groups with similar use or phenotype, and it separates PC1 from PC2 and Oil1 from Oil2. For experiment B, self-compatibility was not recorded.

Within the *B. rapa* species, the Broccoletto, Caixin and oil types have elongated stems or branches. Broccoletto originated from Italy and has a strong stem and short internode length

(data are not shown). The edible parts of this type are the small flower heads that appear when the plants are about 20 cm tall. The edible part is quite similar to that of Chinese Caixin, also called Flowering Chinese cabbage, which is also utilized in the early flower stage. Prior to flower opening, the leafy features of Caixin are similar to those of Pak choi.

Turnips also have their specific group characteristics (data not shown), which consist of a swollen hypocotyl and a taproot. Turnips vary widely in shape and color, but as these characteristics are not associated with specific AFLP patterns they could not differentiate between groups.

Discussion

The most interesting information revealed by the dendrogram assembled in this investigation (Fig. 2) is that different morphotypes are often more related to other morphotypes from the same region (East Asia versus Europe) than similar morphotypes from different regions, suggesting either an independent origin in both regions and/or a long and separate domestication and breeding history in both regions. The low bootstrap values for many of the groups show that most polymorphisms do not contribute to the phenotypic variation, which indicates that only a few genes are involved in causing the extreme morphologies. This may also explain why the different morphotypes could emerge independently in the different geographic regions.

Chinese turnip rapes (Chinese oil types) cluster in the PC2 group, and flowering cabbages cluster with the early-flowering PC1 group. While the clustering of Caixin with Pak choi indicates a close relationship, it is impossible to determine which type from which. Despite the fact that selection resulted in the use of the same organs, the two flowering cabbage groups viz. the Chinese Caixin types and the Italian Broccolleto types are not at all related to each other. The Caixin types are related to the Pak choi cabbages and form a separate branch with PC1 or PCa in both experiments, whereas as the Broccolleto cultivars form a clearly separate group somewhat related to European turnip (T2) and oil types. Similarly, the Chinese oil types (Chinese turnip rape) are related to Pak choi and form a subgroup within the PC2 cluster, but do not cluster with the oil types or turnips from different geographical origins.

Wutacai is also called flat Chinese cabbage because of its remarkable flat shape. This Chinese vegetable resembles Pak choi at the seedling stage and its leaves are similar structure and color as some Pak choi types, however the rosette has many more very small dark-green

leaves, and the plants bolt very late. One Wutacai (PC-105) accession in experiment A does not group with any the other accessions, and it clearly deviates from the Wutacai's of experiment B which are related to Pak choi types (PCb group). The reason why PC-105 separated from PC group cannot be explained clearly, although its distinctiveness might suggest that Wutacai types have originated from several types independently due to a re-occurrence of a major mutation. Based on RFLP studies (Song et al. 1988b), one Narinosa (Wutacai) accession also seemed to fit neither group.

Turnip types that originate mainly from Japan form a variable intermediate group, which also includes some turnip greens (BRO-127 from Japan resembles turnip greens more than Broccoletto) (Fig. 2a). This group of oriental turnips is clearly different from the European fodder and vegetable turnips, and it also flowers earlier. The Chinese cabbage accession CC-94, originating from Japan, does not fit in CC, but is positioned close to Japanese vegetable turnip types. This geographical distinction of the turnips can also be seen in morphological and physiological characters such as leaf shape and flowering time and might either be due to a long separation of breeding of the different turnip type or even an independent origin. Chinese turnips are located mainly in the PCb group in experiment B, and it will be interesting to see whether they are closely related to the Chinese oil types in the PC2 group.

The turnip greens characterized by many narrow leaves, which in our collection are mainly of Japanese origin, form a very diverse group that either clusters with the Japanese turnips or forms two very separate clusters. MIZ-079 in particular deviates greatly from all other *B. rapa* accessions and is characterized by many unique AFLP bands. MIZ-079 is similar to the other Mizuna types at the early seedling stage in having a large number of soft and serrated feathery leaves. However, the internodes of MIZ-079 elongated quickly up to a height of about 90 cm during later development, and this line is completely self-compatible, a condition which separates it from the typical Mizuna accessions. In experiment B, Shuicai accessions that resemble Mizuna form no clearly separate cluster and group in the Pak choi cluster. This suggests that similar phenotypes were selected in both China and Japan.

When the results from experiments A and B are compared, it is remarkable that the grouping is quite similar; namely, there are two main groups each of Chinese cabbage and Pak choi, with a corresponding position for the two common Pak choi accessions and the rapid-cycling accession. Unlike the common accession CC-147/cCC102, the common accession CC-

161/cCC94 has no corresponding position in both trees. In order to better compare the trees from both experiments, we analyzed the data of experiment A after removing all the types that are not represented in experiment B (Oil1, Oil2, T1, T2, Bro). This subsequent comparison between the two trees illustrated that in experiment B the two Chinese cabbage groups are much more distinct than in experiment A, while relationship between Pak choi types is similar in both trees. It is important to mention that in experiment A, 4 *Pst/Mse*I primer combinations were used, while in experiment B 12 *EcoR*I*/Mse*I primer combinations were used. *Pst*I does not cut methylated DNA and thus avoids repetitive DNA sequences, like the DNA located around centromeres. We do not know whether *Pst*I and *EcoR*I target different DNA regions, which would result in different polymorphism rates and consequently contribute to the higher polymorphism rate in experiment A.

In addition, distinguishable subgroups are formed by self-incompatible, dark-seeded winter oil seed types from Pakistan (Oil1) and early-flowering yellow-seeded self-compatible Sarson types from India and Bangladesh (Oil2), both of which are not directly related to either East Asian or European types. A previous taxonomic study of oil type *B. rapa* (ssp. *oleifera*) using RAPD and AFLP fingerprints also divided the accessions into groups corresponding to seed color and self-compatibility (Das et al. 1999). The origin of the accessions was not provided, so that we cannot directly compare the studies.

The phenetic groups based we found in this investigation based on AFLP data are consistent with previous proposed groups based on morphology, origin, isozymes and nuclear RFLPs (Vaughan 1977; Prakash and Hinata 1980; Song et al. 1988b). Previous studies have suggested that these two groups represent two centers of origin for *B. rapa,* each originating from distinct wild *B. rapa* populations (Song et al. 1988b). Since the two large groups in experiment A have similar genetic distances, it can be concluded that the genetic variation in both centers is of the same order of magnitude and that this might be the consequence of the number of independent domestication events, intercrossing and breeding history.

P.S.: Later flowcytometric analysis indicated that Mizuna MIZ-079 had an almost double DNA content compared to *B. rapa*, suggesting that it is not *B. rapa* but is allopolyploid *Brassica*.

Acknowledgements

We thank P. van der Berg for technical support and Plant Sciences Experimental Centre of Wageningen University for taking care of the plants. We thank Noortje Bas from the Centre

for Genetic Resources the Netherlands (CGN) Wageningen-UR, Li Xixiang from the Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences and Liu Shengyi, Oil Crop Research Institute, Chinese Academy of Agricultural Sciences, for kindly supplying the accessions used in this study. This project is sponsored by the Asian Facility (project AF01/CH/8 “Sino-Dutch Genomic Lab and Vegetable Research Center”).

Chapter 3

Genetic variation for phytate, phosphate and several agronomic traits in *Brassica rapa*: an association mapping approach

Jianjun Zhao, Maria-Joao Paulo, Diaan Jamar, Ping Lou, Fred van Eeuwijk, Guusje Bonnema, Maarten Koornneef, Dick Vreugdenhil

Abstract

Association mapping was used to investigate the genetic basis of variation within *Brassica rapa,* which is an important vegetable and oil crop. We analyzed the variation for phytate and phosphate levels in seeds and leaves, and additional developmental and morphological traits in a set of diverse *B. rapa* accessions and tested association of these traits with AFLP markers. Associations among markers related either to genetic linkage or to population structure. The analysis of population structure found four subgroups in the population with different trait values, suggesting an association between population structure and trait values even for traits such as phytate and phosphate levels. Therefore, population structure was taken into account in the association mapping. In total, 54 markers were found to be significantly associated with various traits, 16 of which had known map positions and some of them were confirmed in other QTL mapping studies.

This chapter has been submitted for publication

Introduction

The traditional method to investigate the genetic basis of variation within the germplasm of a species is genetic mapping based on segregating populations. Such populations are derived from crosses of two parental lines differing for the trait(s) of interest. Major limitations of this procedure are: i) only those loci for which parents differ will be detected; ii) the power to identify loci depends on genetic and statistical parameters such as size of population and number of segregating loci; iii) the development of populations takes time (Flint-Garcia et al. 2003; Gupta et al. 2005). An alternative strategy is association mapping (Flint-Garcia et al. 2003), which uses linkage of molecular markers with traits of interest in natural available populations and makes use of linkage disequilibrium (LD), which is the nonrandom association of alleles at different loci. In natural populations that have no known pedigrees, more variation can be observed than in segregating populations. Association can be detected in unrelated genotypes because linkage has not completely been broken by recombination events. LD analysis has been used widely in humans (Kruglyak 1999) and animals (Farnir et al. 2000; Nsengimana et al. 2004), and identified markers closely linked to genes affecting complex diseases (Lander and Schork 1994; Jorde 2000).

Association mapping has also been extended to plants, both at the level of individual genes and at the whole genome level (Gupta et al. 2005). Understanding the degree of LD across the genome in sampled populations will facilitate the choice of appropriate methods and germplasm collections for genetic association mapping (Varshney et al. 2005). Studies in Arabidopsis (Nordborg et al. 2002, 2005), maize (Remington et al. 2001; Tenaillon et al. 2001; Palaisa et al. 2003), rice (Garris et al. 2003) and barley (Kraakman et al. 2004) showed the impact of biological and historical factors on the extent of LD explaining the variable degrees of LD. The development of high throughput genotyping techniques, and new statistical methodologies allow a more efficient use of this genetic approach, resulting in a growing number of publications describing research on marker-trait associations in germplasm or cultivar collections, also for plants (Kraakman et al. 2006; Breseghello and Sorrells 2006; Aranzana et al. 2005; Skøt et al. 2005).

The genus *Brassica* has a long history of world wide cultivation and comprises a large and diverse group of important vegetable, oil, fodder and condiment crops. *Brassica rapa* morphotypes, including leafy vegetables, turnips and oil types, differ based on which organs

are consumed as food. Therefore, morphological characteristics like rosette morphology, leaf shape and structure, enlarged taproot, branching habit and size of the seedpods differ probably because of directed selection for specific variants. Flowering time and leaf number varies also greatly within *B. rapa*, which is possibly important for the selection of plants to meet growth environments and consumer needs. However, other traits have not or much less rigorously been subjected to human selection and might show variation independent from the use of the crop. An example of this might be nutrient composition, which is important for future plant breeding programs, provided sufficient variation is present. *Brassica* species and varieties provide a useful resource of protein, vitamin C and secondary metabolites, like glucosinolates, phosphate and other minerals for human and animals. However, the unwanted side effect of anti-nutritional substances such as phytate in *B. napus* feed for animals was also described (Peng et al. 2001), which is an example of nutrient compounds for which reduction is desired. In fact, there is considerable intra-specific variation in phytate concentration in plant edible portions (White and Broadley 2005). A three-fold difference in phytate levels between cultivars was observed in *B. napus* (Mollers et al. 1999). The screening of a number of *Arabidopsis thaliana* accessions revealed a wide range of variation in phytate levels, varying from 7.0 mg to 23.1 mg of phytate per gram of seed (Bentsink et al. 2003).

In this study we analysed the variation of phytate and phosphate in a diverse set of *B. rapa* accessions and tested if association mapping could be used to identify genomic regions controlling these traits. Additionally, we compared the outcome of the association mapping approach for phytate and phosphate content with that for the traits flowering time, leaf edge shape, the presence of leaf hair and leaf number for which it is assumed that selection depending on the use of the crop has taken place. The software STRUCTURE was used to identify different groups in the population. One strategy to deal with substructure is to firstly identify relevant genetic groups on the basis of neutral markers, for example, by a strategy as embodied in the Bayesian clustering method implemented in STRUCTURE, and secondly use that grouping data in a subsequent mixed model to model genetic correlation between individuals belonging to the same group.

Materials and Methods

Plant Materials

A collection of 160 *Brassica rapa* accessions encompassing a wide range of morphological types and geographical origins was used in this study. The accessions were obtained from the Dutch Crop Genetic Resources Center (CGN) in Wageningen, the Chinese Academy of Agricultural Sciences (CAAS)-Institute for Vegetable and Flowers (IVF), -the Oil Crop Research Institute (OCRI) and from Dr. Osborn (University of Wisconsin, Madison, Wis., USA), who provided three parental lines of mapping populations. The collection includes traditional cultivars, breeding material and modern cultivars originating from different geographical locations. All of the accessions used in the study and their origins are described in Zhao et al. (2005) and chapter 2 of this thesis.

Phenotyping

Three plants per accession were vernalized after germination for two weeks in a cold dark room (4-6°C) after germination and thereafter seedlings were grown in the greenhouse with supplementary light (16h day length) from December 2002 to March 2003 in Wageningen. The number of days to flowering of vernalized plants (VDF) was recorded, from sowing to the appearance of the first open flower. One plant from each accession was selected and its young leaves or flower buds were collected for DNA isolation and AFLP genotyping as described in a previous study (Zhao et al. 2005), and 3 batches of mature seeds from one plant of each accession were used for phytate and phosphate analysis (SPHY and SPHO, respectively). Another set of four non-vernalized plants per accession was grown under similar soil and light conditions in the greenhouse from November 2002 to February 2003 in Wageningen, and was used to score the number of days to flowering (NDF). For some very late turnip and Chinese cabbage types, which did not flower within the experimental period, NDF was set to 120 days. One whole leaf was collected from each of four five-weeks-old plants of each accession to measure phytate and phosphate contents (LPHY and LPHO, respectively). Leaves of two plants were ground together to represent one biological replication. The third set of three non-vernalized plants per accession was grown under similar soil and light conditions in the green house from March to May 2003 in Wageningen, and was used to measure leaf number (LN) at seedling stage four-weeks after sowing. Additionally, the leaf edge shape (LES: 1 - entire; 2 -slightly serrated; 3 - intermediate

serrated; 4 - much serrated) and density of leaf trichomes (LT: 0 - no trichomes; 1 - few trichomes; 2 - many trichomes) were scored.

The phytate and phosphate levels were determined using HPLC as described by Bentsink et al. (2003) with minor modifications. Data analysis for summary statistics, one-way analysis of variation (ANOVA) and correlation were performed in Genstat 8.1.

Genotyping

The *B. rapa* accessions in this study have been genotyped using AFLP fingerprinting (Zhao et al. 2005). In total, 437 AFLP scorable amplification products ranging from 50 bp to 500 bp were generated with 4 primer combinations (pAT/mCCA, pGG/mCAA, pAG/mCAC and pTA/mCAT). Of the 437 AFLP bands, 389 markers were polymorphic.

Map positions of markers were derived from an integrated map with AFLP and SSR markers that was based on two Double Haploid (DH) populations DH-38 (PC-175 X YS-143) and DH-30 (VT-115 X YS-143), sharing the common parent YS-143 (Lou et al. submitted). The three parental lines (Yellow sarson YS-143, Pak choi PC-175 and Vegetable turnip VT-115) are included in the AFLP fingerprinting study allowing the comparison of the AFLP markers based on the band size. Of the AFLP markers used to detect association among the 160 accessions, 76 were mapped on the integrated map, with 3 to 11 markers per chromosome (= R-linkage group). The remaining nearly 300 AFLPs could not be mapped because most of them were either not polymorphic or were absent in parental lines of the DH mapping populations.

Population structure

The program STRUCTURE version 2.1 was used to identify groups in the population, using a Bayesian approach (Pritchard et al. 2000; Falush et al. 2003). The accessions are classified into a pre-set number of clusters based on their allele frequencies, such that accessions within groups are in Hardy-Weinberg equilibrium and LD is found only between groups.

We tested a model with population admixture, which assumes that genotypes can have a mixed ancestry, and assumed independent allele frequencies between subpopulations. The number of subpopulations was set to vary between 1 and 10, and for each fixed number of subpopulations, 2 independent MCMCs (Markov Chain Monte Carlo) were run using 600,000 iterations for each, and the first 100,000 iterations were discarded as burn-in.

Analysis of LD between markers

In order to investigate the significance of linkage disequilibrium between pairs of markers, LD was calculated between all pairs of markers using Chi-square tests on two-by-two tables showing band presence or absence for the pairs of markers. Of the 389 polymorphic bands, 233 markers with allele frequencies larger than 6% were selected for further analysis.

To study heterogeneity in AFLP band frequencies across the different structured subpopulations, Fisher's exact test was applied to contingency tables of marker vs subpopulation using SAS software (SAS, 1999. SAS/STAT User's Guide. SAS Institute, Cary, NC.).

Association analysis of quantitative traits

Association analysis of quantitative traits (LN, NDF, VDF, LPHO, LPHY, SPHO and SPHY) was performed in three steps with a series of increasingly complex mixed models, and was carried out in Genstat (Payne and Arnold 2002) using restricted maximum likelihood (REML).

Model 1: phenotypic response = marker + error

This model corresponds to a series of simple t-tests, without correction for substructure and additional QTL present elsewhere in the genome. The design matrix corresponding to the fixed effects (marker) is a vector corresponding to the marker scores, i.e., a vector having the value 1 if a band is present and 0 otherwise.

Model 2: phenotypic response = structure + marker + error

This model corrects for population substructure by adding a random term to Model 1, named structure, containing the subgroup membership probabilities (Q matrix) obtained from STRUCTURE (Pritchard et al. 2000; Falush et al. 2003). The design matrix for the random term (structure) contains the membership probabilities for each subgroup. This model is very similar to that described in Yu et al. (2006), with the difference that here we use the membership probabilities matrix, instead of 0/1 scores as described in Yu et al. (2006). Our improved design matrix measures not only the differences among subgroups but also among accessions.

Model 3: phenotypic response = structure + co-factors + marker + error

This model corrects for population substructure like model 2, and additionally it includes as co-factors the set of identified putative QTL (markers) from step 2, after a cleaning-up by

backward selection. This method is like Composite Interval Mapping (Jansen and Stam 1994) in an association mapping context, and aims at reducing the genetic background noise.
Another set of traits, such as LES and LT, were measured as ordered categories, and for those traits association analysis was based on ordinal regression (Dobson 2002), using an analogue of model 3.

Results

Population structure

In a previous study (Zhao et al. 2005), cluster analysis using Unweighted Pair Group Method with Arithmetic averages (UPGMA) produced a phenetic tree that suggested two main groups based on the origin of the material (Asia versus Europe), plus a small group of spring oil types from mainly Bangladesh. The model-based approach of STRUCTURE suggests the presence of 4 subpopulations consisting of 3 large groups (Fig. 1): S1 consisting of 60 accessions, S2 consisting of 40 accessions, S4 consisting of 51 accessions, and one small group S3 with 9 accessions (Table 1). Most of oriental accessions are grouped into S1 and S4, and most of western accessions are grouped into S2, while the spring oil types group into S3.

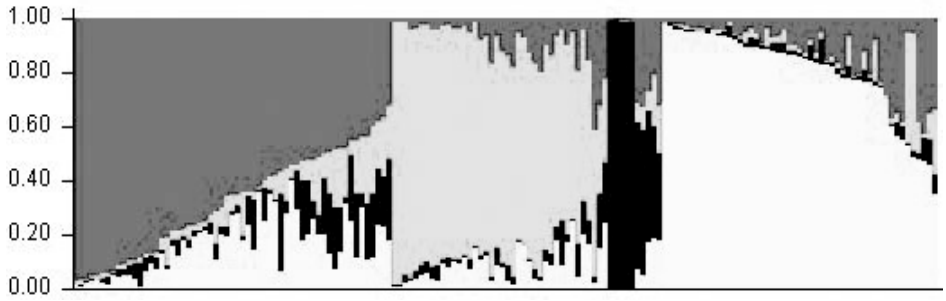

Fig. 1 Results from structure under the assumption of cluster number K=4.
Accessions are represented by a bar which is partitioned into several segments with different gray shades according to the individual's estimated membership fractions of the 4 clusters (S1, S2, S3, S4).

The S1 subpopulation of accessions encompasses all Pak choi types including two Caixin and one Wutacai accession and most of turnip rape accessions from China, four winter oilseed rape and one turnip accessions from Pakistan, nine Japanese and one Soviet Union turnips, three Japanese Neep Greens including two Komatsuna and one turnip green accessions. Two Mizuna accessions from Japan and the Netherlands are also grouped into S1 with an

admixture of S2, S3 and S4. S4 is another subpopulation of oriental origin and compasses 41 accessions of Chinese cabbage cultivars (CC) from Asia (China, Korea and Japan), with 7 western accessions. The S4 group also includes one French turnip accession with S2 admixture, plus one Chinese turnip rape and one Japanese turnip green admixed with S1 and S2. The accessions of subpopulation S2 mainly originate from western countries, including 25 European turnips and 3 turnips from Uzbekistan and India, 6 Italian Broccoletto and 6 oil types from America, Canada, Sweden and Pakistan. The small but distinct subpopulation S3 is formed by spring oil types including 2 Yellow Sarson's originating from India, 6 spring oilseed rape cultivars from Bangladesh and one rapid cycling line from the USA, which probably is derived from these Sarson types from the Indian sub continent.

Table 1 Composition of subpopulations S1-S4 is listed, with number of accessions per cultivar group and their origin

Subpopulation	Total	Cultivar group											Origin	
		T	BRO	CC	PC	NG	MIZ	YS	RC	WO	SO	OR	East	West
S1	60	10	0	0	32	3	2	0	0	4	0	9	56	4
S2	40	28	6	0	0	0	0	0	0	4	2	0	5	35
S3	9	0	0	0	0	0	0	2	1	0	6	0	5	4
S4	51	1	0	48	0	1	0	0	0	0	0	1	43	8

[T], turnip; [BRO], Broccoletto; [CC], Chinese cabbage; [PC], Pak choi, Caixin and Wutacai; [NG], Neep green; [Miz], Mizuna; [YS],Yellow Sarson; [RC], rapid cycling; [WO], Winter turnip rape; [SO], Spring turnip rape; [OR], Chinese turnip rape

Analysis of marker-marker associations

We found that 3.1% of all AFLP marker pairs had a significant correlation at α=0.01. Out of the 233 polymorphic markers, 119 markers were associated with the 4 structured subgroups ($p < 0.0001$) thus having influence on differentiation between the cultivar groups.

Variation in observed traits

The distributions of traits and the correlations between traits are shown in Fig. 2, separately for the subgroups S1 to S4. Statistics of all observed traits, organized in different subpopulations is summarized in Table 2.

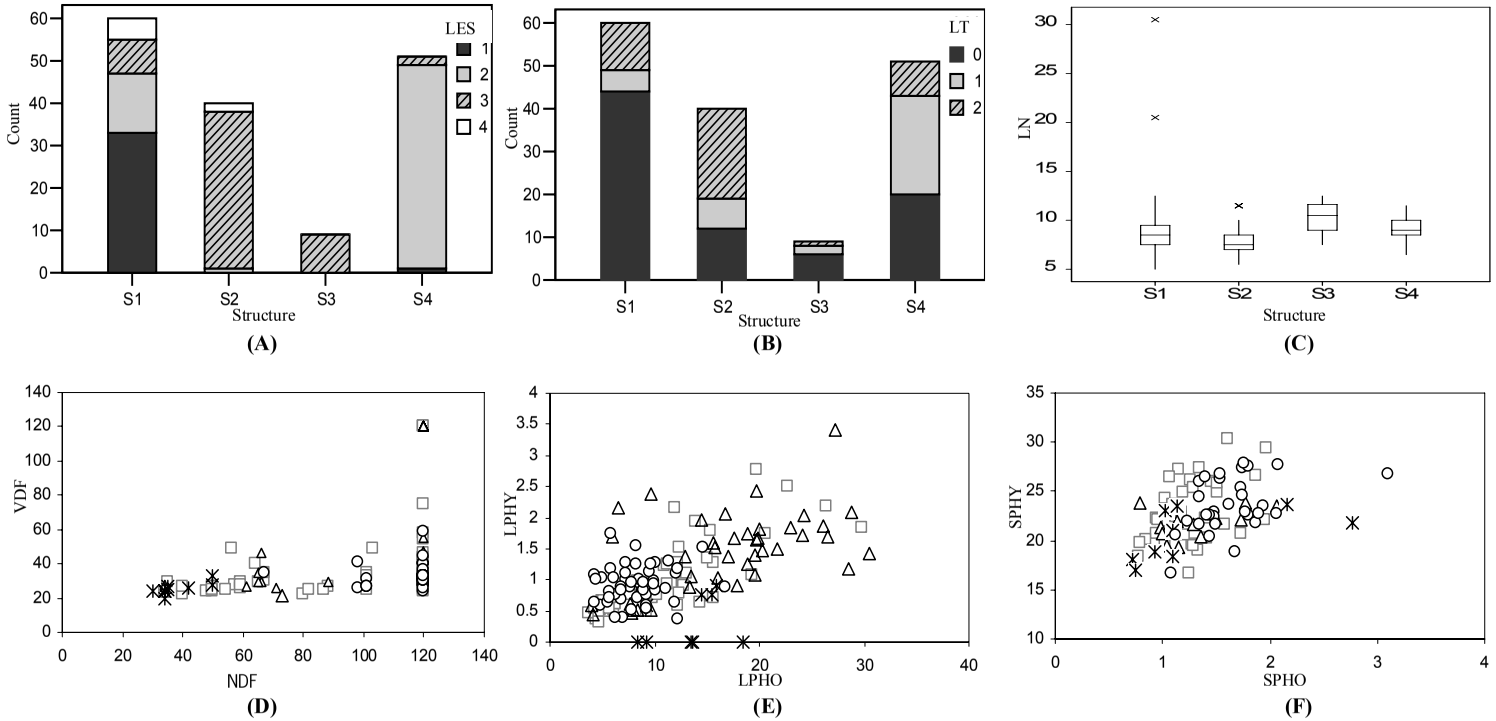

Fig. 2 Natural variation of leaf edge shape (LES: 1-smooth entire; 2-very light dissected; 3-few dissected; 4-much dissected) in **(A)**, leaf trichomes (LT: 0-no trichomes; 1-few trichomes; 2-many trichomes) in **(B)**, leaf number (LN) in **(C)**, non-vernalization flowering time (NDF, days) and flowering time after vernalization (VDF, days) in **(D)**, phytate and phosphate in leaves (LPHY and LPHO, mg/g) in **(E)** and seeds (SPHY and SPHO, mg/g) in **(F)** of 160 *B. rapa* accessions.

The different symbols refer to the different subgroups as illustrated in Fig. 1 (□ S1 △ S2 ✳ S3 ○ S4).

Table 2 Statistical description of observed traits in the subpopulations S1-S4

Traits	S1			S2			S3			S4		
	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
LES**	1.8	1.0	1-4	3.0	0.3	2-4	3.0	0.0	3-3	2.0	0.2	1-3
LT**	0.5	0.8	0-2	0.2	0.9	0-2	0.4	0.7	0-2	0.8	0.7	0-2
LN*	9.1	3.6	5-31	7.9	1.3	6-12	10.2	1.6	8-13	9.2	1.1	7-12
NDF**	96.3	30.7	34-120	105.6	26.5	34-120	38.2	7.4	30-50	116.6	9.6	67-120
VDF**	36.2	19.1	22-120	88.0	42.2	21-120	25.7	1.2	20-33	32.6	5.7	24-59
LPHY**	1.0	0.6	0.3-2.8	1.5	0.6	0.4-3.4	0.3	0.4	0-0.9	0.9	0.3	0.4-1.74
LPHO**	10.7	5.9	3.7-29.8	17.0	7.2	4.0-30.5	13.1	3.5	8.3-18.5	11.2	6.1	4.2-16.7
SPHY*	22.9	3.3	16.7-30.4	21.7	1.6	19.2-23.8	20.6	2.6	16.9-23.6	23.7	2.8	16.6-27.9
SPHO**	1.3	0.3	0.8-1.9	1.3	0.4	0.8-2.1	1.3	0.7	0.7-2.8	1.6	0.4	1.1-3.1

*Significant level of one-way analysis of variation (ANOVA) between subpopulations at $P < 0.05$ * and $P < 0.01$**; SD, Standard Deviation; The scale of traits: number for LN, days for NDF and VDF, mg/g for LPHY, LPHO, SPHY and SPHO; See Fig. 2 for definition of trait abbreviations.

Leaf characteristics including leaf edge shape (LES), leaf trichomes (LT) and leaf number (LN) are important morphological traits distinguishing vegetable *B. rapa* types. The distribution of LES and LT is related to the structured subpopulations (Fig. 2A, B; Table 2). More than 50% of the accessions in S1 have entire leaves (LES-1), most of accessions in S2 and S3 have moderately serrated leaves (LES-3), most of accessions in S4 have very lightly serrated leaves (LES-2), and one Mizuna and some winter oil accessions in S1 and S2 have severely serrated leaves. The different classes of leaf trichome frequency (LT-0, LT-1, LT-2) are distributed within each subpopulation, however most of accessions in S1 (mainly Pak choi's) have a hairless leaf surface and most of accessions in S4 (Chinese cabbage) have few or no trichomes. The variation for LES and LT between different subpopulations was significant at $P < 0.01$. For LN, the range in S2, S3 and S4 is similar from 8 to 13 leaves (Fig 2C; Table 2). Within S1, the variation of LN is higher (5-31) because of two Mizuna and one Wutacai accessions with many leaves. The mean value of LN value in S2 is lower although two accessions (one spring and one winter oil type) had a high value of 11.5. The variation for LN between different subpopulations was different with $P = 0.01$.

Flowering time is a very important developmental trait in *B. rapa* and wide variation was observed among the collections in days to flowering; more than 3-fold difference was found between accessions without vernalization (NDF, 30-103 days) and with vernalization (VDF, 20-75 days) when non-flowering plants were excluded (Fig 2D; Table 2). Under non-vernalization conditions, only 52 accessions flowered including 26 accessions in S1 (15 Pak Choi, 2 Caixin, 2 Komatsuna, one vegetable turnip, 4 winter turnip rape and 2 Chinese oil type cultivars), 6 Broccoletto and 4 oil type accessions in S2, all spring oil accessions in S3, and 6 Chinese cabbage cultivars and one Chinese oil type in S4. Late forms of Chinese cabbage in S4 responded strongly to vernalization, whereas only 8 Japanese turnips in S1 and 2 other turnips in S2 flowered upon vernalization. Most turnip accessions in S2 and 2 turnips in S1 did not flower after vernalization, which may indicate that these accessions require a longer period of cold or vernalization at a later stage of development to induce flowering. The differences in flowering time are also associated with population structure as illustrated in Table 2. The variation for NDF and VDF between different subpopulations was significant at $P < 0.01$.

The levels of phytate in seeds were 10 times higher than in leaves. However phosphate levels in leaves were 10 times higher than in seeds. Comparing the accessions, a positive correlation between the two compounds was detected at $P < 0.01$, with a correlation coefficient of $r = 0.52$ in leaves and of $r = 0.44$ in seeds (Fig. 2E, F). The correlation between phytate in seeds and phosphate in leaves was low ($r = -0.21$) and not significant. In leaves, the variation of phytate (LPHY, 0 - 3.4 mg/g) and phosphate (LPHO, 3.7 - 30.5 mg/g) was about ten-fold. In some oil type accessions, phytate levels were below detection level in leaves. Variation in seeds was less, being two-fold for phytate (16.7 - 30.4 mg/g) and four-fold for phosphate (0.7 - 3.1 mg/g).

Variation for phytate and phosphate levels was observed within each subpopulation (Fig. 2E, F; Table 2). The variation between different subpopulations was significantly different at $P < 0.05$ for SPHY and at $P < 0.01$ for SPHO, LPHY and LPHO. However for SPHO and SPHY, only 87 accessions were evaluated since many accessions (mainly turnips in S2) did not produce enough seeds. Although the phytate and phosphate concentration in Chinese cabbages in S4 and spring oil accessions in S3 was lower than that of Pak choi's in S1 and turnip's in S2, the range of variation in each subgroup is overlapping in both seeds and leaves,

which made it possible to analyze the association of markers and traits within this *B. rapa* collection.

Association mapping

Association between markers and quantitative traits was examined using three models: simple t-test (model 1), a model correcting for population structure (model 2), and a Composite Interval Mapping procedure (model 3).

Fig. 3 illustrates that the three different models detect quite different numbers of markers that are strongly associated with the traits analyzed. Using the t-test, 185 markers (from 28 to 112/trait) were found associated to observed traits, with many markers for LPHY, LPHO, NDF and VDF. When population structure was used in the mixed model 2 and 3 to correct for spurious associations, the number of markers associated with the different traits was much lower. The Composite Interval Association Mapping procedure, with only few markers (1 - 12) detected per trait, which in addition reduces the genetic background noise, seems the most appropriate procedure.

Table 3 shows an overview of all significant associations between measured traits and AFLP markers using the Composite Interval Mapping procedure. The linkage group of the associated markers is listed together with previously identified QTL in the same group. Totally, 54 markers, of which 16 had known map positions, were associated with the 9 traits at $P < 0.05$.

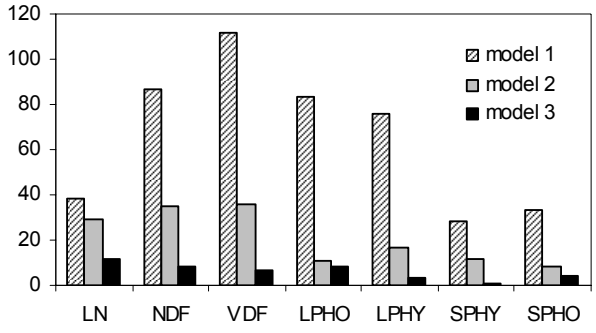

Fig. 3 Number of markers associated with traits as resulting from three different models

Table 3 An overview of all significant associations ($p < 0.05$) between measured traits and AFLP markers using the Composite Interval Mapping procedure

Markers	Linkage group	Previous QTL*	Significant p value in association with various traits								
			LES	LT	LN	NDF	VDF	LPHO	LPHY	SPHY	SPHO
pAG/mCAC0137.6	R02	SPHY, SPHO						0.028			
pAG/mCAC0323.4	R02	FLC2					0.036				
pGG/mCAA0165.8	R02	LN			0.020						
pTA/mCAT0174.4	R03				0.002						
pGG/mCAA0339.8	R05			0.015			0.000				
pTA/mCAT0220.0	R05	SPHY, SPHO								0.003	
pAT/mCCA0163.9	R05				0.014						
pGG/mCAA0411.1	R06	SPHY, SPHO									0.001
pAT/mCCA0114.7	R06	SPHY							0.015		
pAG/mCAC0316.7	R07	LPHY						0.024	0.029		
pAG/mCAC0244.5	R07			0.045							
pAT/mCCA0430.9	R07	LN			0.018						
pAG/mCAC0459.3	R08			0.000							
pAG/mCAC0175.5	R08			0.022							
pAG/mCAC090.5	R08		0.002								
pTA/mCAT0265.2	R10							0.021			
pAG/mCAC0465.9	unmapped			0.001							
pAG/mCAC0454.5	unmapped										0.047
pAG/mCAC0422.0	unmapped					0.005					
pAG/mCAC0386.2	unmapped					0.017					
pAG/mCAC0364.4	unmapped										0.010
pAG/mCAC0274.5	unmapped			0.000							
pAG/mCAC0255.5	unmapped				0.038						
pAG/mCAC0253.3	unmapped					0.045					
pAG/mCAC0247.5	unmapped						0.035				
pAG/mCAC0218.5	unmapped							0.010			
pAG/mCAC0216.7	unmapped			0.039							
pAG/mCAC0148.6	unmapped							0.021			
pAG/mCAC0132.4	unmapped			0.018							
pAG/mCAC085.6	unmapped		0.001								
pAG/mCAC073.0	unmapped				0.003	0.023					
pAG/mCAC053.6	unmapped										0.030
pGG/mCAA0449.1	unmapped		0.026								
pGG/mCAA0329.6	unmapped					0.005					
pGG/mCAA0188.7	unmapped			0.013							
pGG/mCAA0181.3	unmapped		0.003								
pTA/mCAT0419.3	unmapped				0.016						
pTA/mCAT0362.5	unmapped							0.001	0.044		
pTA/mCAT0312.9	unmapped				0.038						
pTA/mCAT0280.4	unmapped		0.027		0.031						
pTA/mCAT0152.1	unmapped				0.000						
pTA/mCAT0148.9	unmapped		0.005								
pTA/mCAT077.7	unmapped						0.004				
pTA/mCAT071.6	unmapped						0.009				
pTA/mCAT068.7	unmapped					0.041					
pAT/mCCA0420.6	unmapped							0.027			
pAT/mCCA0395.0	unmapped						0.012				
pAT/mCCA0355.8	unmapped					0.013					
pAT/mCCA0264.7	unmapped		0.004								
pAT/mCCA0210.6	unmapped						0.018				
pAT/mCCA0184.2	unmapped				0.007						
pAT/mCCA0175.1	unmapped		0.046								
pAT/mCCA089.6	unmapped		0.032		0.030						
pAT/mCCA076.9	unmapped					0.007		0.038			

*Candidate genes or QTL identified in previous studies in similar genomic region. QTL were identified in QTL analysis in 4 DH populations, one F2 population and one BC1 population (Chapter 4 and 5), and FLC2 as candidate gene for flowering QTL as described by Schranz et al. (2002), Kim et al. (2006) and chapter 5 of this thesis. See Fig. 2 for definition of trait abbreviations.

For leaf traits (LES, LT and LN), two unmapped AFLP markers (pTA/mCAT0280.4 and pAT/mCCA089.6) were associated with both LES and LN. Of the 9 markers associated with LES, only one marker (pAG/mCAC0090.5) was mapped, namely at the top of R08. Four markers associated with LT were mapped on R05, R07 and R08. Leaf number was associated with 12 markers, 4 of which had known map position and were distributed over R02, R03, R05 and R07 (Table 3). One unmapped AFLP marker (pAG/mCAC073.0) was associated with both LN and NDF, which suggests a correlation between the two traits. Several markers were associated with days to flowering. However, the same associations were not found for NDF and VDF, which indicated that the two traits were not strongly correlated. Seven markers were associated to VDF. One associated marker (pAG/mCAC0452.5) was located on R02, close to the position of a flowering gene *FLC* that had been reported before (Kim et al. 2006; Chapter 5). Another associated marker (pGG/mCAA0339.8) was located on R05. The remaining 8 markers correlated to NDF had no known map position.

For phytate and phosphate levels in seeds and leaves, 14 associated AFLP markers were detected with $P < 0.05$. One mapped marker (pAG/mCAC0316.7) and one unmapped marker (pTA/mCAT0262.5) were associated with both LPHY and LPHO, which illustrates close linkage or pleiotropy of the two traits. However, association with the same markers with SPHY and SPHO was not detected. For 8 markers associated with these four traits, the map position was known and some of them were confirmed in other QTL mapping studies. The marker pAG/mCAC0316.7 related to LPHO and LPHY was mapped on R07, co-localizing with a strong QTL region related to LPHY on R07 based on three DH populations analyzed in a previous study (Chapter 4, this thesis). Two markers located on R02 and R10 (pAG/mCAC0137.6 and pTA/mCAT0265.2) were associated with LPHO. Furthermore, for a marker (pAT/CCA0114.7) mapped on R06, an association to LPHY was detected with $P = 0.015$. Four markers were associated to SPHO, but only one of them (pGG/mCAA0411.1) was mapped on R06, a region where a QTL was detected for SPHY/SPHO in a previous study (Chapter 4, this thesis).

Discussion

In the present study we analysed a number of traits in a set of *B. rapa* genotypes representing the various cultivar types in the germplasm of this species. For these genotypes AFLP analyses had been performed, which indicated a loose population structure based on UMPGA

(Zhao et al. 2005 and chapter 2 of this thesis). This analysis showed two main subgroups, one originating from East Asia, which could be subdivided in Chinese cabbage (CC) and Pak choi (PC) groups and winter oil types from Pakistan, and another group originating from Europe, which mainly include turnips and Broccoletto types and a small subgroup consisting of early oilseed types including the Yellow Sarson types from India. Based on allele frequencies, the analysis of the same AFLP data set with the STRUCTURE program confirmed the above result and suggested 4 subgroups (Pak choi group S1, turnip group S2, spring oil group S3 and Chinese cabbage group S4). The structured subgroups S1 and S4 mainly belong to the UPGMA Group 1, and the structured subgroup S2 mainly belonged to the UPGMA Group 2. From the STRUCTURE results, the admixture between accessions could also be detected. For example several Pak choi accessions and Chinese oil types in S1 share genetic background with S4, which is possibly related to their breeding history.

This subdivision was taken into account in the analysis of variation of a number of traits determined in these materials. The number of leaves (LN) was evaluated at the four-weeks-old plant stage, which does not reflect the whole vegetable developmental process because some vegetable types form more leaves during later development. Flowering time of *B. rapa* species used as vegetables or turnips is agronomically important because it relates to yield and quality. Flowering time was assessed under vernalization and non-vernalization treatment in the present study. Chinese cabbage and turnip types displayed a different vernalization requirement compared to other cultivar groups which suggests that different genes affecting flowering time are present in these groups. Vernalization greatly reduced the range of variation in flowering time when non-flowering plants were excluded.

Considerable variation in phytate and phosphate accumulation was observed. The extensive variation of leaf phosphate might be used to breed for better phosphate use efficiency. Phytate content is relevant for oilseed types and a two-fold range of variation in seed phytate level exists. We also observed a positive correlation of phytate and phosphate as has been reported in Arabidopsis (Bentsink et al. 2003) and corn (Raboy et al. 2001). Despite this general correlation, a few accessions were identified with relatively high phosphate and low phytate levels in seeds, such as SO-032 (phytate 21.8 mg/g and phosphate 2.8 mg/g) and WO-082 (phyate 18.3 mg/g and phosphate 3.6 mg/g). For leaf content also some genotypes with a strongly altered relationship between phosphate and phytate levels were found. VT-015 and

WO-024 have a higher phosphate level (30.5 mg/g and 28.6 mg/g) but a low phytate level (1.43 mg/g and 1.19 mg/g), and all spring oil accessions in S3 have lower or non-detectable phytate levels (from 0 to 0.9 mg/g). In future *Brassica* breeding programs, it is possible to combine high phosphate with low phytate levels, and to select ideal genotypes as parents of mapping populations for QTL identification.

Naturally occurring genetic variation is a useful resource for the genetic mapping of complex phenotypic traits (Koornneef et al. 2004). We applied association mapping in *B. rapa* for identification of genetic markers associated with leaf traits, flowering time and phosphate levels, and to compare the outcome of association mapping with QTL detected in DH populations that we developed for this purpose. The presence of population structure may affect LD and produce false positives. The associations among markers themselves were also examined; markers that differ in allele frequency between subpopulations provide an example of LD due to population structure. Some of these markers may be causally responsible for observed phenotypic differences between the groups. However, marker frequencies between groups can also differ due to chances. We cannot discriminate between chance and non-chance associations at the level of the phenotypic differences between groups, but differences in marker frequency are explored in the program STRUCTURE to form the groups. Since trait values differed significantly between subgroups, an association between population structure and these trait values is suggested even for traits such as phytate and phosphate levels, for which we assumed no selection had occurred.

Markers associated with the traits analyzed are presented for the model 1 to 3. These markers that are causally responsible for the phenotypic differences but also related to structure, should be listed in the outline of model 1. Markers that show association after correction for substructure can reliably be interpreted as being linked to QTL. More than 20 markers were found associated with leaf traits (LES, LT and LN) by Composite Interval Mapping, and in addition 15 markers were associated with flowering time. An earlier report about morphological variation in *B. rapa* (Song et al. 1995) described that the degree of pubescence (presence of leaf hair) was controlled by polygenes and three related QTL in 3 linkage groups were found. However, since these linkage groups were not assigned it is not possible to compare these QTL to the LT-associated markers located on R05, R07 and R08 in our study. Fifteen markers associated with flowering time were detected in this study, of which one

VDF-associated marker was located on R02, at a position where a flowering time QTL was found (Schranz et al. 2002).

The traits studied in this paper have also been analysed in a set of mapping populations including 8 parents (Chapter 4 and 5 of this thesis). These QTL analyses identified a number of QTL related to the different traits and co-locations with associated loci identified in this study. For LPHY, we found one significant QTL on R07 in 3 DH populations analyzed likely because the common parent Yellow sarson YS-143 has very low LPHY value compared to the other three parental lines. In the present study, one LPHY-associated marker located on R07 was also identified.

Since association mapping will also allow the analysis of variation in multiple genotypes it is expected that additional associated loci will be identified. However most of the AFLPs in this study are not mapped we cannot conclude if these associated markers refer to additional genetic positions. One objective of future studies is to use mapped markers for association mapping. Many SSR markers and increasing sequencing information of *B. rapa* is already available (http://www.brassica.info), which makes it possible to profile SSR and gene target markers across all accessions which allows determination of the LD level across the genome, facilitating the identification of QTL in *B. rapa*. A characteristic of association mapping is that only those alleles that have a sufficient high frequency can yield significant association implying that rare alleles, even when they are strongly linked to the trait, will remain undetected.

Confirmation of some of the marker-trait associations by QTL analysis indicated that association mapping allows the detection of linkage with moderately frequent alleles, which thereafter can be confirmed by linkage analysis in mapping populations.

Acknowledgements

We thank the Wageningen Plant Sciences Experimental Centre of Wageningen University for taking care of the plants. We thank Noortje Bas from the Centre for Genetic Resources the Netherlands (CGN) Wageningen-UR, Li Xixiang from the Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Liu Shengyi, Oil Crop Research Institute, Chinese Academy of Agricultural Sciences, and Dr. Osborn (University of Wisconsin, Madison, Wis., USA) for kindly supplying the accessions used in this study. The project is sponsored by the Asian Facility (project AF01/CH/8 “Sino-Dutch Genomic Lab and

Vegetable Research Center") and a fellowship from an exchange student programme (Huygens) between Chinese Scholarship Council and Nuffic in the Netherlands.

Chapter 4

QTL analysis of phytate and phosphate content in seeds and leaves of *Brassica rapa*

Jianjun Zhao, Diaan Jamar, Ping Lou, Xiaofei Song, Ying Li, Jian Wu, Xiaowu Wang, Guusje Bonnema, Maarten Koornneef, Dick Vreugdenhil

Abstract

Phytate, being the major storage form of phosphorus in plants, is considered to be an anti-nutritional substance for human, because of its ability to complex essential micronutrients. In the present study we describe the genetic analysis of phytate and phosphate accumulation in *Brassica rapa* using five segregating populations, involving 8 parental accessions representing different cultivar groups. A total of 27 QTL affecting phytate and phosphate contents in seeds and leaves were detected in the used populations, most of them located on linkage groups R01, R03, R06 and R07. Two QTL affecting seed phytate, 2 QTL affecting seed phosphate, 1 QTL affecting leaf phosphate, and 1 major QTL affecting leaf phytate were detected in at least 2 populations. Co-localization of QTL suggested single loci to be involved in the accumulation of phytate or phosphate in seeds or leaves. Some co-localizing QTL for seed phytate and seed phosphate had parental alleles with effects in the same direction suggesting that they control total phosphorus level. For other QTL the allelic effect was opposite for phosphate and phytate suggesting that these QTL are specific for the phytate pathway.

Introduction

Phosphorus is an essential element for all living organisms. The major form in which phosphorus is stored in seeds of plants is *myo*-Inositol-1,2,3,4,5,6-hexakisphosphate (IP6, or phytic acid), which release phosphorus and *myo*-Inositol during seed germination. Phytic acid especially accumulates in seeds in which it can account for up to several percent of seed dry weight and about 65-85% of seed total phosphorus (Raboy 2001, 2003). The highly negatively charged phosphate groups in IP6 form a complex with cations, such as the essential minerals potassium, magnesium, iron and zinc to form phytate. Therefore a high level of phytate in plant tissue causes problems, since essential elements are bound in phytate and hence not bioavailable for humans and animals. In non-ruminant animals (pig, chicken) a large portion of total phosphorus, supplied in feed, is excreted, contributing to soil and water pollution in the regions with high concentrations of non-ruminant animals production farms, which is a significant problem in some developed countries. Secondly, high phytate levels often cause micronutrient (iron and zinc) deficiencies in humans when their mineral supply largely depends on seed-derived diets such as cereals. The latter problem is most prominent in developing countries.

To reduce the phytate concentration in food and feed and to increase the bio-available contents of essential elements, several strategies have been adopted. The traditional way is to supply additional elements or ingredients to food and feed as a supplementation. However, this approach does not alleviate the related problems such as environmental pollution caused by animals. Recently, an alternative solution for increasing mineral contents in edible portions of crop plants termed ‘biofortification’ has been proposed (Bouis 2003; Welch and Graham 2004). Transgenic methods can be used to break down phytate and release free elements via knocking out enzymes of the phytate biosynthetic pathway or by over-expressing phytase in edible parts (White and Broadley 2005). In addition, breeding for low phytate accumulation but also for higher micronutrient content is considered as a possible solution of this problem (Raboy 2001). Low phytate mutants have been obtained after chemical mutagenesis in maize (Raboy and Gerbasi 1996; Raboy et al. 2000; Pilu et al. 2003), barley (Larson et al. 1998; Rasmussen and Hatzack 1998), rice (Larson et al. 2000), soybean (James et al. 2000; Wilcox et al. 2000; Hitz et al. 2002) and wheat (Guttieri et al. 2004). These mutants are characterized by reduced phytate levels, which is matched by increased inorganic phosphorus contents,

thereby retaining the same total phosphorus levels as in wild type. The biochemical and molecular characterization of the maize mutants revealed that the *lpa1* mutant phenotype is associated with the reduced expression of an inositol-1-P synthase (*MIPS*) gene, one of the first steps of the phytate biosynthesis pathway (Shukla et al. 2004), The *lpa2* mutant is associated with reduced expression of an inositol phosphate kinase (*ZmlPK*) gene, which encodes an enzyme affecting the phytate biosynthesis pathway downstream of inositol biosynthesis (Shi et al. 2003). The *LPA3* gene encodes an inositol kinase (*MIK*) that plays a role in phytate biosynthesis but not in inositol phosphate intermediates in developing seeds (Shi et al. 2005).

In addition to mutants, natural variation can serve as starting material for gene identification (Koornneef et al. 2004). Natural variation in Arabidopsis is abundantly present among the many accessions of this species found all around the world. Using a recombinant inbred line (RIL) population derived from the accessions L*er* and Cvi, some genomic regions controlling phytate and phosphate contents in seeds and leaves and mineral levels in seeds of Arabidopsis were identified (Bentsink et al. 2003; Vreugdenhil et al. 2004). It was found that the accumulation of phytate and certain minerals in seeds can be separated genetically which indicates the possibility to breed for reduced phytate content without affecting micronutrients levels (Vreugdenhil et al. 2004).

The *Brassica* genus, comprising a large and diverse group of plant species, is closely related to *Arabidopsis thaliana*, which is a member of the same Brassicaceae family. *Brassica rapa* is an important vegetable and oil crop with variation for many different morphological characteristics. Within this species, genetic variation exists for agronomic characteristics but also for nutritional components (Zhao et al. 2005). Little is known about the contribution by *Brasscia* vegetables to the mineral supply in the human diet (Ma et al. 2005). In China, rapeseed oil represents approximately 35% of edible oil consumption and in recent years rapeseed meal accounts for about 25% of plant seed meal consumption by animals (Wang 2004). Today “double low” (low content of erucic acid and glucosinolates) commercial varieties of oilseed rape dominate the oilseed *Brassica* production area in the world, and their nutritional value is being improved. However, the phytate level is still about 2.5-3.0% in “double low” rapeseed varieties and the phosphorus in phytate represents 75-80% of seed total phosphorus content, which implies that the utility of seed meal is limited (Peng et al.

2001). In addition to their use as oilseed crop, a wide range of *Brassica* species and varieties is also used as vegetables and provides a useful resource for phosphate and minerals. Phytate levels are generally low in non-seed tissues. Until now, the use of plant breeding to reduce phytate levels and increase the available minerals has not been exploited in *Brassica* species.

In the present study we describe the genetic analysis of phytate and phosphate accumulation in *B. rapa* using five segregating populations. The aim is the identification of putative loci involved in regulation of phytate and phosphate levels, which can be the basis for improvement of the nutritional quality of this important vegetable and oilseed crop species. A number of *B. rapa* accessions representing a range of geographical origins and cultivar groups have been collected and analyzed for their genetic diversity using AFLP fingerprinting (Zhao et al. 2005), and the variation of this collection for phytate and phosphate levels in seeds and leaves has also been established (chapter 3, this thesis). The data indicated that there is ample variation between different accessions for phytate and phosphate in both seed and leaves, thus quantitative trait locus (QTL) mapping of phytate and phosphate accumulation in *B. rapa* seems feasible. Identification of QTL will facilitate breeding for phytate and phosphate both in vegetable and oil seed *B. rapa* and possibly also in *B. napus*. Anchored SSR markers will allow comparison of QTL between different genetic maps. The syntenic relationship of *Brassica* to the model plant Arabidopsis allows a direct comparison of map positions of the two species (Parkin et al. 2005; Suwabe et al. 2006; Schranz et al. 2006) and might assist the identification of candidate genes that are already known or will be known in Arabidopsis.

Materials and methods

Population development

Several mapping populations were developed from wide crosses between *B. rapa* accessions. The parents of crosses represent different cultivar groups in *B. rapa* that are polymorphic for both morphological traits and AFLP pattern (Zhao et al. 2005).

One F2/3 population (RC-CC F2/3) produced from selfing of a single F1 plant, resulting from a cross between a Rapid cycling line RC-144 (accession number: FIL501) and a vegetable type Chinese cabbage line CC-156 (cultivar: Huang Yang Bai; accession number: VO2A0030) was analysed. This population consists of 178 F2 individuals. Seeds of F2 were grown in pots and seedlings with 4-5 leaves were transferred to soil in a greenhouse of the Institute of Vegetables and Flowers, Chinese Academy of Agriculture Science during the

spring of 2004. Full grown leaves of F2 plants and mature F3 seeds were used for phytate and phosphate analysis.

Furthermore, three double haploid (DH) populations were developed from crosses between the oil type Yellow Sarson YS-143 (accession number: FIL500) and the vegetable types: Pak choi PC-175 (cultivar: Nai Bai Cai; accession number: VO2B0226), the Asian Vegetable turnip VT-115 (cultivar: Kairyou Hakata; accession number: CGN15199) and Mizuna MIZ-019 (cultivar: Bladmoes; accession number: CGN06790) using microspore culture according to Custers et al. (2001). A total of 165 lines including 71 lines from population DH-38 (PC-175 X YS-143), 64 lines from population DH-30 (VT-115 X YS-143) and 30 lines from population DH-03 (MIZ-019 X YS-143) were analyzed for phytate and phosphate in seeds and leaves. The construction of these lines and their genotyping are described elsewhere (Lou et al. submitted). These DH plants (5 plants per line) were grown in a greenhouse under uniform soil conditions during autumn and winter (September to December) of 2004 in Wageningen University.

An additional large mapping population of 183 double haploid (DH) lines, called DH-CC, was developed by the Institute of Horticulture Science of Henan Academy of Agriculture Science in China (Zhang et al. 2001). DH-CC is derived from a cross between a Chinese cabbage DH line obtained from the Japanese cultivar CC-Y177 and a Chinese cabbage DH line from the Chinese cultivar CC-Y195. For phytate and phosphate analysis, seeds were harvested from the 183 DH-CC lines, which were grown in soil in a greenhouse during spring (January to April) of 2005 in the Institute of Vegetables and Flowers (Beijing), Chinese Academy of Agriculture Science.

Analysis of phytate and phosphate

The HPLC analyses of phytate and phosphate in seeds and leaves were performed as described by Bentsink et al. (2003) with minor modifications. Ten to fifteen mature seeds per line, harvested from individual F2 plants for the RC-CC F2/3 population, and from one plant per line for all DH populations were used. Before flowering, one leaf was sampled and lyophilized, the leaf being collected from one F2 plant per line for the F2/3 population. For the DH-38, DH-30 and DH-03 populations, one leaf was collected from each plant of each accession and leaves of two plants were ground together to represent one biological replication.

All samples were ground in a membrane disruptor for 20-60" at 1500 rpm. The samples varied from 5.0 - 7.0 mg and were extracted by boiling for 15 min in 0.5 ml of 0.5 N HCl and 50 mg/l cis-aconitate as an internal standard. The extracts were centrifuged at 14,000 rpm for 10 min. The supernatants were diluted 10-times (seeds) or 5-times (leaves) with ultrapure water, and 20 µl was analyzed using a Dionex ICS2500 HPLC system (Dionex Corporation, Sunnyvale, Calif., USA). Anions were separated on an AS 11 (4 x 250 mm) column at 30°C, preceded by an AG 11 guard column and eluted with a NaOH gradient. The elution profile was: 0-3 min isocratic at 5 mM of NaOH, followed by a 3-15 min linear gradient with 5-100 mM of NaOH. After each run the column was washed for 15-20 min with 0.5 M NaOH, followed by a 20-35 min equilibration at 5 mM. Flow rates were 1 $ml.min^{-1}$ throughout the run. Contaminating anions in the eluents were removed using an ion trap column (ATC), installed between the pump and the sample injection valve. Anions were determined by conductivity detection. Background conductivity was decreased using an ASRS suppressor, with water as a counterflow (5 $ml.min^{-1}$), operated at 248 mA, controlled by an SRS controller (Dionex Corporation Sunnyvale, Calif., USA). Peaks were identified and quantified by co-elution with known standards.

Construction of a genetic map of the RC-CC F2/F3 population

For the RC-CC F2/3 population, genomic DNA was extracted from fresh leaves of F2 plants according to the procedure described by Van der Beek *et al*. (1992). Fresh leaf tissue was ground by shaking tubes containing leaf material and iron bullets in a Retsch shaker at maximum speed (Retsch BV, Ochten, The Netherlands). The AFLP procedure was performed as described by Vos *et al.* (1995). Total genomic DNA was digested using two restriction enzymes, *EcoR* I and *Mse* I, and ligated to adaptors. Pre-amplifications were performed with *EcoR* I + A/*Mse* I + C primers. Five-microliter of the twenty fold diluted pre-amplification product was used as template for the selective amplification. Only *EcoR* I primers were labeled with IRD-700 or IRD-800 fluorescent dyes at the 5´ end for the selective amplification. Following the selective amplification, the reaction products were mixed with an equal volume of formamide-loading buffer (98% formamide, 10 m M EDTA pH 8.0 and 0.1% Bromo Phenol Blue), denatured for 5 min at 94°C, cooled on ice and run on a 5.5% denaturing polyacrylamide gel with a NEN® Global Edition IR2 DNA analyzer (LI-COR® Biosciences, Lincoln, NE). The AFLP gel images were mainly scored as dominant markers on

the basis of the presence or absence of the band at a corresponding position among the segregating population. When two polymorphic bands are derived from different parents within the same primer combination and segregate complementary, the two polymorphic bands were assigned as two alleles from one co-dominant marker (Alonso-Blanco et al. 1998). Two allelic segregating bands of this type were manually scored as one co-dominant marker. Only clear and unambiguous bands in the range of 50 bp to 500 bp were scored for genotyping. Segregating AFLP markers in the mapping population were named according to the primer combinations employed, followed by the parental line from which they were derived.

Public SSR primer pair sequences information of *Brassica* was obtained from the *Brassica* information website (http://www.brassica.info), and previous publications (Suwabe et al. 2002; Kim et al. 2006).

PCR reactions were performed in 96-well plates in a volume of 10 ul. The composition of the mix included 1 unit of Taq DNA polymerase, 5 mM of dNTP, 2.5 ul 10x supertaq buffer and 50 ng of each primer (forward and reverse primers). DNA was present in the PCR reaction to a concentration of 1 ng/ul. The PCR was performed on GeneAmp PCR system 9700 (Applied Bio-system) with the following program: 94 °C for 2 min; 10 cycles with 94 °C denaturation for 1 min, 65 °C annealing for 1 min, 72 °C elongation for 1.5 min, with a 1 °C decrease in annealing temperature at each cycle; then 30 cycles with 94 °C denaturation, 55 °C annealing, 72 °C elongation, 1 min each step; then a final elongation step of 5 min. PCR products were loaded on 2% agarose electrophoresis gels with loading buffer in 0.5x TBE buffer. Alleles were scored as co-dominant markers visually and bands of the same size were assumed to be identical. Multiple segregating loci detected with one SSR primer pair were indicated by addition of a suffix (a, b) to the locus names.

Linkage analysis and map construction was carried out using the program Joinmap 3.0 (Van Ooijen and Voorrips 2001). The initial step involved calculating the Logarithm of odds (LOD) scores and pairwise recombination frequencies between markers. The segregating markers were grouped at a wide range of LOD scores (4.0 to 7.0) to identify the linkage groups. The Kosambi mapping function was adopted for map distance calculation. Linkage maps were visualized using Mapchart (Voorrips 2002).

QTL analysis

The computer software MAPQTL 5.0 was employed to perform QTL analysis (Plant Research International, Wageningen University and Research Centre, Wageningen, The Netherlands) using both interval mapping (IM) and multiple-QTL model mapping (MQM) methods as described in its reference manual (http://www.kyazma.nl/manuals/MapQTL5Manual.pdf). The analysis started with the interval-mapping test to find putative QTL by applying the permutation test to each data set (1000 repetitions) to decide the LOD thresholds ($p = 0.05$). Markers located in the vicinity of QTL were selected as an initial set of cofactors. MQM analysis was then performed to precisely locate QTL after the automatic selection of cofactors. Only significant markers at $p < 0.02$ were used as cofactors in the multiple QTL detection. A map interval of 5cM was used for both IM and MQM analyses. LOD 2.9 for F2/3, 2.5 for DH38, DH30, DH-03 and DH-CC was used as a significance threshold for the presence of a candidate QTL. For each QTL, two-LOD support intervals were established as approximately 95% confidence intervals (van Ooijen 1992). Genetic maps were constructed using Mapchart software (Voorrips 2002). DH-38, DH-30 and DH-03 maps are described by Lou et al. (submitted) and the DH-CC map is described by Wu et al. (submitted).

Results

Construction of a genetic map of the RC-CC F2 population with AFLP and SSR markers

A set of 50 *EcoR* I/*Mse* I and 70 SSR primer combinations were tested on parents of the F2 population (RC-144 and CC-156) to evaluate their polymorphisms. A total of 28 pairs of *EcoR* I/*Mse* I and 16 SSR markers (Table 1) were selected and used for genotyping the mapping population resulting in 332 AFLP and 17 SSR markers.

Of the 332 AFLP fragments 18 bands (5.4%) showed a clearly alternating segregation in pairs of alleles, resulting in 9 bi-allelic co-dominant markers. The F2 linkage map was based on 256 AFLP (from which 7 were co-dominant) and 13 SSR markers, representing 11 linkage groups covering a total map length of 943.2 cM (Table 2; Fig. 1). All the markers were arranged into 11 linkage groups at a LOD value of 4 to 7, while the haploid chromosome number of *B. rapa* is ten. Using the SSR markers, eight of the 11 linkage groups could be assigned to R01, R02, R03, R05, R06, R07, R08 and R09 of the international reference *B. rapa* map (Kim et al. 2006; Suwabe et al. 2006) and DH-30 and DH-38 maps (Lou et al.

submitted). Three linkage groups LG1, LG2 and LG3 could not be assigned to R group, and may represent R04 and R10. Most of the linkage groups showed no apparent clustering of linked markers, with the exception of R09. The number of markers in each linkage group varied from 7 (R01) to 42 (R09), with an average interval size of 3.63 cM ranging from an interval size of 1.86 cM in R09 to an interval size of 11.70 cM in R01 (Table 2). Of 269 mapped markers, 87 (32.3 %) deviated ($P \leq 0.01$) from the expected 3:1 (dominant loci) or 1:2:1 (co-dominant loci) ratio showing distortion in the segregation values. Most markers with distorted segregation ratios mapped on R03, R05, R06, R08 and R09.

Table 1 Primer combinations applied on the RC-CC F2/F3 population

Types	Primer combinations			
EcoR I/*Mse* I	E31-AAA/M60-CTC	E32-AAC/M49-CAG	E32-AAC/M54-CCT	E32-AAC/M60-CTC
	E32-AAC/M61-CTG	E33-AAG/M47-CAA	E33-AAG/M48-CAC	E33-AAG/M50-CAT
	E33-AAG/M51-CCA	E33-AAG/M59-CTA	E34-AAT/M50-CAT	E35-ACA/M47-CAA
	E35-ACA/M62-CTT	E36-ACC/M47-CAA	E37-ACG/M59-CTA	E37-ACG/M60-CTC
	E38-ACT/M50-CAT	E38-ACT/M51-CCA	E38-ACT/M56-CGC	E38-ACT/M59-CTA
	E38-ACT/M62-CTT	E39-AGA/M47-CAA	E39-AGA/M51-CCA	E41-AGG/M50-CAT
	E41-AGG/M62-CTT	E42-AGT/M51-CCA	E44-ATC/M47-CAA	E44-ATC/M62-CTT
SSR	BRMS096R01	Ra2G09R01	BRMS037	FLC2R02
	Na12H09R02	BRMS042R03	BRMS043R03	BRMS054R04
	Ra3H10R05	BRMS014R06	Ra2A01R07	BRMS036R07
	Ra2E12R08	BRNS051R09	BRMS019R10	FLC1R10

Table 2 Characteristics of an F2 (CC-156 x RC-144) genetic map of *B. rapa*

Linkage group	No. of markers (AFLP+SSR)	Density (marker/cM)	Average interval (cM)	No. of distorted	Length (cM)
R01	7 (4+3)	0.09	11.70	1	81.9
R02	28 (27+1)	0.23	4.28	6	119.9
R03	30 (28+2)	0.34	2.91	12	87.4
R05	26 (25+1)	0.34	2.91	10	75.7
R06	32 (31+1)	0.29	3.42	12	109.5
R07	22 (20+2)	0.23	4.36	9	96
R08	24 (23+1)	0.24	4.10	11	98.4
R09	42 (40+2)	0.54	1.86	14	78.1
LG1	20 (20+0)	0.24	4.16	6	83.2
LG2	13 (13+0)	0.26	3.80	1	49.4
LG3	25 (25+0)	0.39	2.55	5	63.7
Sum/Mean	269 (256+13)	0.28	3.63	87	943.2

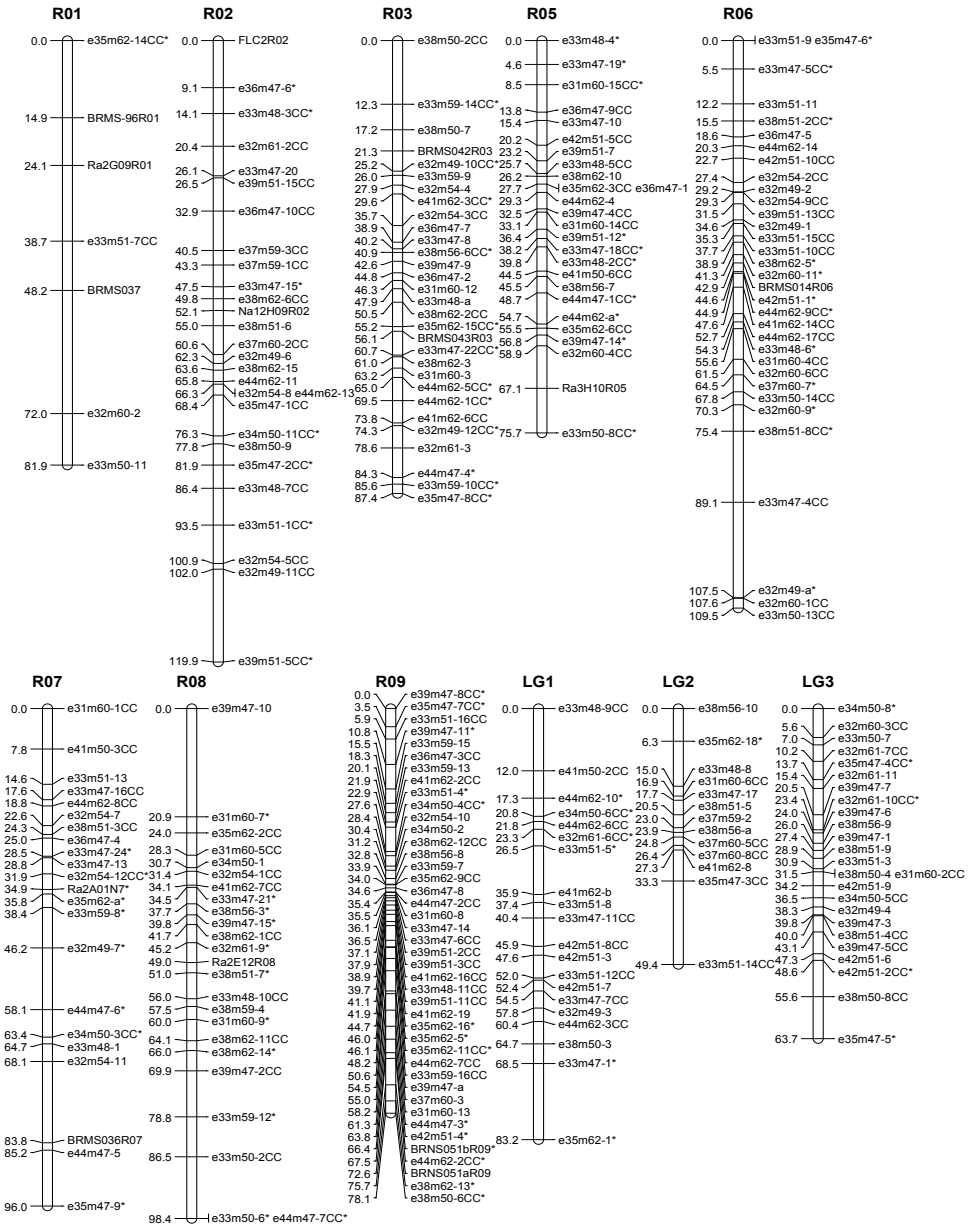

Fig. 1 A linkage map of *B. rapa* based on the F2 (CC-156 x RC-144) population with AFLP and anchor SSR markers.

Skewed marker loci are indicated with *, indicating a significant level at $P \leq 0.01$.

Variation in phytate and phosphate content

To identify the genetic loci responsible for the genetic variation in phytate and phosphate contents, the concentration of both compounds was determined in seeds and (or) leaves of individuals in all segregating populations. Phytate in seeds were much higher than in leaves, whereas phosphate levels were higher in leaves (Table 3) as commonly found when seeds are compared with leaves. The amount of phytate and phosphate in seeds was higher (~ 2.0 fold) in the RC-CC F2 and DH-CC plants but leaf phosphate level was lower (~ 2.7 fold) in the F2 population, compared to plants of DH-38, DH-30 and DH-03, which may be caused by the different growing conditions. Seeds of DH-38, DH-30 and DH-03 were harvested from greenhouse grown plants in the winter in Wageningen, whereas seeds of F2 plants and seeds of DH-CC were collected from greenhouse grown plants in the spring in Beijing. The F2/3 and DH-CC populations also showed a larger variation in seed phosphate (SPHO) (~ 10 fold) than the other DH populations (~ 4-6 fold) (Table 3). In DH-38, DH-30 and DH-03, the variation coefficients for different traits were similar (around 0.25), except for SPHO in DH-30 which had a higher variation coefficient (0.46), and leaf phytate (LPHY) in three DH populations (variation coefficient > 0.50). The DH-CC showed lower (0.18) variation coefficient of seed phytate (SPHY). However, the variation coefficient for SPHO was higher (0.57) than for all other populations.

Table 3 shows the correlation coefficients between phytate and phosphate levels in seeds and leaves. The correlation coefficient value between the two components in seeds and leaves was lower than that in the analysis of a collection of 160 accessions (see chapter 3 of this thesis), a significant positive correlation ($p < 0.05$) was only detected in DH-38 and the F2/3 populations. No significant correlation was observed between phytate in seeds and phosphate in leaves, which implies that the phytate level in seeds might not represent the overall higher phosphorus status in the plant.

For the DH-38, DH-30 and DH-03 populations, the frequency distributions of phytate and phosphate levels showed transgression in both directions except for the phytate levels in DH-03 (Fig. 2). This implied that the parental accessions YS-143 and PC-175, VT-115, MIZ-019 carry alleles that decrease levels at some loci but increase levels at other loci.

Table 3 Statistical analysis of phytate and phosphate levels in five populations

Population	Statistical Parameters	SPHY	SPHO	LPHY	LPHO	$R^{sphy/spho}$	$R^{lphy/lpho}$
F2/3	Range (mg/g)	34.8-108.8	0.7-8.2	-	2.11-18.71	0.22*	-
	Standard deviation	18.27	1.30	-	2.77		
	Mean (mg/g)	61.36	3.36	-	7.08		
	Variation coefficient	0.30	0.39	-	0.39		
DH-38	Range (mg/g)	16.2-62.4	0.6-2.6	0.2-1.8	7.9-26.9	0.27*	0.36**
	Standard deviation	11.33	0.39	0.36	3.76		
	Mean (mg/g)	32.95	1.22	0.56	15.97		
	Variation coefficient	0.34	0.32	0.64	0.24		
DH-30	Range (mg/g)	13.6-38.9	0.6-3.9	0.1-2.1	8.8-35.1	0.14	-0.01
	Standard deviation	6.51	0.81	0.55	6.01		
	Mean (mg/g)	26.31	1.78	0.85	22.94		
	Variation coefficient	0.25	0.46	0.65	0.26		
DH-03	Range (mg/g)	21.5-54.8	0.5-2.4	0.4-2.8	10.5-34.4	-0.18	0.14
	Standard deviation	8.88	0.40	0.75	4.77		
	Mean (mg/g)	33.29	1.35	1.39	19.70		
	Variation coefficient	0.27	0.29	0.53	0.24		
DH-CC	Range (mg/g)	34.2-102.9	0.9-9.7	-	-	0.15	-
	Standard deviation	11.11	1.49	-	-		
	Mean (mg/g)	61.54	2.62	-	-		
	Variation coefficient	0.18	0.57	-	-		

SPHY, seed phytate; SPHO, seed phosphate; LPHY, leaf phytate; LPHO, leaf phosphate; Rsphy/spho, Rlphy/lpho: Correlation coefficient between SPHY and SPHO, and between LPHY and LPHO; *: Significant level for correlation coefficient at $p < 0.05$* and $p < 0.01$**; -, not analyzed.

Identification of QTL for phytate and phosphate

To detect association between molecular markers and phytate and phosphate levels, QTL analysis was performed. Some loci significantly affecting phytate and phosphate content in seeds and leaves were identified in all mapping populations (Table 4). In total, 27 QTL for phytate and phosphate content in seeds and leaves were detected in five populations distributed over 8 linkage groups. A large percentage of phenotypic variation (38.6-72.1%) was explained by a LPHY QTL on the higher middle of R07, which was detected in DH-38, DH-30 and DH-03. For the other three traits (SPHY, SPHO and LPHO) in F2/3, DH-38 and DH-30 populations, the additive effects of QTL accounted for 40.8%, 53.6% and 75.0% of the variation for SPHY, 50.3%, 32.8% and 51.3% for SPHO, and 8.9%, 28.1% and 46.2% for LPHO. In DH-CC, three QTL affecting phytate and phosphate contents in seeds were detected, explaining only 12.7% of the variation for SPHY and 16.1% of the variation for SPHO.

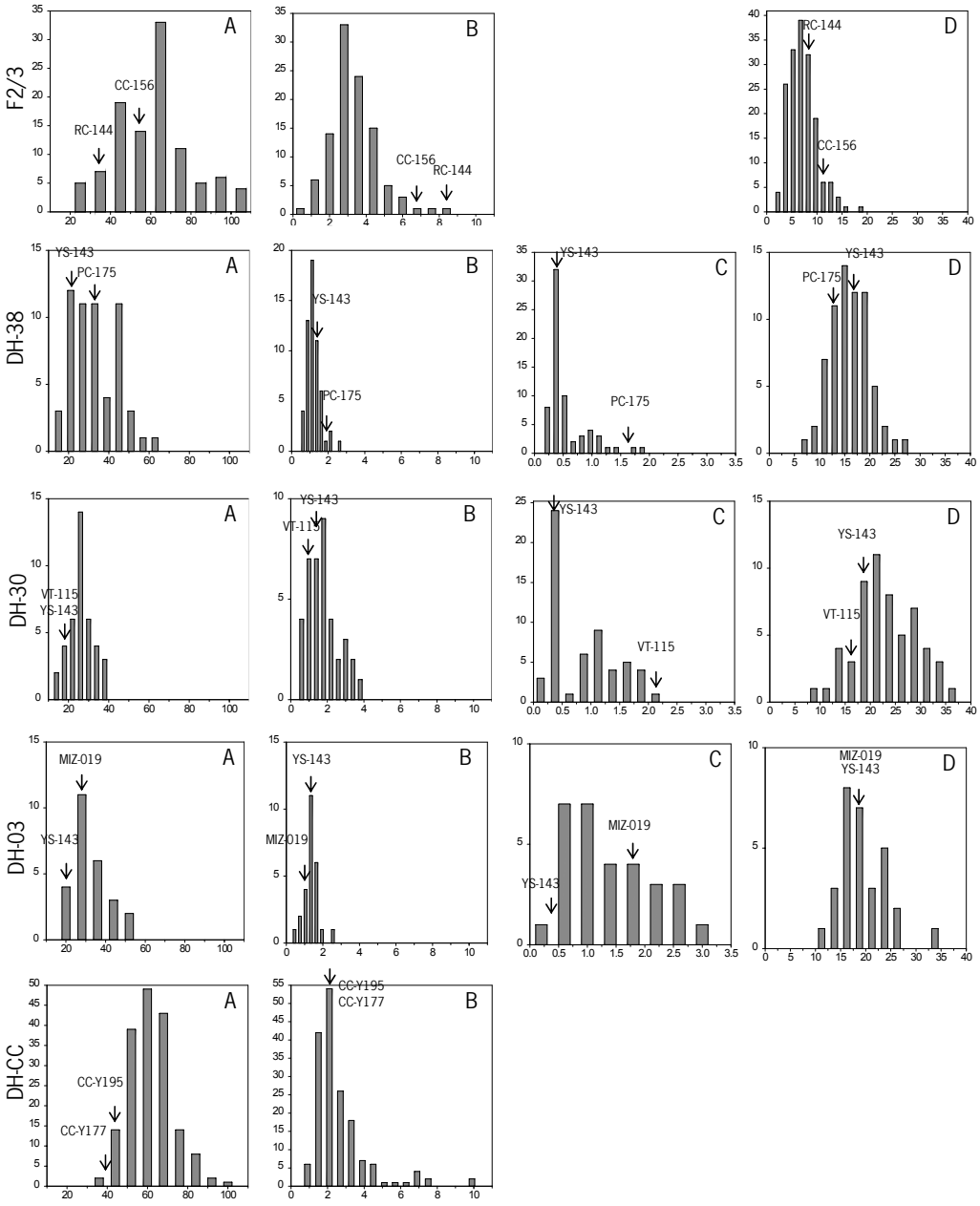

Fig. 2 Frequency distribution of the concentration of SPHY (A), SPHO (B), LPHY (C) and LPHO (D) of five segregating populations. Arrows indicate the levels in the parental lines. The horizontal axes indicate concentrition (mg/g); the vertical axes indicate number of genotypes.

Table 4 Observed QTL affecting phytate and phosphate levels in seeds and leaves in five populations

Population	Trait	Linkage group	Position (cM)	Nearest marker	LOD	Explained variance (%)	Effect
F2/F3	Sphy	R01	24.15	Ra2G09R01	3.92	16.1	-
		R05	7.64	E31M6015-CC	2.93	12.7	+
		R06	18.61	E36M47-5	3.06	12.0	+
	Spho	R03	68.31	E38M50-7	3.65	17.6	-
		R05	25.71	E33M48-5CC	3.80	13.1	-
		R09	35.43	E44M47-2CC	5.67	19.6	+
	Lpho	R03	24.18	E31M60-3	2.99	8.9	-
DH-38	Sphy	R01	31.95	E36M15M197.9Y	2.51	13.2	-
		R06	51.29	P23M48254.2y	3.35	18.1	-
		R06	86.20	P23M47254.2	4.03	22.3	+
	Spho	R01	31.95	E36M15M197.9Y	2.65	15.3	+
		R06	86.20	P23M47254.2	3.47	17.5	+
	Lphy	R07	26.32	E32M16409.0	12.7	59.7	-
	Lpho	R03	39.50	E32M16496.7	2.63	12.6	-
		R08	46.63	E44M20190.4	3.09	15.5	+
DH-30	Sphy	R01	31.69	E32M19378.2	3.14	22.7	+
		R02	71.13	E34M16194.7	3.09	20.4	-
		R06	17.32	E34M15446.7y	3.95	31.9	+
	Spho	R01	17.90	E34M15420.9Y	2.88	20.3	+
		R03	33.99	P23M48281.8	4.39	31.0	+
	Lphy	R07	36.31	BRMS018R07	12.6	72.1	-
	Lpho	R01	27.40	E34M16237.5Y	4.46	22.8	-
		R03	32.73	E46M16575.7	4.31	23.4	+
DH-03	Lphy	R07	24.85	E34M16221.9	3.17	38.6	-
DH-CC	Sphy	R06	77.75	E38M50-6	4.16	12.7	+
	Spho	R01	46.71	E36M31-7	3.00	8.8	-
		R02	83.78	E33M56-2	2.47	7.3	-

Positive (+) effect indicated that from one parent (RC-144 for F2/3, YS-143 for DH38, DH-30 and DH-03, Y-195 for DH-CC) alleles at that marker increase levels of this trait, negative (-) effect indicated that from another parent (CC-156 for F2/3, PC-175 for DH-38, VT-115 for DH-30, MIZ-19 for DH-03, Y-177 for DH-CC) alleles increase levels of this trait. Trait abbreviations are indicated in Table 3.

The locations of all significant QTL and their support intervals are indicated in Fig. 3, where the linkage maps were compared, based on the common SSR or AFLP markers. Some QTL affecting a same trait were detected in different populations in the same linkage group (R01, R03, R06 and R07). For SPHY, 3 QTL were identified in the higher middle of R01, and 4 QTL were identified in the middle of R06 in multiple populations, suggesting that these represent only two different loci. For LPHO, 2 QTL explaining 12.6% and 23.4% in DH-38 and DH-30, were found on linkage group R03, which could also represent the same gene. One locus on R07 in the DH-38, DH-30 and DH-03 populations explained 59.7%, 72.1% and 38.6% of the variation for the LPHY, which appears to be the major locus responsible for the difference in phytate content in leaves between YS-143, which hardly contains phytate in leaves and the vegetable *B. rapa* parents. For SPHO, 3 QTL were detected in linkage group R01 in F2/3, DH-30 and DH-CC, respectively. However, we could not confirm whether these QTL have identical position because of lack of common markers in the regions where these QTL were detected.

Three other SPHY QTL, one at the bottom of R02, another one at the top of R05, and last one at the bottom of R06, were only detected in a single population and explained 12.7-22.3% of the variation. Additional QTL were detected for SPHO, one on R01 in DH-30 and one on R08 in DH-38, which explained 22.8% and 15.5% of the variation observed.

Since there is a low positive correlation between phytate and phosphate levels in leaves and seeds (Table 4) it was interesting to investigate whether QTL affecting both traits could be detected. Possible co-locations of QTL for SPHY, SPHO and LPHO in R01, SPHO and LPHO in R03, SPHY and SPHO in R05 and R06 were observed in different populations. No genomic regions affecting all 4 traits simultaneously were detected in a single population. Some QTL, where the different parental alleles had either both positive or opposite effects on the trait were detected for SPHY and SPHO. For example, the SPHY and SPHO QTL on R06 in DH-38 and on R01 in DH-30 co-localized, where the YS-143 allele increases the content of both phytate and phosphate in seeds. In the DH-38 population a QTL affecting both SPHY and SPHO was detected on R01. Here the YS-143 allele decreased SPHY and increased SPHO. The YS-143 allele for the LPHY QTL (on R07) decreases the content in all DH populations (DH-38, DH-30 and DH-03) where it was detected. However the YS-143 allele effects for LPHO (QTL on R01, R03 and R08) were either positive or negative explaining the transgressive variation.

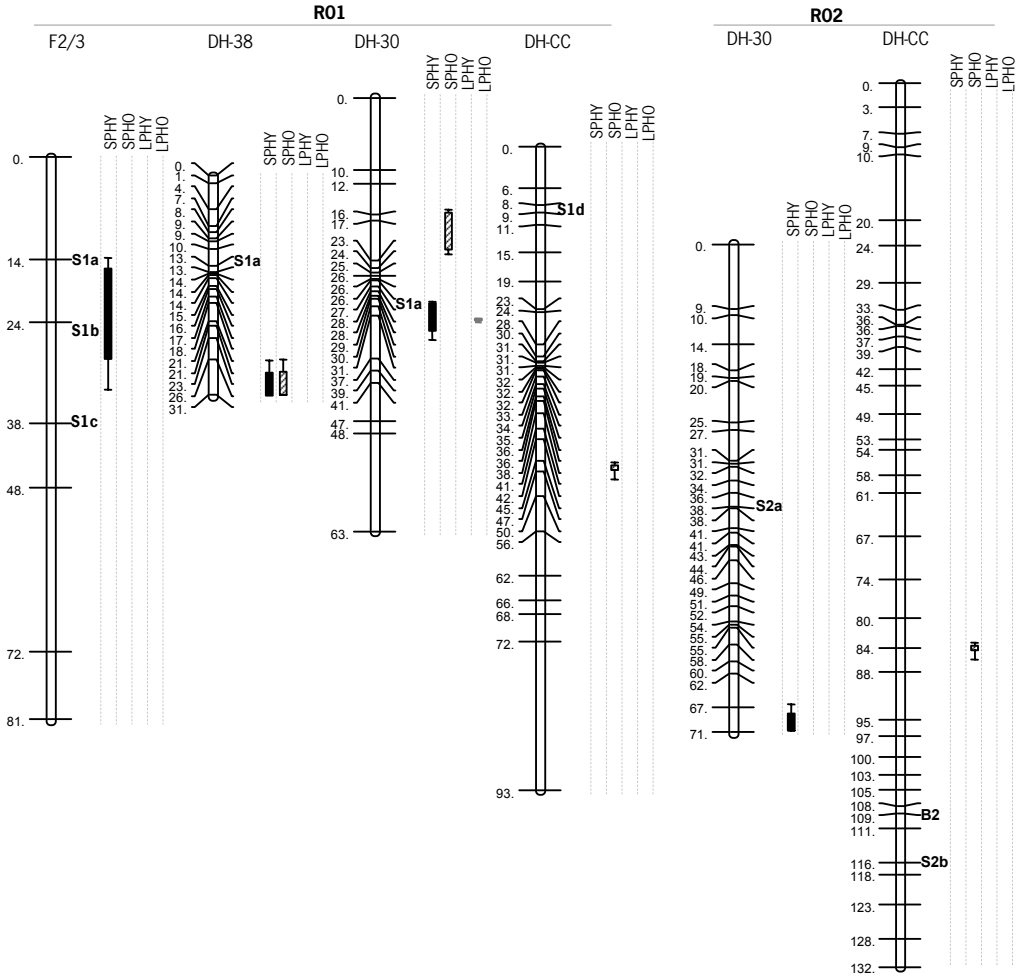

Fig. 3 The genetic locations of QTL (different boxes) affecting SPHY, SPHO, LPHY and LPHO levels indicated above each column in five mapping populations.

Boxes and whiskers represent 1-LOD and 2-LOD confidence intervals (95%) respectively for significant QTL. Linkage groups designations followed the international R group of *B. rapa* (Kim et al. 2006; Suwabe et al. 2006). Markers of DHs map are the same as described in Lou et al (submitted) and Wu et al. (submitted). Position of the same linkage groups in different populations is compared based on common AFLP and SSR (B1-B10; S1-S10) markers. Centimorgan (cM) position is indicated to the left of each linkage group. Trait abbreviations are indicated in Table 3.

S1a, BRMS096R01; S1b, Ra2G09; S1c, BRMS037; S1d, BRMS056; S2a, Na12H09; S2b, BrMAF-2; B2, BC-48; S3a, BRMS043; S3b, BRMS042; S5, Ra3H10; S6a, BRMS014; S6b, Na12H07; B6-BC51; S7a, BRMS018; S7b, Ol12E03; S7c, Ra2A01; S8, Ra2E12; S9, BRMS051.

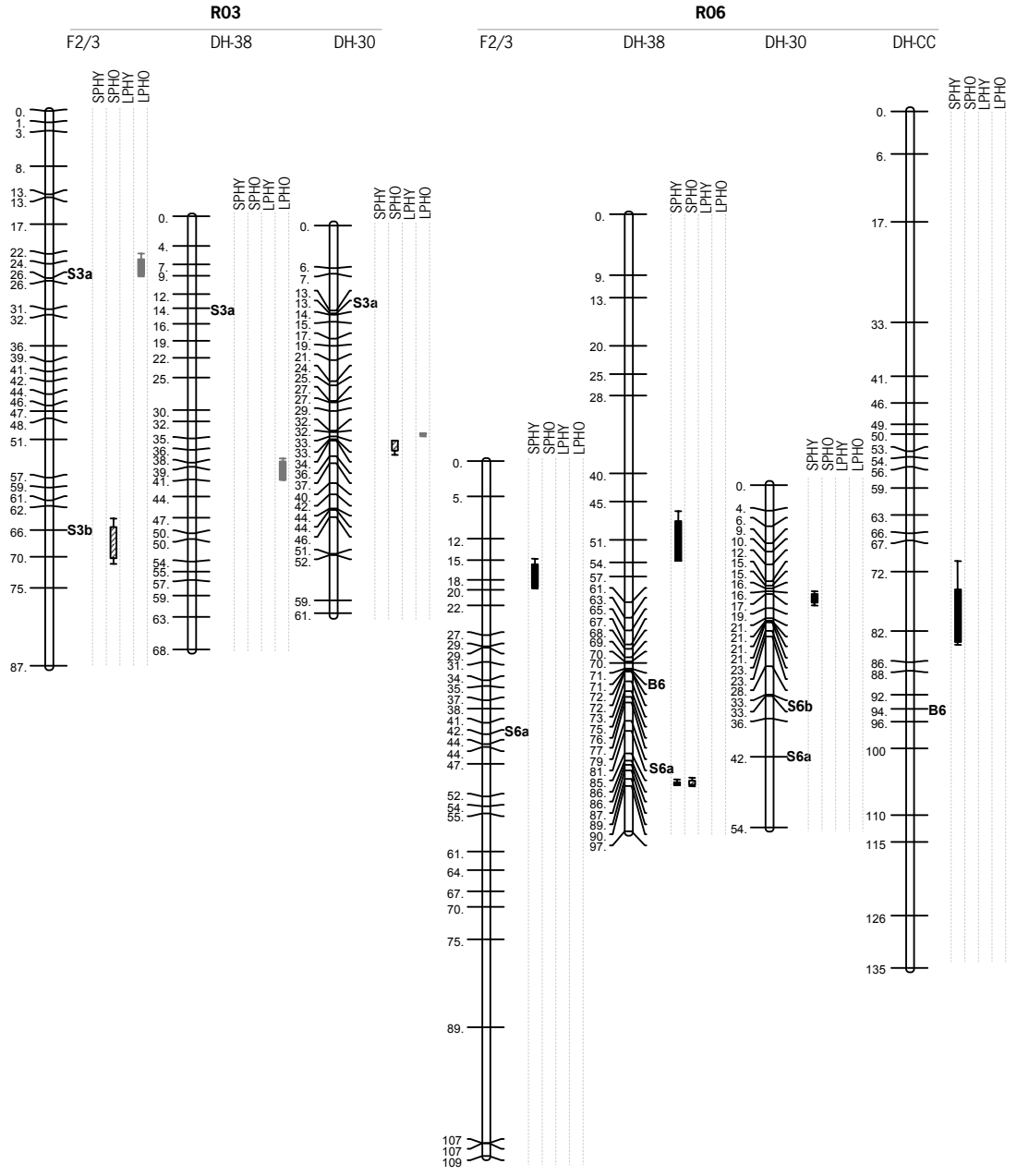

Fig. 3 continued

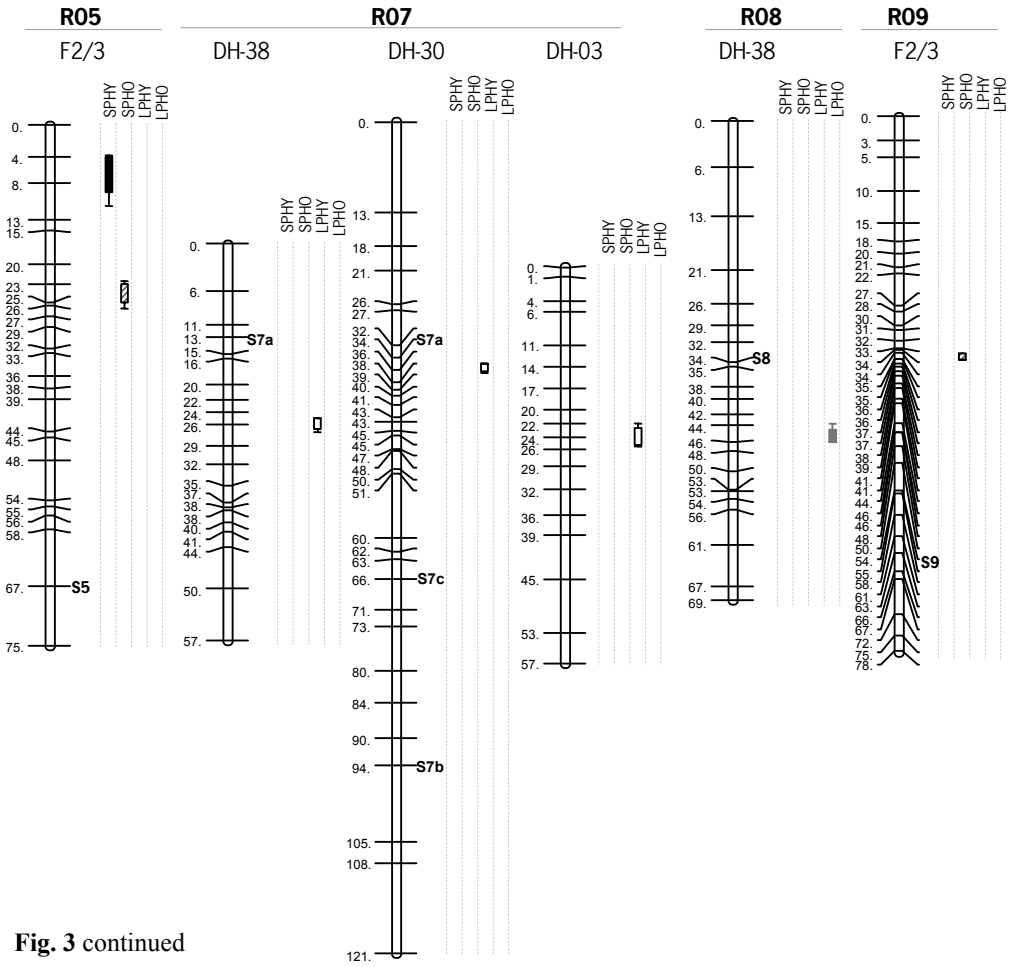

Fig. 3 continued

Only two possible co-locations of QTL for seed phytate/phosphate and leaf phytate/phosphate were detected, corresponding to the absence of significant correlations between seeds and leaves. For this co-location, the different parental alleles had either both positive or opposite effects on the trait. The SPHO and LPHO QTL on R03 in DH-30 co-localized, and the YS-143 allele increases the phosphate content both in seeds and in leaves. The phenotypic effects of SPHY/SPHO and LPHO QTL on R01 in DH-30 were different: YS-143 alleles for SPHY/SPHO QTL decreased the content, whereas YS-143 alleles for the LPHO QTL increased the content.

Discussion

In *B. rapa*, SSR markers have been used to construct linkage maps and contributed to assignment of linkage groups (R groups) (Parkin et al. 2005; Kim et al. 2006; Suwabe et al. 2006). The linkage maps of 4 segregating populations: DH-38, DH-30, DH-03 (Lou et al. submitted) and DH –CC (Wu et al. submitted) were based on AFLP and SSR markers. A new linkage map of the F2/3 population, derived from a cross between a Rapid cycling line RC-144 and a vegetable type Chinese cabbage line CC-156, includes 11 linkage groups covering a total map length of 943.2 cM. Unfortunately SSRs could not be mapped onto 3 linkage groups, so that those could not be assigned to R groups. However, the other 8 linkage groups with SSR markers allowed the map comparison with the DH populations.

The genetic regulation of phytate and phosphate levels in seeds and leaves was studied in these 5 segregating populations involving 8 parental accessions (CC-156, RC-144, YS-143, PC-175, VT-115, MIZ-019, CC-Y195 and CC-Y177). In *B. napus* (Lickfett et al. 1999), it was shown that phytate and phosphate concentrations in seeds are affected by growth conditions. Such environmental effects might explain why the levels of SPHY and SPHO in F2/3 and DH-CC, grown in a greenhouse in Beijing were higher than these levels in DH-38, DH-30 and DH-03, grown in a greenhouse in Wageningen. In the F2/3 population, the level of LPHO was much lower than the level in DHs, and LPHY levels were below detection level. The variation between parental accessions was generally small. However, considerable variation and transgression was observed in most populations, also in the DH-CC population, which is derived from a cross within the Chinese cabbage cultivar group. This indicated that the parental accessions of the used populations carry alleles that both decrease and increase levels at several loci. The genetic analysis for the traits showed that the direction of allelic effects could indeed explain the transgression observed in most cases, indicating that both parents of the populations have QTL alleles with positive and negative effects. For SPHO QTL in DH-38, DH-30 and DH-CC, and LPHO in F2/3, parents carried only one directional QTL, while the segregation for SPHO and LPHO in these populations was transgressive. Possibly additional QTL with opposite phenotypic effects escaped detection in this study.

Totally 27 QTL for leaf and seed phytate and phosphate levels, probably representing 10 different loci, were identified in the different populations, most of them located on R01, R03, R06 and R07. Several QTL affecting the same trait co-localized in different populations. Two

SPHY QTL on the middle of R01 and R06 were detected in multiple populations, explaining 12.0%-31.9% of the phenotypic variations indicating that these are possibly the two major QTL distinguishing SPHY accumulation between the oil type parents and the vegetable parents. In agreement with this the R01 SPHY QTL was not detected in DH-CC. One QTL affecting LPHO level was located on a same position in linkage group R03 in DH-38 and DH-30, and one major QTL affecting LPHY level at the same position in linkage group R07 in DH-38, DH-30 and DH-03 probably represents the same locus. Some QTL affecting a particular trait were only found in a single population, suggesting differences between populations (parents) or genotype x environment effects on phytate and phosphate accumulation in seeds or leaves. The use of other accessions may identify additional loci affecting this trait.

Although genome regions affecting all 4 traits were not detected, possible co-locations of QTL for 2 or 3 traits were found in R01, R03, R05 and R06 in different populations. Co-localization of these QTL suggests that single loci are involved in the accumulation of phytate or phosphate in seeds or leaves. Such loci could control overall phosphorus levels in the plant or specifically in the different organs. The weak but positive correlation between phytate and phosphate within seeds and within leaves is in agreement with co-locations of QTL affecting these traits. Four QTL for SPHY/SPHO (on R01, R05 and R06) possibly co-located with each other within a specific population. Two of them (on the bottom of R01 and R06) only co-localized in DH 38, in which population the only strong correlation between the two traits was detected. For 2 SPHY/SPHO QTL the allelic effects are in the same direction. This was also found for the major QTL on the top of chromosome 3 in Arabidopsis that controlled both the levels of phytate and phosphate in seeds and leaves. It was concluded that this QTL is a phosphate accumulation QTL (Bentsink et al. 2003). For co-locating QTL with an apparent antagonistic effect (like QTL for SPHY/SPHO on R01 in DH-38) the biosynthesis of phytate may be altered, thus altering the ratio of phytate compared to the phosphorus level, as has been described in mutants of maize, barley and rice (Raboy et al. 2001). It must be emphasized that co-location of QTL may indicate that a single gene underlies the QTL, or implies that different but closely linked genes are involved. In this paper comparisons of QTL positions across populations are also complicated by in inaccuracies of QTL mapping. In DH-38, DH-30 and DH-03, we could not detect a LPHO-QTL that co-localized with the major

LPHY-QTL on R07. This is possibly due to the low LPHY content in leaves, which implies that changes in phytate content hardly affect phosphate levels.

Based on homologous SSR loci, Suwabe et al. (2006) analyzed the synteny between *B. rapa* and Arabidopsis. For QTL identified in this study on R01, R03, R05, R06 and R07 of *B. rapa*, syntenic regions in each chromosome of Arabidopsis could be identified. Based on the comparative mapping between *B. napus* (the A genome component of *B. napus* N1-N10) and Arabidopsis (Parkin et al. 2005), Schranz et al. (2006) summarized the organization of Arabidopsis genomic blocks that make up the A genome in *B. rapa*. Some of the anchored SSRs on the SSR-based map (Suwabe et al. 2006) were also mapped in a sequenced-tagged map (Kim et al. 2006) and in this study, which allows the comparison of the maps presented in this study to the genomic blocks as defined by Schranz et al. (2006) and Parkin at al. (2005). This comparison makes it possible to directly compare the location of *B. rapa* phytate and phosphate QTL to those QTL identified in Arabidopsis. In Arabidopsis, 5 QTL affecting phytate and phosphate levels in seeds and leaves were detected, one major QTL being located on the top of chromosome 3 and additional QTL being located on chromosomes 1, 2 and 4 (Bentsink et al. 2003). In Fig. 4, synteny between Arabidopsis and *B. rapa* genomic blocks is depicted only for those syntenic blocks with QTL in both Arabidopsis and *B. rapa*. The QTL on R03 and R05 could be related to the major genomic region affecting phytate and phosphate on the top of chromosome 3 in Arabidopsis. The QTL on R01 and R07 are also possibly related to the QTL on chromosome 4 and 1 of Arabidopsis, respectively. However, we cannot reliably compare the SPHY/SPHO QTL on R06 of *B. rapa* with SPHO QTL on the top of chromosome 1 in Arabidopsis because of lack of common SSR markers in these regions. Adding additional SSR and gene targeted markers to the *B. rapa* linkage groups will improve the accuracy of identification of syntenic Arabidopsis-*B. rapa* QTL. For other QTL identified in this study and in Arabidopsis, we did not identify QTL in syntenic regions.

Our results provide evidence for a genetic regulation of phytate and phosphate levels in seeds and leaves of *B. rapa*, and a preliminary genomic comparison with QTL identified in Arabidopsis at syntenic positions. The used populations will be further applied to perform genetic analysis for morphological traits and are suitable to study the link between phytate levels and plant vigour and seeds traits.

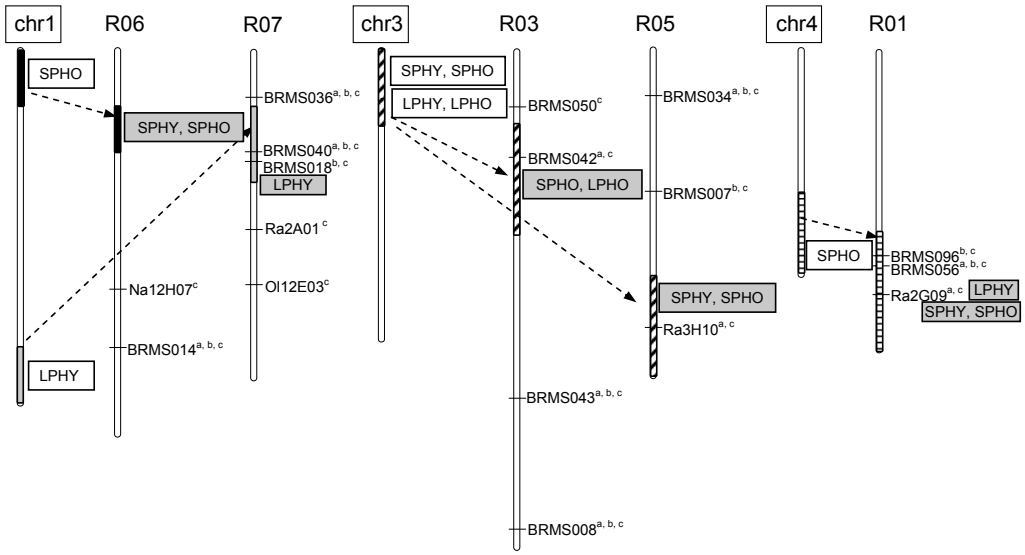

Fig. 4 A comparative *B. rapa-Arabidopsis thaliana* map with phytate and phosphate QTL.
B. rapa linkage groups R03, R05, R06 and R07 are presented with QTL positions (gray boxes) as identified in this study and positions of SSRs (a, mapped in Kim et al. 2006; b, mapped in Suwabe et al. 2006; c, mapped in this study) used for map comparison. The white boxes represent the QTL identified in Arabidopsis chromosomes (chr) 1, 3 and 4, which have been described by Bentsink et al. (2003). Synteny between Arabidopsis and *B. rapa* genomic blocks is indicated with similar patterns and shading (information from Parkin et al. (2005), Suwabe et al. (2006) and Schranz et al. (2006)); only those syntenic blocks with QTL in both Arabidopsis and *B. rapa* are depicted. Trait abbreviations are indicated in Table 3.

Acknowledgements

We thank Prof. Zhang Xiaowei (Henan Academy of Agricultural Sciences, China) for the development of one Double Haploid population (DH-CC). We thank Jiangling Xiong for testing of some SSR markers. We are grateful to staff in Plant Science Experimental Centre of Wageningen University, the Netherlands, and Yanguo Zhang and Yanling Liu in the Institute of Vegetables and Flowers (Beijing), Chinese Academy of Agriculture Science, China for taking care of the plants and assistances. The project is sponsored by the Royal Dutch Academy of Sciences (KNAW), the Asian Facility (project AF01/CH/8 "Sino-Dutch

Genomic Lab and Vegetable Research Center") and a fellowship from an exchange student programme (Huygens) between Chinese Scholarship Council and Nuffic in the Netherlands.

Chapter 5

Mapping quantitative trait loci for morphological traits in multiple populations of *Brassica rapa*

Ping Lou*, Jianjun Zhao*, Xiaofei Song, Dunia Pino Del Carpio, Shuxing Shen, Dick Vreugdenhil, Xiaowu Wang, Maarten Koornneef, Guusje Bonnema

Abstract

Wide variation for morphological traits exists in *Brassica rapa* and the genetic basis of this morphological variation is largely unknown. Here we report on quantitative trait loci (QTL) analysis of flowering time, seed and pod traits, growth-related traits, leaf traits, and turnip traits in *B. rapa* using multiple populations. The populations resulted from wide crosses between parental accessions: Rapid cycling, Chinese cabbage, Yellow sarson, Pak choi and Vegetable turnip. A total of 27 QTL affecting 20 morphological traits were detected, 7 QTL for flowering time, 6 for seed traits, 3 for growth-related traits, 10 for leaf traits and one major QTL for turnip formation. Principal component analysis and co-localization of QTL indicated that some components of the genetic control of leaf and seed-related traits and of flowering time and turnip formation might be the same. One major QTL, controlling turnip formation, was mapped on the top of R02 and co-localized with the major flowering time QTL. The possible gene(s) underlying this QTL and comparative analyses between the four populations and with *Arabidopsis thaliana* are discussed.

*These authors contributed equally to this work.

Introduction

Brassica rapa is an important species of the genus *Brassica*, which provides both rapeseed oil, fodder and vegetables contributing to the world economy and to the health of people as a source of beneficial nutrients. During the long history of breeding and selection, a variety of forms have been selected for use as oilseeds, leafy vegetables and turnips (chapter 1, this thesis).

Although *B. rapa* breeding is ongoing for a long time, very limited information is available on the inheritance of morphological traits in this species. To date, a number of genetic studies aimed at the identification of loci controlling morphological variation have been conducted, illustrating the complex genetic control of the many quantitatively inherited traits. In a study by Song et al. (1995) 28 phenotypic traits including leaf, stem and flowering characteristics were analysed and 0-5 QTL were detected for each of the traits in an F2 population (derived from a cross Chinese cabbage X Spring broccoli). Using the Chinese cabbage DH population (Chiifu X Kenshin), QTL with significant effects on head weight, leaf blade width, head compactness and head length were mapped on 3 linkage groups (Choi et al. 2004). A number of studies have been conducted to study the genetics underlying the specific morphology of the curd in *Brassica oleracea* (Lan and Paterson 2000; Sebastian et al. 2002), which acts as an example to understand the genetic basis for morphological and developmental traits in other *Brassica* species. In recent studies on Arabidopsis (reviewed by Koornneef et al. 2004), QTL for floral, leaf morphology and other growth-related traits were mapped and presented clear evidence for a modular genetic architecture where similar loci control a number of related processes. Modularity is that the quantitative expression of particular traits tend to vary in a coordinated and structured manner. One of the mechanisms that can cause modularity is genetic correlation among traits, due to pleiotropy or extremely tight linkage (Conner 2002). Linking the studies of Arabidopsis with *Brassica* is well feasible nowadays because the syntenic relationships are better established (Parkin et al. 2005).

Among the agronomic traits, flowering time is one of the most important traits and wide variation exists among *B. rapa*. It is affected by the growing season (day length, temperature) and thus varieties are bred for specific geographical regions and seasons (spring/autumn). In the Brassicaceae family (Cruciferae), many studies have examined QTL affecting flowering time in different environments using different populations. In Arabidopsis, the largest

difference in flowering time among ecotypes appears to be due to allelic variation at the *FLC* (*Flowering locus C*, a MADS box transcription factor) and loci, such as *FRI* (*FRIGDA*), that interacts with *FLC* (Koornneef et al. 2004). Using mutant approaches several genes such as *CO (CONSTANS)*, *EMF (EMBRYONIC FLOWER)*, *FY* and *FLC* that contribute to flowering time were mapped on the top of Arabidopsis chromosome 5. In *B. oleracea*, Kennard et al. (1994) found two regions, representing genome duplications, each containing QTL for flowering time, stem and leaf traits. Bohuon et al. (1998) described three regions containing QTL for flowering time and additional QTL for flowering time were revealed by using substitution lines (Rae et al. 1999). In *B. nigra,* Lagercrantz's group observed that a genomic region, which is co-linear with the top of chromosome 5 of Arabidopsis, was associated with flowering time variation and suggested *CO* as a likely candidate gene for this flowering time QTL. Furthermore, they compared the genetics of flowering time in four *Brassica* genomes and again concluded that *CO*, and not *FLC*, duplicated copies were likely candidates for flowering time QTL (Lagercrantz et al. 2002; Kruskopf Österberg et al. 2002).

In *B. rapa*, several QTL (*VFR1*, *VFR2* and *VFR3*; *FR1*, *FR2* and *FR3*) for flowering time were identified in an F2 population (Teutonico and Osborn 1995) from a cross between an annual (yellow sarson) and a biennial oil type and in a recombinant inbred population (Osborn et al. 1997). *VFR2* was estimated to have a large effect and appeared to be homologous to *FLC* of Arabidopsis (Kole et al. 2001). A further study indicated that *VFR2* corresponded to *BrFLC1*, *FR1* corresponded to *BrFLC2*, *FR2* corresponded to *BrFLC5,* and *VFR1* was mapped on R02 close to *MAF* (MADS *Affecting Flowering*) region (Schranz et al. 2002). These three *B. rapa* flowering time genes *BrFLC2, BrFLC3* and *BrFLC1* were assigned to linkage group R2, R3 and R10 respectively (Schranz et al. 2002; Kim et al. 2006).

All the research described above used oil-type *B. rapa* for mapping flowering time genes and it will be interesting to know what is the genetic variation for flowering time in the other *B rapa* types. Bolting time was also analyzed under different conditions in a population derived from a cross between two heading Chinese cabbages, and 10 QTL located on 6 linkage groups were identified (Ajisaka et al. 2001; Nishioka et al. 2005). However, these linkage groups were not assigned to the reference linkage groups so it is not possible to compare these QTL to other flowering time QTL. In general it seems that the multiple copies of *Brassica* genes homologous to flowering time genes especially on the top of chromosome 5 of Arabidopsis,

such as *FLC* and *CO*, contribute to the wide variation in flowering time in the genus *Brassica*. In a previous study the relationship between accessions was revealed by AFLP fingerprinting in a large collection of *B. rapa* (Zhao et al. 2005). A finding was that genetic distance was more related to geographical origin (East Asia vs. Europe) than to the different morphotypes. This prompted us to further investigate the genetic relationships by crossing genotypes with different morphotypes and geographical origins. In this study a number of segregating populations with parents selected from the three main groups (oil-, leafy- and turnip types) that are distinguished in *B. rapa*, are used to genetically dissect plant morphology. Yellow sarson, an Indian oil-type, is characterized by its early flowering, self-compatibility and yellow seed coat. Within the leafy types, there are two important Chinese subgroups, Chinese cabbage and Pak choi, that differ in leaf characteristics like leaf surface, color and shape, flowering time and heading form (in case of Chinese cabbages). Turnips were mainly produced in European countries and are characterized by swollen taproots that are used for human and animal consumption. In addition, another 'morphologically simple' rapid cycling genotype (Williams and Hill 1986) with a very short life cycle, selected by accumulating QTL for earlier flowering, was included as parent. Five distant parental lines were crossed and resulted in 4 populations (1 F2/3, 2 DHs (Double Haploid) and 1 BC1 (backcross)) that are used to study the genetics of several morphological traits.

Materials and Methods

Plant materials

Four populations were developed from wide crosses between *B. rapa* accessions. The parental accessions were selected based on their origins, morphological types and their AFLP patterns, which are described in previous study (Zhao et al. 2005).

One F2/3 (RC-CC) population was produced from selfing of a single F1 plant, resulting from a cross between a Rapid cycling line RC-144 (accession number: FIL501) and a vegetable type Chinese cabbage line CC-156 (cultivar: Huang Yang Bai; accession number: VO2A0030) was analysed. The F2 and F3, obtained by selfing individual F2 plants (bud-pollination was used for self incompatible F2 plants), were used to evaluate flowering time, plant height, leaf traits and seed weight in three experiments. The first experiment was carried out during the spring (January to April) of 2004 in the Institute of Vegetables and Flowers (Haidian district of Beijing), Chinese Academy of Agriculture Science, where 178 plants were

grown in soil in the greenhouse. The second experiment was carried out during the spring (January to April) of 2005 in the Institute of Vegetables and Flowers (Beijing), where F3 seeds (125 lines, 10 plants per line) were grown in soil in the greenhouse at the Nankou farm, Changping district of Beijing. The third experiment was carried out during the winter (September to December) of 2005 in Wageningen University, where F3 seed (115 lines, 10 plants per line) were grown in pots in the greenhouse.

Two double haploid (DH) populations were developed from crosses between the oil type Yellow sarson YS-143 (accession number: FIL500) and the vegetable types: Pak choi PC-175 (cultivar: Nai Bai Cai; accession number: VO2B0226) and Vegetable turnip VT-115 (cultivar: Kairyou Hakata; accession number: CGN15199). A total of 135 lines including 71 lines from population DH-38 (PC-175 X YS-143), 64 lines from population DH-30 (VT-115 X YS-143) were analyzed for flowering time, leaf traits, seed colour and seed pod traits. DH-30 was also used to evaluate turnip formation. Three sowings were made per DH population. One set of DH lines (5 plants per line) were grown in pots in the greenhouse during the winter (September to December) of 2004 and in the spring (March to May) of 2005 at Wageningen University. Another set of DH lines (5 plants per line per replication) with 2 replications were grown in the open field during the autumn (July to October) of 2005 at Wageningen University.

An additional backcross (BC1) population of 136 pants, ((VT-115 X YS-143) X VT-115), was developed from a cross between one F1 plant (VT-115 X YS-143) and one plant of parental accession VT-115. Flowering time and turnip formation were analyzed in this population. The individuals were grown in the open field during the autumn (July to October) of 2005 at Wageningen University.

Trait measurement

In total, 22 traits related to flowering, seed, growth (plant height and branches number), leaf and turnip formation were recorded in 1 - 4 populations. The traits, their description and scale, years, locations and populations of the trials are shown in Table 1.

The leaf characteristics were scored on a fully developed leaf before flowering stage at a fixed date and subdivided in lamina length (LL), lamina width (LW), petiole length (PL) and leaf edge shape (LES) as illustrated in Fig. 1B. The values of leaf area (LA), LL, LW and PL in DH-38 and DH-30 were obtained by analyzing the leaf photographs using Scion Image (Scion

Corparation, MD, USA, http://rsb.info.nih.gov/nih-image), where the leaf photographs were digitally processed with the Irfanview program (http://www.irfanview.com). The values of LL and LW of F3 plants were measured by using a ruler. The mature and dried seedpod traits were measured once on harvested siliques in the greenhouse experiment of spring 2005. Seedpod characteristics are shown in Fig. 1A.

Data analysis

Statistical analysis for distribution and correlation were performed in Genstat 8.1. We also conducted a principle component analysis (PCA) in Genstat 8.1 on the line means for the flower, seed, leaf and turnip related traits to evaluate the correlations between the various traits.

Genotyping and map construction

Linkage analysis and map construction for F2/3 (chapter 4), DH-38 and DH-30 (Lou et al. submitted) and BC1 were carried out using the program Joinmap 3.0 (Van Ooijen and Voorrips 2001).

QTL analysis

The computer software MAPQTL 5.0 was employed to perform QTL analysis (Plant Research International, Wageningen University and Research Centre, Wageningen, The Netherlands) using both interval mapping (IM) and multiple-QTL model mapping (MQM) methods as described in its reference manual (http://www.kyazma.nl/manuals/MapQTL5Manual.pdf). The analysis started with the interval-mapping test to find putative QTL. Markers located in the vicinity of QTL were selected as an initial set of cofactors. MQM analysis was then performed to precisely locate QTL after the automatic selection of cofactors. Only significant markers at $p < 0.02$ were used as cofactors in the multiple QTL detection. A map interval of 5 cM was used for both IM and MQM analyses. A permutation test was applied to each data set (1000 repetitions) to decide the LOD (Logarithm of odds) thresholds ($p = 0.05$). LOD values of 2.9 for F2/3, 2.0 for DH-38, DH-30, DH-03 and BC1 was used as a significance threshold for the presence of a candidate QTL. For each QTL, two-LOD support intervals were established as approximately 95% confidence intervals (Van Ooijen 1992). Graphics were produced by Mapchart software (Voorrips 2002).

Table 1 List of traits analyzed

Trait type	Trait code	Trait name	Trait description[1]	Scale	Time of trial	Location[2]	Population of trial
Flowering trait	FL04sp	Flowering time 04sp	Days to flowering from sowing to appearance of the first open flower	days	Spring of 2004	GH/IVFa	F2
	FL04wi	Flowering time 04wi	Days to flowering from sowing to appearance of the first open flower	days	Winter of 2004	GH/WU	DH-38, DH-30
	FL05sp	Flowering time 05sp	Scored as 1 -early, 2 -middle, 3-late	1-3	Spring of 2005	GH/IVFb	F3
	FL05sp	Flowering time 05sp	Days to flowering from sowing to appearance of the first open flower	days	Spring of 2005	GH/WU	DH-38, DH-30
	FL05wi	Flowering time 05wi	Days to flowering from sowing to appearance of the first open flower	days	Winter of 2005	GH/WU	F3
	FL05au	Flowering time 05au	Days to flowering from sowing to appearance of the first open flower	days	Autumn of 2005	OF/WU	DH-38, DH-30, BC1
Seed-related trait	SPL	Seed pod length	Length between pedicel of silique and top of beak (Fig. 1A), measured by vernier caliper	mm	Spring of 2005	GH/WU	DH-38, DH-30
	SPW	Seed pod width	Width at the lengthwise midpoint of each silique (Fig. 1A), measured by vernier caliper	mm	Spring of 2005	GH/WU	DH-38, DH-30
	SBL	Seed pod beak length	Length between top of silique and top of beak (Fig. 1A), measured by vernier caliper	mm	Spring of 2005	GH/WU	DH-38, DH-30
	SC	Seed colour	Scored as 1-yellow, 2-yellow brown, 3-light brown, 4-brown, 5-dark brown	1-5	Autumn of 2004	GH/WU	DH-38, DH-30
	SW04sp	Seed weight 04sp	The mean seed weight per seed, obtained by weighting 2 to 5 seed lots each of 20 seeds	mg	Spring of 2004	GH/IVFa	F2
	SW05sp	Seed weight 05sp	The mean seed weight per seed, obtained by weighting 2 to 5 seed lots each of 20 seeds	mg	Spring of 2005	GH/IVFb	F3
Growth-related trait	PH	Plant height	Height from ground to the apical point of plant at flowering stage	cm	Spring of 2005	GH/IVFb	F3
	PB	Branches number	The number of main branches	number	Autumn of 2005	OP/WU	DH-30, BC1
Leaf trait	LES04sp	Leaf edge shape 04sp	Scored as 1-entire, 2-slightly serrated, 3-intermediate serrated, 4-much serrated	1-4	Spring of 2004	GH/IVFa	F2
	LES05wi	Leaf edge shape 05wi	Scored as 1-entire, 2-slightly serrated, 3-intermediate serrated, 4-much serrated	1-4	Winter of 2005	GH/WU	F3
	LT04sp	Leaf trichomes 04sp	Hair on leaf surface scored as 0-hair absent, 1-hair present before flowering	0-1	Spring of 2004	GH/IVFa	F2
	LT05sp	Leaf trichomes 05sp	Hair on leaf surface scored as 0-hair absent, 1-hair present before flowering	0-1	Spring of 2005	GH/IVFb	F3
	LN04sp	Leaf number 04sp	Number of leaves before flowering	number	Spring of 2004	GH/IVFa	F2
	LN05wi	Leaf number 05wi	Number of leaves before flowering	number	Winter of 2005	GH/WU	F3
	LB05wi	leaf lobes 05wi	Scored as 0-absent, 1-present	0-1	Winter of 2005	GH/WU	F3
	LB05au	leaf lobes 05au	Number of lobes	number	Autumn of 2005	OF/WU	DH-38, DH-30
	LL	Lamina Length (LL)	From base of petiol to tip of lamina (Fig. 1B)	cm	Spring of 2005	GH/IVFb/WU	F3, DH-38, DH-30
	LW	Lamina width (LW)	Lamina width at the widest point (Fig. 1B)	cm	Spring of 2005	GH/IVFb/WU	F3, DH-38, DH-30
	PL	Petiole length (PL)	From base of petiol to bottom of lamina (Fig. 1B)	cm	Spring of 2005	GH/WU	DH-38, DH-30
	LA	Leaf area (LA)	The whole surface of full leaf	cm^2	Spring of 2005	GH/WU	DH-38, DH-30
	LI	Leaf Index (LI)	Ratio of LL to LW, LL/LW	ratio	Spring of 2005	GH/WU	DH-38, DH-30
Turnip trait	TF	Turnip formation	Qualitative score of turnip formation (1 - 4 scale, Fig. 1C)	1-4	Autumn of 2005	OF/WU	DH-30
	TS	Turnip shoots	Number of shoot on the turnip (Fig. 1C)	number	Autumn of 2005	OF/WU	DH-30, BC1
	TL	Turnip length	Length from the top to bottom of turnip, measured by vernier caliper (Fig. 1C)	mm	Autumn of 2005	OF/WU	DH-30, BC1
	TWi	Turnip width	Width at the widest point, measured by vernier caliper (Fig. 1C)	mm	Autumn of 2005	OF/WU	DH-30, BC1
	TWe	Turnip weight	The mean weight of each turnip after harvesting	g	Autumn of 2005	OF/WU	DH-30, BC1

1: Details in Materials and Methods. 2: IVFa, Institute of Vegetables and Flowers, Haidian district of Beijing; IVFb, Institute of Vegetables and Flowers, Nankou farm, Changping district of Beijing; WU, Wageningen University; GH, Green house; OF, Open field.

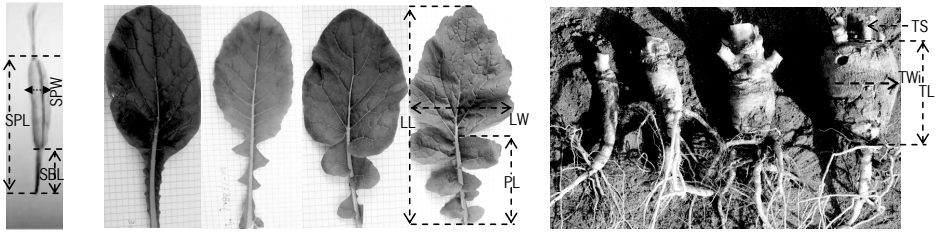

Fig. 1 Pictorial representation of measurement of seedpod (A), leaf (B) and turnip (C) traits.
Leaf edge shape (LES) classifications were indicated in B from left to right 1 - 4. Turnip formation classifications were indicated in C from left to right 1 - 4. For detail descriptions see Table 1.

Results

Variation in traits

The 5 parental lines belong to different morphotypes and displayed variation for flowering time, seed traits, plant height, leaf traits, and turnip traits (Table 2; Fig. 2). Transgression beyond the parental values within the used populations was observed for most of the measured traits including those for which parental values hardly differed, such as seed pod width (SPW) and leaf width (LW) in DH-38. The flowering time ranged from 17 days to 132 days within populations, depended on growing season, the parental genotypes and locations, and was transgressive in both directions in DH-38, DH-30 and BC1. In the RC-CC F2/3 population transgression for flowering time was only towards lateness as the RC-144 parent always had the shortest flowering time. The flowering time for F2/3 and DH populations was determined 3 times, and the mean values and ranges differed considerably. However, a strong positive correlation between different experiments was observed within populations, with correlation coefficients r = 0.31-0.61 in F2/3, r = 0.76-0.81 in DH-38 and r = 0.87-0.90 in DH-30.

Nine leaf traits were measured in the RC-CC F2/3 and the two DH populations before flowering. YS-143, the common parent of the two DH populations, had an average petiole length (PL) of 12.3 cm; the PC-175 parent had a short petiole of only 3.8 cm while VT-115 has no petiole. Within populations the petiole length ranged from 0 to 13 cm. Turnip related traits, like weight, length and width of the turnip, could only be measured in the DH-30 and BC1 population. In the BC1, all progenies had some degree of taproot thickening (Fig. 2), and the mean value of TL (57.5 mm), TWi (43.0 mm) and TWe (94.5 g) was higher compared to plants of DH-30.

Table 2 Phenotypic values of parental lines and corresponding populations

Trait	F2/3 (CC-156 X RC-144)				DH-38 (PC-175 X YS-143)				DH-30 (VT-115 X YS-144)			BC1 (VT-155 X YS-143) X VT-115)	
	CC-156	RC-144	Mean	Range	YS-143	PC-175	Mean	Range	VT-115	Mean	Range	Mean	Range
FL04sp	107.8	40.8	70.9	40.0-122.0	-	-	-	-	-	-	-	-	-
FL04wi	-	-	-	-	-	-	86.3	39.0-128.0	-	89.2	46.0-132.0	-	-
FL05sp	3.0	1.0	2.1	1.0-3.0	45.0	54.0	58.2	35.0-87.0	77.0	39.3	17.0-69.0	-	-
FL05wi	nd	27.4	46.6	27.3-66.2	-	-	-	-	-	-	-	-	-
FL05au	-	-	-	-	45.0	51.0	52.0	39.0-63.0	59.0	50.9	39.0-69.0	56.4	48.0-71.0
SPL	-	-	-	-	70.0	45.3	43.0	20.4-67.4	54.4	39.2	0.0-61.9	-	-
SPW	-	-	-	-	5.0	5.7	4.9	2.6-7.9	3.7	4.5	0.0-6.9	-	-
SBL	-	-	-	-	20.5	6.4	9.4	2.5-18.8	10.9	10.7	32.2	-	-
SC	-	-	-	-	1.0	5.0	2.9	1-5	5.0	2.4	1-5	-	-
SW04sp	2.25	1.05	1.3	0.5-2.2	-	-	-	-	-	-	-	-	-
SW05sp	-	1.15	1.8	1.0-2.7	-	-	-	-	-	-	-	-	-
PH	16.5	19.1	33.3	10.0-67.5	-	-	-	-	-	-	-	-	-
PB	-	-	-	-	6.0	-	-	-	8.0	7.5	3.8-15.0	8.0	0.0-17.0
LES04sp	2.0	3.0	2.3	1.0-4.0	-	-	-	-	-	-	-	-	-
LES05wi	2.0	3.0	1.9	1.0-3.9	-	-	-	-	-	-	-	-	-
LT04sp	1.0	0.0	0.3	0..0-1.0	-	-	-	-	-	-	-	-	-
LT05sp	1.0	0.0	0.3	0.0-1.0	-	-	-	-	-	-	-	-	-
LN04sp	9.0	5.0	6.8	4.0-13.0	-	-	-	-	-	-	-	-	-
LN05wi	18.8	5.0	13.2	6.5-20.8	-	-	-	-	-	-	-	-	-
LB05wi	0.0	1.0	0.4	0.0-1.0	-	-	-	-	-	-	-	-	-
LB05au	-	-	-	-	2.0	0	2.2	1.0-4.0	0	1.9	1.0-4.0	-	-
LL	33.4	4.8	24.7	8.8-44.8	12.2	9.0	10.7	4.7-19.2	32.9	13.9	6.4-25.7	-	-
LW	21.5	1.7	10.7	4.3-19.8	8.5	8.4	10.9	4.0-18.23	9.4	7.6	4.3-11.3	-	-
PL	-	-	-	-	12.3	3.8	5.7	0.1-13.0	0.1	6.2	0.1-13.8	-	-
LA	-	-	-	-	71.6	59.4	103.4	15.5-266.1	160.6	68.8	27.7-160.9	-	-
LI	-	-	-	-	41.0	29.7	39.5	15.6-67.5	74.1	37.8	21.6-66.9	-	-
TF	-	-	-	-	1	-	-	-	4.0	2.0	1-3.8	-	-
TS	-	-	-	-	0.0	-	-	-	5.0	2.6	0.0-14.0	3.7	0.0-11.0
TL	-	-	-	-	0.1	-	-	-	72.0	51.4	25.2-82.6	57.7	35.0-80.0
TWi	-	-	-	-	2.0	-	-	-	22.0	14.0	5.18-39.0	43.0	18.0-96.0
TWe	-	-	-	-	5.4	-	-	-	243.0	11.3	0.77-53.6	94.5	23.0-271.0

For trait abbreviation see Table 1; nd: no data because of no flowering at 130 days after sowing; -: not measured in the corresponding population.

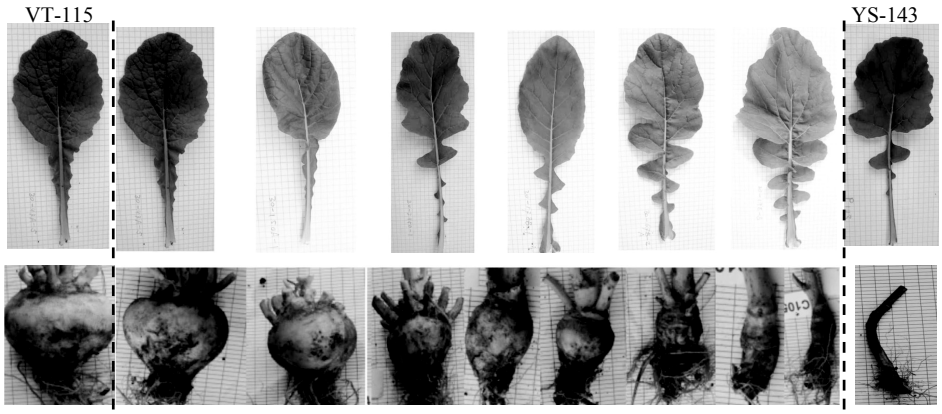

Fig. 2 An example of the variation for leaf traits (up) and turnip traits (down) of parental accessions VT-115 and YS-143 and selected individuals from population DH-30 (up) and BC1 (down) displaying all phenotypic variations.

We grouped all the morphological traits into 5 classes: flowering time measured in different growth seasons, seed-related traits, growth-related traits, leaf traits and turnip traits (Table 2). A PCA analysis was performed for all 17 traits representing the 5 classes in the RC-CC F2/3 and DH populations (Table 3). The different PCA components generally represent different traits and revealed correlations between the various traits. Flowering time contributed most to the principle component 1 (PCA-1) in each population. PCA-1 also reveals negative loading for leaf traits in F2/3 (LN05wi, LL and LW) and DH-38 (LL, LW, LA and LI), and turnip related traits in DH-30. This indicates that flowering time is more related to leaf-size traits than to leaf-edge-shape traits (LES), when the sink organ (turnip) is not formed. For principle component 2 (PCA-2), leaf traits and flowering time in DH-38, leaf traits and seedpod traits in DH-30, and leaf edge shape in F2/3 were the main variables. For the principle component 3 (PCA-3) seed related traits (SW04sp, SW05sp in F2/3; SPL, SPW, SBL in DH-38 and DH-30) and leaf traits (except for LES in F2/3, PL in DH-38, LL and LI in DH-30) were most important, indicating that these traits are partly correlated.

The positive correlation between flowering time and turnip traits was further analysed in BC1 populations (Fig. 3). Strong correlations were revealed between different turnip traits and flowering time in a backcross population (BC1) of 136 individuals. A number of significant genetic correlations were detected among the different turnip traits. The turnip width, - length and - weight were positively correlated with each other (correlation coefficient r = 0.49-0.89), but also with flowering time (correlation coefficient r = 0.57-0.67).

Table 3 Principle component analysis of morphological traits in F2/3, DH-38 and DH-30

Trait	F2/3			DH-38			DH-30		
	PCA-1	PCA-2	PCA-3	PCA-1	PCA-2	PCA-3	PCA-1	PCA-2	PCA-3
FL04sp	-0.2555	-0.2319	-0.3100	-	-	-	-	-	-
FL04wi	-	-	-	-0.3339	0.1838	0.1297	-0.3947	-0.0307	-0.081
FL05sp	-0.3519	-0.1311	0.1299	-0.2877	0.2934	0.0994	-0.3808	0.0326	-0.1302
FL05au	-	-	-	-0.3392	0.2036	0.0223	-0.3827	-0.0004	-0.0709
FL05wi	-0.4049	0.0404	-0.0073	-	-	-	-	-	-
SPL	-	-	-	-0.1917	-0.0603	0.3144	0.0716	-0.0710	-0.4717
SPW	-	-	-	-0.1871	-0.1852	0.5366	-0.1005	-0.2489	-0.2125
SBL	-	-	-	-0.1024	-0.1661	0.6434	0.007	-0.2614	-0.0371
SW04sp	0.1144	0.3243	-0.3832	-	-	-	-	-	-
SW05sp	0.1153	0.1522	-0.6333	-	-	-	-	-	-
PH	0.3580	0.1555	-0.2396	-	-	-	-	-	-
LES04sp	-0.0315	0.5516	0.1519	-	-	-	-	-	-
LES05wi	-0.0181	0.5367	0.0525	-	-	-	-	-	-
LN04sp	-0.1484	0.3862	0.3456	-	-	-	-	-	-
LN05wi	-0.3963	0.1563	-0.0471	-	-	-	-	-	-
LL	-0.3982	0.0705	-0.2635	-0.3400	-0.2516	-0.1282	0.1206	0.4608	0.0745
LW	-0.3917	0.0636	-0.2545	-0.2801	-0.3800	-0.2235	-0.0188	0.2415	-0.5333
PL	-	-	-	0.2019	-0.3180	0.0812	-0.1437	-0.3458	-0.2851
LA	-	-	-	-0.3195	-0.3413	-0.2034	0.0052	0.4104	-0.3858
LI	-	-	-	-0.3235	-0.3427	-0.2021	0.1201	0.4751	-0.0473
TF	-	-	-	-	-	-	-0.3886	0.0947	0.0606
TL	-	-	-	-	-	-	-0.1971	0.1515	0.0916
TWe	-	-	-	-	-	-	-0.3427	0.128	0.2388
Twi	-	-	-	-	-	-	-0.3442	0.1694	0.2394
PVE (%)	36.8	17.6	11.5	42.5	21.8	13.0	33.8	23.7	12.4

For trait abbreviation see Table 1; PVE: percent variance explained; -: not measured in the corresponding population.

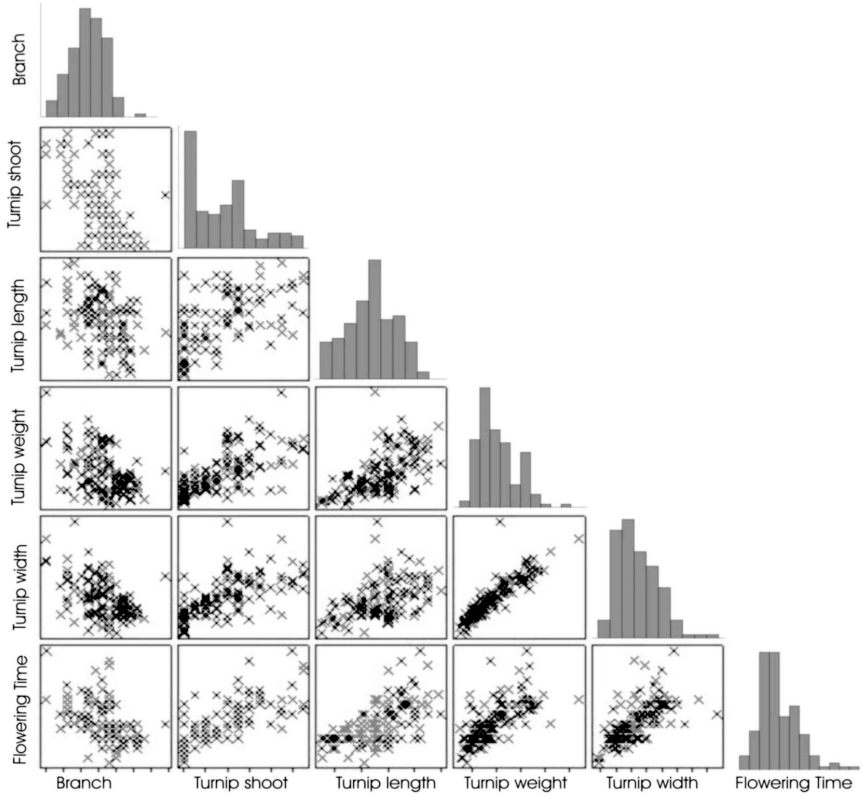

Fig. 3 Scatter plot matrix of turnip and flowering time traits generated from BC1 populations. The histograms along the diagonal provide a visual representation of the phenotypic variance for each of the traits. The off-diagonal scatter plots provide a visual representation of the correlation among the traits.

Construction of genetic maps

The genetic map of RC-CC F2/3 was described in chapter 4 and the genetic maps of DH-38 and DH-30 were described by Lou et al. (submitted). For the BC1 population, the AFLP pattern of F1 and parental genotypes (YS-143 and VT-115) revealed that the recurrent parent VT-115 was heterozygous at several loci, resulting in maximal 4 possible segregating alleles (one YS-143 allele from the F1 parent; one VT-115 allele from the F1 parent and two VT-115 alleles from the recurrent parent) in the backcross population. Thus the data were analyzed with the cross population (CP) algorithm of Joinmap to construct a linkage map. Only 58 YS-143 markers were used to construct the map, and less abundant VT-115 markers were not

used. Finally the BC1 linkage map includes 8 linkage groups with 40 AFLP markers covering 356 cM and 4 linkage groups could be assigned to R groups R01, R02, R03 and R08 based on common AFLP markers between the BC1 and DH-30.

QTL Mapping

Flowering time

A total of 7 QTL for flowering time (FLQTLn) were identified on R01, R02, R03, R06, R07 and R10 in 4 different populations evaluated in different growing seasons, 3 in F2/3, 3 in DH-38, 3 in DH-30 and 1 in BC1 (Table 4; Fig. 4). In RC-CC F2/3, the explained variation per QTL was generally lower (8.9-24.6%) than for the other populations (13.4-59.3%), which maybe caused by the fact that more QTL affecting flowering time were involved in the F2/3 population, and not all QTL could be detected in this study. A large percentage of phenotypic variance (20.0-59.3%) was explained by FLQTL-2 on R02 in the DH and BC1 populations. This QTL was detected in all populations, growing seasons and conditions. FLQTL-3, FLQTL-4, FLQTL-5, FLQTL-6 and FLQTL-7 were only detected in a single population and each QTL explained 15.1-19.7% of the variation. To investigate whether the same or different QTL positions were identified in the different populations, linkage maps were compared based on the common AFLP or SSR markers (Fig. 4). For FLQTL-2 on the top of R02, there was strong overlap of the 2-LOD support intervals across all the four populations, co-segregating with SSR marker *BrFLC2* (S2b in Fig. 4) in F2/3. We assume that FLQTL-2 in DH-38 map at the same location as in F2/3, DH-30 and BC1, even though the DH-38 map does not extend to that region. FLQTL-3 in DH-38 and FLQTL-4 in F2/3 are located on the same linkage group R03 but at opposite ends, therefore they represent two different flowering time QTL.

The FLQTLn detected in the populations were not always identical in different growing seasons and at different locations, suggesting genotype x environment effects, which could reflect the effects of temperature and day-length response on the expression of these flowering time QTL. The major QTL FLQTL-2 on R02 was detected in all experiments (FL04sp, FL04wi, FL05sp, FL05wi, FL05au) but other QTL are sensitive to environment, for example, FLQTL-5 on R06 was not detected in the open field experiment in DH-30 and FLQTL-7 on R10 was not detected in spring season 2005 in the greenhouse in DH-38.

In the F2/3 and DH populations, the earlier parent RC-144 or YS-143 always contributed alleles that decrease flowering time except for FLQTL-4 and FLQTL-3 on R03, which were

detected in spring 2005 in Beijing and autumn 2005 in Wageningen respectively (Fig. 4), where CC-156 and PC-175 contributed to the earlier flowering time alleles.

Table 4 Results of QTL analyses of measured traits in 4 *B. rapa* populations

QTL	Trait	Population	Linkage group	Position[a]	LOD	Exp%
FLQTL-1	FL04wi	DH-30	R01	middle	2.25	13.4
FLQTL-2	FL04sp, FL04wi, FL05sp, FL05wi, FL05au	F2/3, DH-38, DH-30, BC1	R02	top	2.17-9.66	8.9-59.3
FLQTL-3	FL05au	DH-38	R03	bottom	2.00	19.7
FLQTL-4	FL05sp	F2/3	R03	top	5.01	9.3
FLQTL-5	FL04wi	DH-30	R06	middle	2.42	18.3
FLQTL-6	FL04sp, FL05au	F2/3	R07	higher-middle	4.26-6.02	14.3-15.1
FLQTL-7	FL05au	DH-38	R10	top	2.70	17.9
SPQTL-1	SPL	DH-38	R01	middle	2.70	11.1
SPQTL-2	SBL	DH-38	R05	top	3.87	20.3
SPQTL-3	SPL, SBL, SPW	DH-38, DH-30	R07	higher-middle	4.68-6.39	27.2-38.1
SPQTL-4	SBL	DH-30	R09	higher-middle	3.08	25.5
SCQTL-1	SC	DH-38, DH-30	R09	middle	10.18-12.58	61.7-65.5
SWQTL-1	SW04sp, SW05sp	F2/3	R03	middle	2.30-3.25	10.2-11.6
SWQTL-2	SW05sp	F2/3	R08	higher-middle	4.22	17.6
PHQTL-1	PH	F2/3	R02	higher-middle	9.43	23.9
PHQTL-2	PH	F2/3	R03	higher	6.17	15.7
PHQTL-3	PH	F2/3	R07	higher-middle	3.00	8.9
LQTL-1	LB05wi, LN05wi, LL, PL, LA	F2/3, DH-38	R02	higher-middle	2.20-6.12	10.3-25.8
LQTL-2	LES04sp	F2/3	R02	lower-middle	3.32	6.5
LQTL-3	LES04sp, LES05wi, LW, LI	F2/3, DH-38	R03	bottom	3.04-6.44	20.5-26.4
LQTL-4	LW, LA, LI	F2/3, DH-30	R05	middle	2.33-3.16	7.0-24.2
LQTL-5	LB05wi, LW, LL, LI	F2/3, DH-38, DH-30	R06	middle	2.12-6.07	12.5-22.5
LQTL-6	LES04sp	F2/3	R06	lower-middle	3.47	9.1
LQTL-7	LL, LW, LB05wi, LN04sp, LN05wi, PL, LI	F2/3, DH-30	R07	higher-middle	3.41-6.03	13.7-21.9
LQTL-8	LA	DH-38	R08	higher-middle	2.15	10.6
LQTL-9	LW, LL	F2/3	R08	bottom	2.98-3.05	9.4-11.2
LQTL-10	LW, LA, LL	DH-38	R09	higher-middle	2.00-2.53	10.8-12.6
TuQTL-1	TF, TS, TL, TWi, TWe	DH-30, BC1	R02	top	4.74-7.08	24.0-40.0

For trait abbreviation see Table 1; Exp: Phenotypic variation of explained.

[a]: The about position on the linkage group in Fig. 4 based on common SSR or AFLP markers.

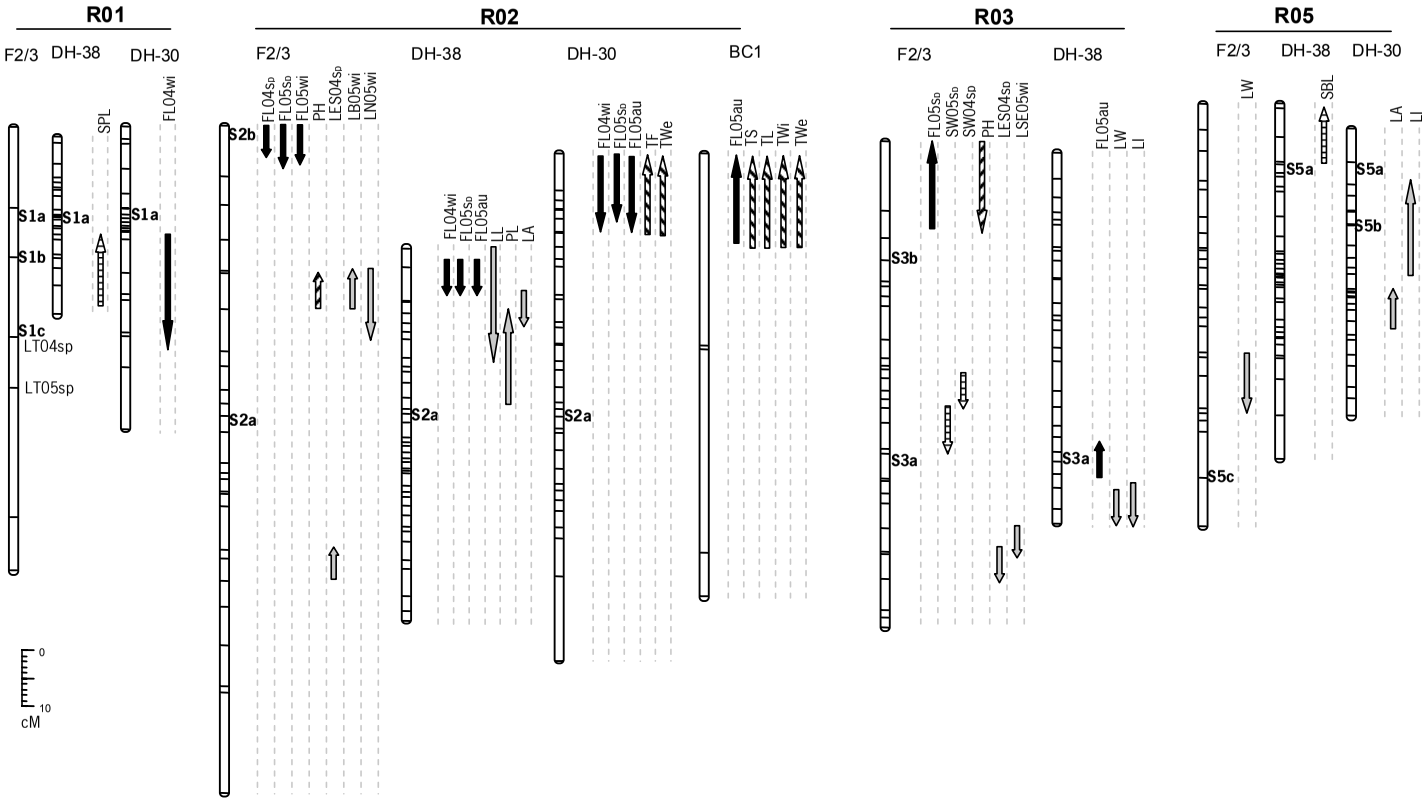

Fig. 4 Locations of QTL for the traits analyzed in the 4 mapping populations.

The linkage groups of different maps are aligned based on common SSR markers (S1-S10) or common AFLP (Lou et al. submitted). The lengths of the arrows indicate the 2-LOD support intervals. The traits (see abbreviations in Table 1) are indicated above each column. The direction of the arrow's head indicates the allelic effect: upward, RC-144 increases and CC-156 decreases for F2/3, YS-143 increases and PC-175/VT-115 decreases in DH-38 and DH-30; downward: CC-156 increases and RC-144 decreases, YS-143 decreases and PC-175/VT-115 increases in DH-38 and DH-30. The filling pattern of arrows refers to different groups of phenotypic traits. Flowering time; Seed; Plant height; leaf; Turnip

S1a, BRMS096; S1b, Ra2G09; S1c, BRMS037; S2a, Na12H09; S2b, BrFLC2; S3a, BRMS043; S3b, BRMS042; S5a, BRMS034; S5b, BRMS007; S5c, Ra3H10; S6a, BRMS014; S6b, Na12H07; S7a, BRMS018; S7b, Ol12E03; S7c, Ra2A01; S7d, BRMS036; S8, Ra2E12; S9a, BRMS051; S9b, Na10A08; S9c, Ol12F02; S9d, Ol10D08; S10, BrFLC1

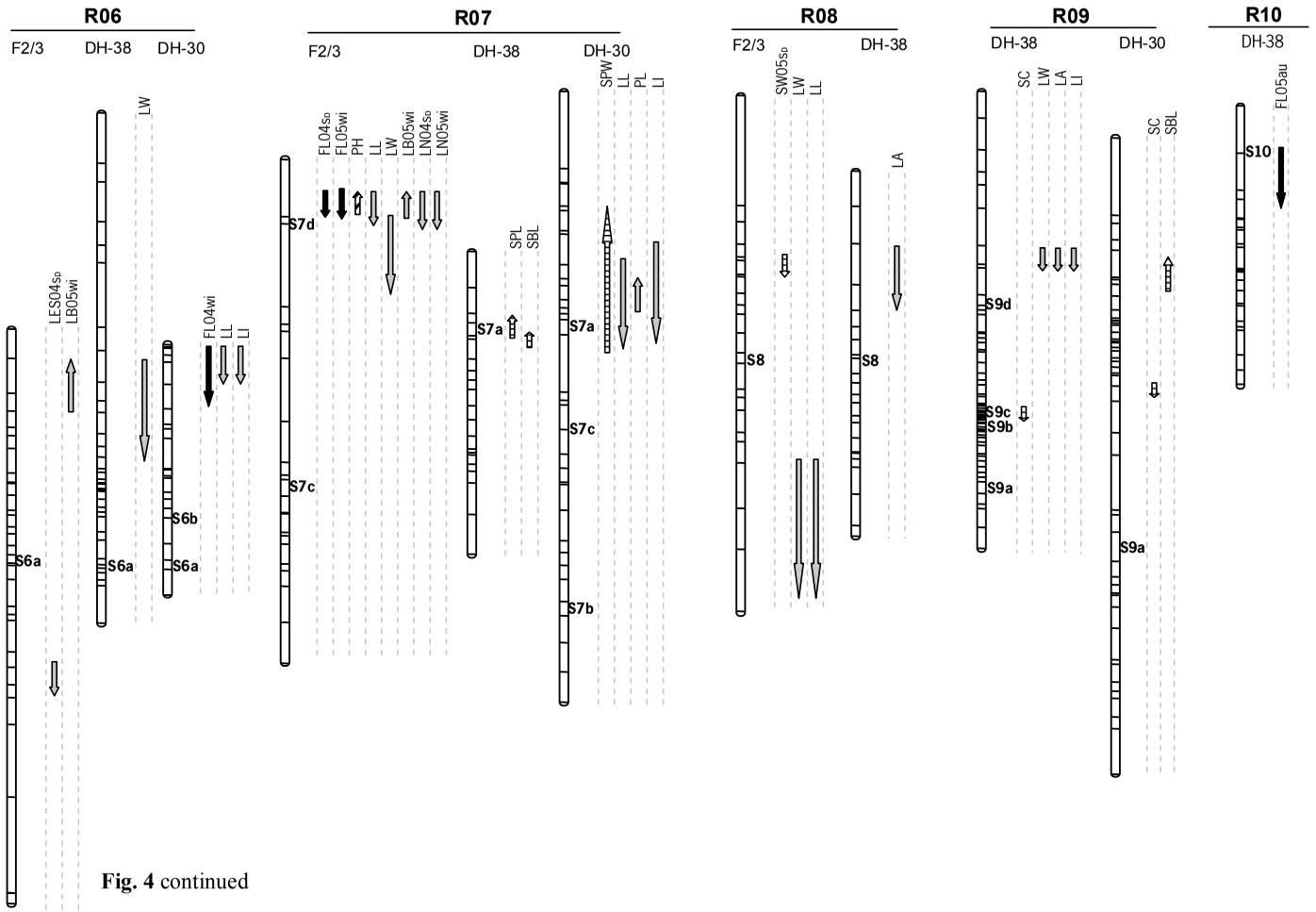

Fig. 4 continued

Seed traits

We detected 4 QTL for seedpod traits (SPQTLn) and 1 QTL for seed coat colour (SCQTL-1) in the two DH populations, and 2 QTL for seed weight (SWQTLn) in the RC-CC F2/3 population. The proportion of total variation explained by each QTL ranged from 11.1 % to 38.1% for seed pod traits, 61.7% to 65.5% for seed coat colour and 10.2% to 17.6% for seed weight (Table 4; Fig. 4).

Two genomic regions, on R05 and R09, each harbored a single seedpod trait (SBL) in DH-38, and one genomic region on R07 affected three seedpod traits (SPL and SBL in DH-38, SPW in DH-30). In both the DH-38 and DH-30 population, one major genomic region on R09 affected seed coat colour (SC) with a high explained phenotypic variation (>60%) and LOD value (>10.0), indicating almost monogenic inheritance of the yellow seed trait of YS-143. In RC-CC F2/3 one genomic region in the middle of R03 affected seed weight (SW) and was detected in the 2004 and 2005, with the RC-144 allele decreasing seed weight.

Growth-related and leaf traits

Three QTL affecting plant height (PHQTLn) were detected on R02, R03 and R07 in the RC-CC F2/3 population, explaining only 8.9-23.9% of the phenotypic variation, which was possibly caused by the scoring of plant height at a later flowering stage affecting plant or stem elongation. Two markers, E33M51-7CC and BRMS037, 10 cM apart on R01, are linked with the number of leaf trichomes (LT04sp and LT05sp). We could not detect QTL for the number of branches (PB) in DH-30 and BC1 populations.

Ten QTL for leaf traits (LQTLn), distributed over 7 linkage groups, were detected in the F2/3 and DH populations; the proportion of total variation explained by each QTL ranged from 6.5% to 26.4% (Table 4). The five different parents contributed alleles with effects in both directions to most of these traits (Fig. 4). Seven genomic regions affected two or more leaf traits, where LW always co-segregated with other leaf traits (LL, LA, LI, LN or LB) except for one region, PHQTL-1,in the high middle on R02. LQTL-1 on R02, LQTL-3 on R03, LQTL-4 on R05, LQTL-5 on R06, and LQTL-7 on R07 were detected in multiple populations and related to multiple traits, appeared to be the major QTL affecting leaf size in the used populations. Two genomic regions (LQTL-2 on R02 and LQTL-6 on R06) affected only a single leaf trait (LES04sp), representing QTL of leaf serration. It is hard to conclude whether LQTL-8 on R08 maps in a same region to LQTL-9 on R08 because only one common SSR

marker connects the R08 maps of F2/3 and DH-38.

Turnip formation

We detected one major QTL for turnip related traits (TuQTL-1) both in DH-30 and the BC1 population (Table 4; Fig. 4) and this QTL co-located with FLQTL-2 on the top of R02. In the BC1 population the TuQTL-1 explained about 24.0% of variation and in DH-30 the QTL explained 36.7-40.0% variation.

Clustering of QTL

We detected some co-locations of QTL where one locus controlled multiple traits, which are possibly physiologically related. On the top of R02 many QTL co-localized, mainly for flowering time, but also for leaf traits in F2/3 and turnip related traits in both the DH-30 and BC1 population (Fig. 4). Clusters of QTL were also detected on the middle of R06 and on the higher middle of R07. Clustering of QTL was consistent with the strong genetic correlations observed among specific traits (turnip and flowering traits, Fig. 3), which is also obvious from the PCA for the traits that have significant loadings for the respective PCA component. Since most of traits are not tested across the different populations, it is difficult to compare PCA QTL positions in the different populations.

Discussion

Flowering time

In this study we mapped QTL for flowering time in 4 different populations derived from crosses between diverse parental morphotypes. This multiple population approach has the advantage that alleles of 5 parental accessions can be evaluated and possibly revealed a large number of genomic regions harbouring allelic variation for flowering time. Seven possible genomic regions on 6 linkage groups affect flowering time, two of them (FLQTL-2, FLQTL-6) were detected in different conditions, suggesting that they are not, or only marginally, affected by the environment. Some of the flowering-related QTL (FLQTLn) that we found in the used populations co-localized with previously published QTL detected in other *B. rapa* populations. In Fig. 5 maps of R02, R03, R10 and Arabidopsis chromosomes 5 are depicted, with the map positions of *FLC* and *MAF* paralogues and in which regions the identified QTL both in this and in other studies (Schranz et al. 2002; Kim et al. 2006).

In previous QTL analyses of flowering time in *Brassica*, *FLC* or *CO* are mentioned as candidate genes underlying flowering time QTL in *B. napus* (Tadege et al. 2001), *B. nigra*

(Lagercrantz et al. 2002; Kruskopf Österberg et al. 2002) and *B. rapa* (Schranz et al. 2002; Kim et al. 2006). *CO* and *FLC* are linked on Arabidopsis chromosome 5 (15 cM apart) and based on the synteny between Arabidopsis and *Brassica* (Parkin et al. 2005) they are linked in *B. rapa* as well. A number of *FLC* paralogues (*BrFLC1, BrFLC2, BrFLC3* and *BrFLC5*) were mapped in this and previous studies (Schranz et al. 2002; Kim et al. 2006). However *CO* has not been mapped yet, so we cannot conclude whether *FLC* or *CO* are candidate genes underlying the flowering time QTL. The use of SSR markers allowed the alignment of our maps to the *B. rapa* reference maps and to compare QTL positions between populations. Four *FLC* homologues, *BrFLC1*, *BrFLC2*, *BrFLC3* and *BrFLC5*, have been cloned in *B. rapa* (Schranz et al. 2002). Using in situ hybridisation, genetic mapping with AFLPs or linked SSR markers, *BrFLC2* and *BrFLC3* (*a* and *b*) are mapped on the top of R02 and on the bottom of R03 between SSR marker BRMS043 and BRMS008; *BrFLC-1* was mapped on the top of R10 and *BRFLC5* was also mapped on the lower middle of R03, 33 cM apart from *BrFLC3* (Kim et al. 2006). *BrFLC2* was mapped on R02 using SSR markers (Yang et al. 2006) in the F2/3, DH-38 and DH-30 populations, and co-localized with the major FLQTL-2 on R02. Although *BrFLC3* is not mapped in our experiment, FLQTL-3 was detected at a similar position on the bottom of R03 between SSR marker BRMS043 and BRMS008 in DH-38. FLQTL-4 is detected on the top of R03 in the RC-CC F2/3 population at which position the flowering time QTL, FR2 (Osborn et al. 1997), co-localizing with *BrFLC5* (Schranz et al. 2002). The orientation of the different maps compared to the maps of Kim et al. (2006) was based on a number of SSRs. However, only one SSR (BRMS043) was common between the F2/3 and the DH populations. The position of FLQTL-4 relative to FLQTL-3 both on R03 could not be determined and therefore *BrFLC5* may or may not be a candidate for FLQTL-4.

We could not find any QTL at the *MAF* locus on the bottom of R02, probably because the parental accessions carry similar alleles at this locus. In DH-38 *BrFLC1* maps on the top of R10 and co-localized with the minor FLQTL-7 QTL that was detected in autumn of 2005. Our data suggest that several of the flowering time loci correspond to the map positions of *FLC* paralogues. In summary, FLQTL-2 on the top of R02, FLQTL-3 on the bottom of R03, FLQTL-7 on the top of R10 might correspond to the *FLC* homologues *BrFLC2*, *BrFLC3* and *BrFLC1* respectively. FLQTL-2 determines the flowering differences between the early oil types and the other middle late morphotypes. Loci that are detected only in some of the

populations we studied represent loci that differ between the different parental vegetable types used.

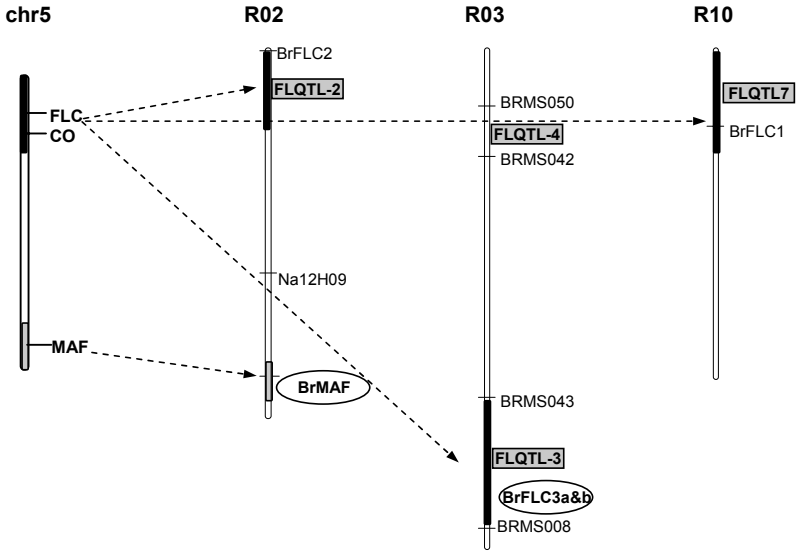

Fig. 5 Comparative map of flowering time QTL (FLQTLn) with flowering time related gene *FLC* in Arabidopsis chromosomes (chr) 5.

The maps of R02, R03 and R10 are from the integrated map of DH-38 and DH-30 (Lou et al. submitted). The SSR markers that are used for alignment with previous *B. rapa* maps are indicated in the map (Schranz et al. 2002; Kim et al. 2006). The gray boxes represent the FLQTL identified in this study. Flowering time related genes that have not been mapped in our populations are indicated in ovals. The information on map positions in *B. rapa* of flowering related genes *FLC* and *MAF* on chr 5 in Arabidopsis is from Kim et al. (2006), illustrating in the left side.

In Arabidopsis a number of additional flowering related genes were described and were shown to interact in a network. Examples are *FT* (*FLOWERING LOCUS T*) and *LFY* (*LEAFY*). In *B. rapa*, one *FT* paralogue was recently mapped on R07 (*BrFTa*) close to SSR marker BRMS036 (personal communication with Jungsun Kim); FLQTL-6 in F2/3 maps near this SSR marker BRMS036 on R07 (S7d, Fig. 4), and likely corresponds to *BrFTa.* One *LFY* paralogue was mapped on the lower middle of R06 (Kim et al. 2006); FLQTL-5 in DH-30

was mapped on the middle of R06. However, we cannot directly compare the synteny between FLQTL-5 and *BrFLY* because of lack of common SSR markers in this region. For the flowering time QTL on R01 (FLQTL-1) identified in this study, candidate genes cannot yet be suggested.

The evaluation of flowering time in different seasons and growing conditions allows one to evaluate the expression of the QTL in different environments. In this study, FLQTL-2 on R02 is the major QTL controlling flowering time. However it explains 38.8% of phenotypic variation in the open field experiment of autumn 2005 in DH-38 compared to 59.3% and 56.6% in the other two experiments (2004 winter and 2005 spring). In DH-30 FLQTL-2 detected in the autumn 2005 experiment is a minor QTL with low LOD score (2.17) and explained variance (20.0%), while it has a larger effect (35% and 50% variation explained) in the other two growth seasons (2004 winter and 2005 spring). The lower expression of FLQTL-2 in autumn 2005 is possibly caused by the low temperature or effects of the interaction with other flowering-related genes affected by the environment. *FLC* is known to be repressed upon vernalization (Koornneef et al. 2004; Salathia et al. 2006), and one homologue (*FLC2*) co-segregated with the flowering time QTL FR1 on R02 (Schranz et al. 2002), which is not related to vernalization response (Osborn et al. 1997). It further suggests that *FLC2* is a candidate for FLQTL-2.

Turnip formation

In the present study a single QTL for each of the traits, turnip width, turnip weight, turnip length was detected at the top of R02 and this QTL for turnip formation co-localizes with the major flowering time QTL (FLQTL-2). There are two possible explanations for this co-localization, pleiotropy or tight linkage. A plant that invests in reproduction needs to flower and allocate its energy to the developing seeds, while turnip formation requires the redirection of most of assimilates to the roots. The direction of the sink can be regulated by transcription factors and hormone levels. However, since several very late oil and vegetable type accessions do not make turnips genetic variation for genes specifically controlling turnip formation must be present in the *B. rapa* gene pool, even when early flowering is epistatic to turnip formation. This should further be investigated by fine mapping and the analysis of other populations segregating for turnip formation, especially those not differing much in flowering time.

Other morphological traits and genetic architecture of trait variation

Besides flowering time and turnip formation, we provide a genetic analysis of other morphological traits. A number of QTL underlying these traits in the used populations are observed at many loci throughout the whole genome, for example, leaf perimeter, leaf length, petiole length. Typically the parental accessions contained alleles that both increased and decreased leaf phenotypes, resulting in large transgressive segregation within the population.

Co-location of QTL for phenotypic traits are found in many cases, indicating that these loci may have an overall effect on plant development and suggest a pattern of genetic integration of morphological traits. For example, the genomic regions at the top of R02 and the higher middle of R07 generally affect seed traits, growth-related traits and leaf traits, besides flowering time and turnip formation, in which multiple genes or pleiotropic loci controlling related developmental characteristics may be involved. For only leaf traits, we identified one QTL LQTL-5 on R06 across all the populations affecting 4 traits (LB05wi, LW, LL and LI). The reason for this is suggested to be the distinct morphology of parental accessions (Rapid cycling, Chinese cabbage, Yellow sarson, Turnip and Pak choi) that logically involved distinct sets of genes underlying the leaf shape and size of the plant.

Furthermore, we found no or less co-locations for specific traits, like LQTL-2 affecting leaf serration (LES04sp) on R02 and the seed coat colour QTL on R09 (SCQTL), indicating that these loci specifically regulate certain developmental processes that contributed less to the other traits. Leaf serration possibly inherits independently with other leaf shape or size traits (LL, LW, LA and LI). In DH-38, a seed colour QTL (SCQTL) maps near SSR markers, Na10A08 and Ol12F02, one of which (Na10A08) also showed strong association with seed coat colour in *Brassica juncea* (Padmaja et al. 2005).

The coincidence of QTL locations generally supports the observed phenotypic correlations, such as flowering time and turnip formation mentioned above. The QTL clusters in this study, such as genomic region on R02, R03, R06 and R07, reflect the genetic correlations between the studied traits. It would be interesting to measure more floral traits (petal and sepal development) and further investigate the correlation between floral and leaf traits, for which in Arabidopsis a weak correlation was observed and differentiation between floral and vegetative modules was suggested (Juenger et al. 2000).

Acknowledgements

We thank Nurmi P. D. Pangesti, and Jiangling Xiong for phenotypic scoring of turnip characteristics, testing of SSR markers and constructing the BC1 genetic map. We are grateful to staff in Plant Science Experimental Centre of Wageningen University, the Netherlands, and Yanguo Zhang and Yanling Liu in the Institute of Vegetables and Flowers (Beijing), Chinese Academy of Agriculture Science, China, for taking care of the plants and assistances. The project is sponsored by the Royal Dutch Academy of Sciences (KNAW), the Asian Facility (project AF01/CH/8 "Sino-Dutch Genomic Lab and Vegetable Research Center") and a fellowship from an exchange student programme (Huygens) between Chinese Scholarship Council and Nuffic in the Netherlands.

Chapter 6

General discussion

Naturally occurring genetic variation is a useful resource for the genetic mapping of complex phenotypic traits (Koornneef et al. 2004). In this thesis, the genetics of phytate and phosphate representing (anti)-nutritional food components and a number of developmental and morphological traits were investigated within cultivated *Brassica rapa*, used as oil or vegetable crop. One of the objectives is to obtain insight in the structure of the gene pool and the genetic basis of natural variation in *B. rapa*. In this chapter, the genetic relationships and population structure of a large *B. rapa* collection of cultivars, the genetic control of the several traits for which variation was observed, their underlying QTL and the genomic comparison with QTL identified in Arabidopsis at syntenic positions, will be discussed.

The genetic variation in the *B. rapa* gene pool

Molecular analysis using AFLP fingerprinting (chapter 2 and 3) has shown that there is considerable genotypic variation in the *B. rapa* gene pool. The amount of genetic variation present within groups of cultivars, as identified in the phenetic tree, contributed considerably to the high polymorphism rate observed. The AFLP fingerprints of the collection of 160 accessions were also analyzed with STRUCTURE software and four subgroups (S1-S4) were identified. For Chinese cabbage, Pak choi, and turnip, the high amount of polymorphisms indicated that a wide genetic variation still exists within the gene pools of these different cultivar groups. This high level of genetic diversity within groups is possibly caused by their wide distribution in geographical regions with different growth habitats and consumer preferences. This also explained why the different morphotypes could emerge independently in the different regions (like for oil-types: Chinese turnip rape, Pakistani Winter turnip rape and Indian Yellow sarsons) and why population structure is present within this species even within China. Within an ecological region, breeding for specific types (like the Pak choi related rosette shaped Wutacai and early flowering Caixin) is necessary to meet specific uses. We can not directly clarify the position of these smaller groups of special cultivars in this study because of their complex and often unknown pedigrees and unclear local origins. However, the data provides insight into possible application of favorable alleles within groups of the gene pool for breeders. In the collection that was used for both chapter 2 (genetic

relationships as illustrated by a phenetic tree) and chapter 3 (genetic analysis using association mapping), only 2 Yellow sarson and 6 spring oil accessions in Oil2 (chapter 2) or S3 (chapter 3) were involved. The narrow genetic basis and independent genetic background of the Oil2 group and the structured subgroup S3 may indicate that it is difficult to breed for new oil-type cultivars. However, we may have missed important accessions from this subgroup, like from the Canadian germplasm that was underrepresented in this study. Other *B. rapa* oil-types, like winter turnip rapes from Pakistan in Oil1 (chapter 2) and subgroup S2 (the turnip group of chapter 3), and Chinese turnip rape in PC2 (chapter 2) and subgroup S1 (chapter 3), could be the candidate donors for the introduction of new favorable alleles also in *B. napus* breeding.

Constructing a UPGMA tree from the AFLP data cannot be regarded as phylogeny, because of more detailed analysis of the non-dependence of AFLPs (Koopman 2005). By constructing a phenetic tree using UPGMA program, the genetic relationships between different cultivar groups could be revealed (Zhao et al. 2005). Cluster analysis showed a clear separation of two main phenetic groups, which are consistent with previous proposed groups representing two main centers of origin for cultivated *B. rapa* (Prakash and Hinata 1980; Song et al. 1988b). The oil types from the Indian subcontinent may represent a third center. The results showed similarities more related to geographic origin (East Asia versus Europe) than to morphological groupings, suggesting either an independent origin and/or a long and separate domestication and breeding history in both regions. However, based on allelic frequencies analysis it was also likely that a part of the genetic background is shared by many accessions. These accessions that are genetically related possibly shared part of their breeding history. One other explanation for the shared genetic background is that in fact this is caused by some clusters of AFLPs occurring at chromosome sites where some important genes have been introgressed. It would be interesting to further investigate the pattern of polymorphism using high-throughout genotyping techniques to obtain insight into the genome-wide haplotype structure of this species.

Even though the amount of genetic variation present within the groups was often comparable to the variation between the different groups, groups can still be differentiated by a few markers/genes that are significantly associated with specific morphotypes. A number of those markers that differed between subpopulations provide an example of linkage disequilibrium

(LD) due to population structure. Marker frequencies between groups may be causally responsible for phenotypic differences, or may differ due to chance processes. Analysis of LD among 76 mapped markers (data not shown) showed that some associated markers on different linkage groups were also possibly related to subgroups, indicating that different regions of the genome can be involved in the morphological differences. One objective of future studies is to use mapped markers such as anchored SSR markers or gene targeted markers to understand the LD level across the genome, facilitating the application of association mapping appropriately in *B. rapa*.

The molecular basis of natural variation for traits in *B. rapa*

The subdivision of the population into structured subgroups was used to analyze the variation of a number of traits. The analysis of natural variation in morphological traits provides insight into the genetics of complex traits that is essential for *B. rapa* breeding. Ultimately it will be useful to discover the molecular functions of genes affecting a particular trait and further characterize the molecular variation of the identified genes within and outside the species. Phytate and phosphate accumulation are some of the determinants of micronutrient availability in leaves and seeds, and their modification provides the potential for extending the product range derived from *B. rapa*. The extensive variation of leaf phosphate might be used for better phosphate use efficiency in leafy vegetable breeding. Phytate content is especially relevant for oil-types, where high phosphate and low phytate in seeds providing the meal, is desirable.

The genetic determinants accounting for natural variation were identified by genetic analysis of in germplasm collections and mapping populations. The loci controlling the traits were detected using a collection (160 accessions) representing various morphological types and geographical origins, and multiple segregating populations (one F2, four DHs and one BC1) based on parents from different cultivar groups that were grown at different locations and conditions. Fig. 1 summarizes all QTL that influence phytate and phosphate accumulation in seeds and leaves, and the morphology traits including flowering time, seed-related traits, growth-related traits, leaf-related traits and turnip-related traits that were identified in this study. The summary shows the genetic factors affecting the studied traits and provides insight into possible regulation underlying traits integration.

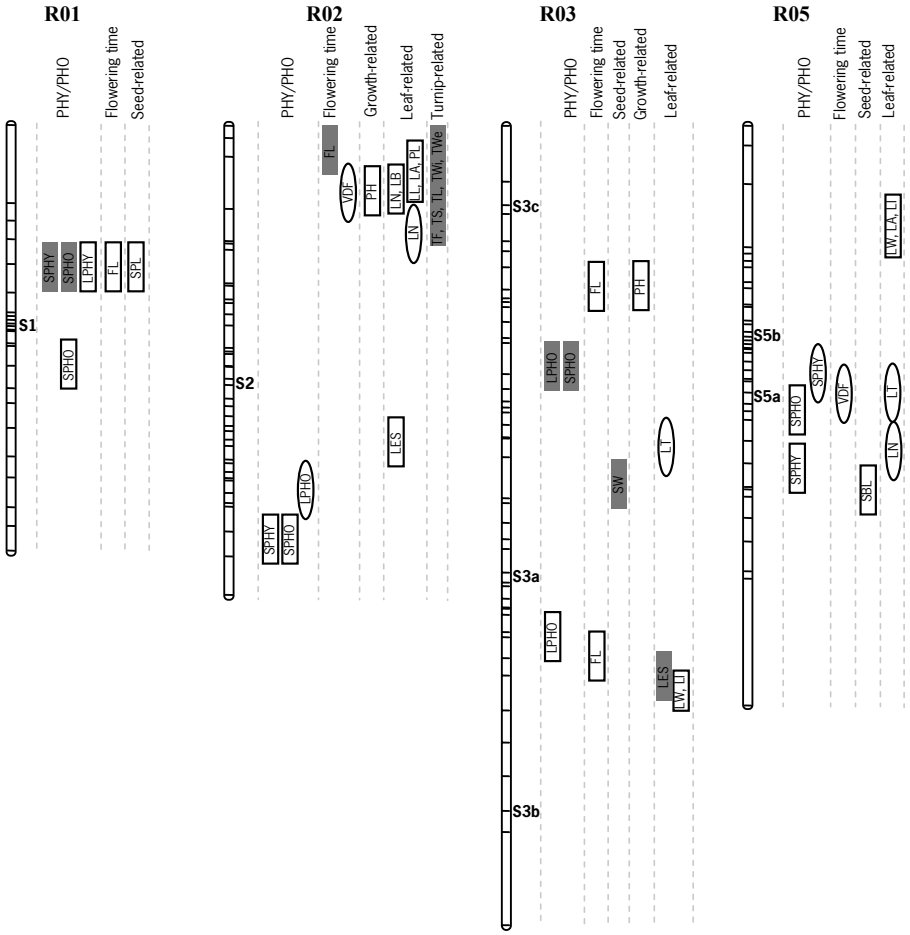

Fig. 1 Schematic view of the *B. rapa* linkage group showing locations of QTL affecting phytate and phosphate (PHY/PHO) traits (SPHY, SPHO, LPHY, LPHO), flowering time (FL, VDF), seed-related traits (SPL, SPW, SBL, SC, SW), growth-related trait (PH) leaf-related traits (LES, LN, LB, LL, LW, PL, LA, LI) and turnip-related traits (TF, TS, TL, TWi, TWe).

The genetic map is an integrated map of two DH populations, described by Lou et al. (submitted). The QTL are identified in the collections in Chapter 3 and several populations (F2/3, DH-38, DH-30, DH-03, DH-CC and BC1) in Chapter 4 and Chapter 5, for which the locations are based on the common SSR or AFLP markers between different maps. The SSR markers are indicated by S1-S10. The gray boxes represent the QTL identified in multiple populations or locations, whereas the QTL identified only in one population are in white boxes, and the QTL identified in collections are in white ovals. The trait abbreviations are illustrated in Chapter 3 (Fig. 2), Chapter 4 (Table 3) and Chapter 5 (Table 1).

S1, BRMS096; S2, Na12H09; S3a, BRMS043; S3b, BRMS008; S3c, BRMS050; S5a, BRMS034; S5b, BRMS007; S6a, BRMS014; S6b, Na12H07; S7a, BRMS018; S7b, BRMS040; S7c, Ra2A01; S7d, Ol12E03; S8, Ra2E12; S9a, BRMS051; S9b, Na10A08; S9c, Ol12F02; S9d, Ol10D08; S10, BrFLC1.

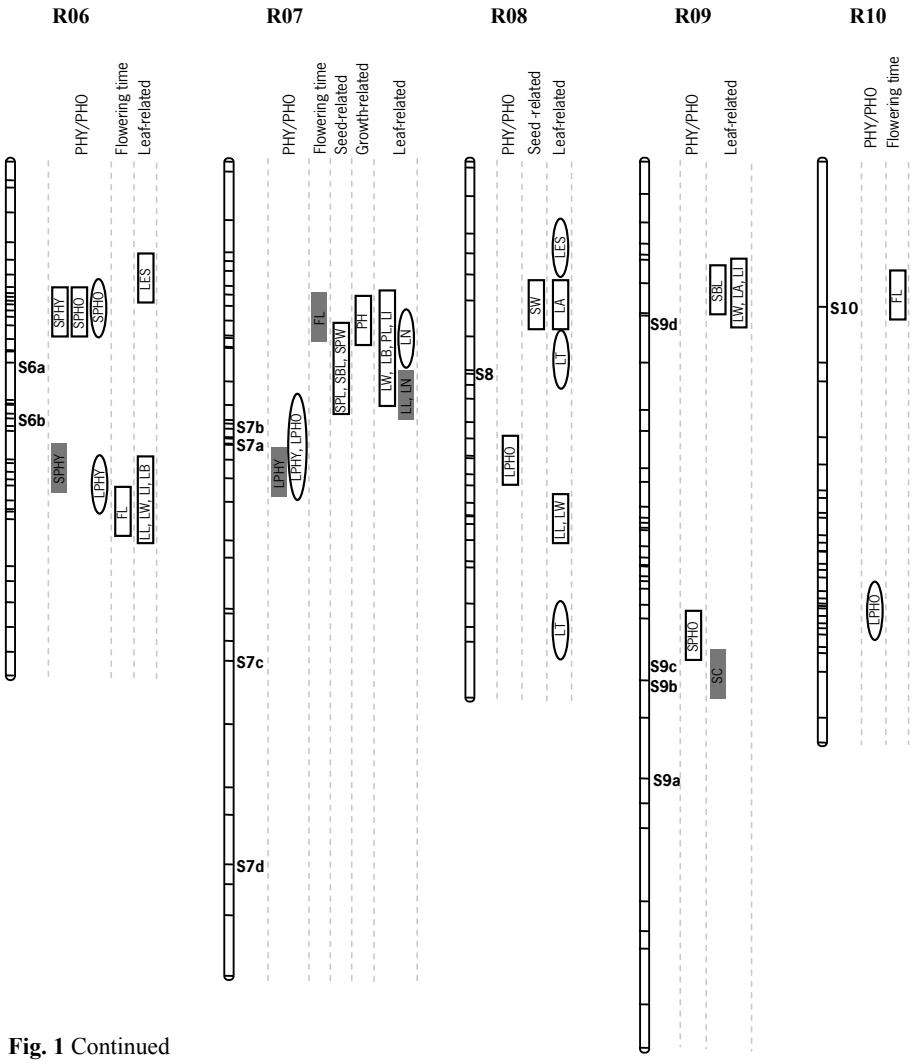

Fig. 1 Continued

Based on grouping the cultivars using STRUCTURE an association between population structure and trait values was detected, even for phytate and phosphate levels, for which we assumed that there has not been deliberate breeding. When screening of the large collection of unrelated *B. rapa* genotypes, association mapping allows the detection of associations between trait values and marker alleles. A number of marker-trait associations have been observed, of which some of them were also detected in the genetic analyses of populations

(Fig 1). For 16 QTL with known map position identified by association mapping as reported in chapter 3, eight were detected at similar positions on the same linkage groups in the QTL mapping studies described in chapter 4 and 5. This illustrates that association mapping is an important genetic approach to detect QTL. Since association mapping was performed on a dataset with many non-mapped AFLP markers, a through comparison of the QTL identified by the two approaches is not possible. In the future, this could be confirmed by linkage analysis in mapping populations designed on the basis of phenotype and marker contrasts derived from association mapping studies.

Part of the genetic variation present among accessions is undetectable by comparing these accessions, but can only be revealed when accessions are crossed and segregating individuals of the offspring display phenotypes outside the range of variation of the parents (Koornneef et al. 2004). Considerable variation and transgression was observed for most of traits in many of populations analyzed in this study, including the DH-CC population, which is derived from a cross within the Chinese cabbage cultivar group. This indicated that the parental accessions of the used populations carry alleles that both decrease and increase levels at several loci. Multiple mapping populations, with some of them sharing one common Yellow sarson parent, were applied in QTL analyses of selected traits. Common AFLP or SSR markers were used to align the different genetic maps and compare QTL positions.

A number of QTL underlying these traits are observed at many loci throughout the whole genome, for example, leaf-related QTL are distributed over 7 linkage groups. Co-locations of QTL for different traits were found in many cases (Fig. 1), which might suggest pleiotropy or tight linkage. For example, the genomic region in the higher middle of R07 generally affects phytate/phosphate levels and various morphological traits including seed-, growth- and leaf-related traits, indicating that a genomic region may have an overall effect on plant development from which genetic variation of other traits derives. Co-localization of QTL for flowering time and turnip related traits is an example of this, as detected on the top of R02, a positive correlation is seen between flowering time and turnip formation. One hypothesis is that transcription factors like *FLC*, co-localizing with this QTL, regulate both traits. It will be interesting to identify QTL for turnip formation in a population derived from a cross between a turnip and other non turnip forming late flowering types, like Winter oil, to identify loci specifically controlling turnip formation. Genomic regions on R01, R02 and R07, with QTL

detected in multiple populations and related to multiple traits represent the major genomic regions affecting related traits. Absence of locations of QTL for specific traits such as leaf-edge-shape (LES) and leaf trichomes (LT), suggest that these traits possibly inherit independently. Some QTL affecting a particular trait were only found in a single population, suggesting differences between populations (parents) or genotype x environment effects on these traits. The use of other accessions may identify additional loci affecting these traits.

For traits, like phytate and phosphate levels that are physiologically related, co-locations of QTL could be explained by the presence of a single locus involved in the accumulation of phosphate, common for those traits. Combining the parental allelic effects, the presence of QTL controlling total phosphorus levels and others specific for the phytate pathway are suggested. Phytate/phosphate levels and ratio's have impact on seed and plant growth, where low phytate resulted in yield losses (Raboy 2001). One SPHO QTL is located on R03 near the locus where QTL affecting seed weight (SW) and plant height (PH) are located, suggesting that these traits are correlated. The further investigation by fine mapping and/or the analysis of additional populations might clarify these co-locations.

The genomic comparison between *B. rapa* and Arabidopsis

The size of the *B. rapa* genome and its close relationship to the model plant Arabidopsis has several advantages for genetic research. The synteny between *Brassica* and Arabidopsis will allow a direct comparison of map positions of the two species (Parkin et al. 2005; Schranz et al. 2006) and might assist the identification of candidate genes that are already known or will be known in Arabidopsis. Some of the anchored SSRs on the previous maps and in this study allow the comparison of the maps presented in this study to the genomic blocks as defined by Schranz et al. (2006) and Parkin at al. (2005).

The preliminary genomic comparison with phytate/phosphate QTL and flowering time genes identified in Arabidopsis at syntenic positions was depicted in chapter 4 and 5. The QTL identified in Arabidopsis could align to multiple regions affecting the same trait in *B. rapa,* contributing to the wide variation in relevant traits in *Brassica* (Schranz et al. 2002). Three copies of *Brassica* orthologues of the Arabidoposis *FLC* genes (*FLC1*, *FLC2* and *FLC3*) were likely aligned to three different regions, corresponding to 3 flowering time QTL identified in this study. For phytate and phosphate accumulation also syntenic regions could be identified

where QTL for the same traits could be located in both species. In those syntenic blocks with phytate/phosphate QTL in both Arabidopsis and *B. rapa*, only 1-2 QTL regions in *B. rapa* corresponded to one syntenic region in Arabidoposis. Adding additional SSR and gene targeted markers to the *B. rapa* linkage groups will further identify other QTL in syntenic regions and improve the accuracy of identification of syntenic Arabidopsis-*B. rapa* QTL.

The growing resource of *B. rapa* genome sequence data (http://www.brassica.info) will become more widely available, sequence-level co-linearity and the molecular signatures underlying conserved blocks between the two species will be identified, which will speed up QTL fine mapping and gene identification.

References

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

Ajisaka H, Kuginuki Y, Shiratori M, Ishiguro K, Enomoto S, Hirai M (1999) Mapping of loci affecting the cultural efficiency of microspore culture of *Brassica rapa* L-syn. *campestris* L using DNA polymorphism. Breed Sci 49:187-192

Ajisaka H, Kuginuki Y, Yui S, Enomoto S, Hirai M (2001) Identification and mapping of a quantitative trait locus controlling extreme late bolting in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis* syn. *campestris* L.) using bulked segregant analysis. Euphytica 118:75-81

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MTR (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a *Ler*/*Cvi* recombinant inbred line population. Plant J 14:259-271

Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C, Toomajian C, Traw B, Zheng H, Bergelson J, Dean C, Marjoram P, Nordborg M (2005) Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. PLoS Genet 1e60:531-539

Axeisson T, Shavorskaya O, Lagercrantz U (2001) Multiple flowering time QTL within several *Brassica* species could be the result of duplicated copies of one ancestral gene. Genome 44:856-864

Bai Y, Huang CC, Hulst RGM van der, Dekens RG, Bonnema AB, and Lindhout WH (2003) QTLs for tomato powdery mildew resistance (*Oidium lycopersici*) in *Lycopersicon parviflorum* G1.1601 colocalize with two qualitative powdery mildew resistance genes. Mol Plant-Microbe Interact 16:169-176

Bentsink L, Yuan K, Koornneef M, Vreugdenhil D (2003) The genetics of phytate and phosphate accumulation in seeds and leaves of *Arabidopsis thaliana* using natural variation. Theor Appl Genet 106:1234-1243

Bohuon EJ, Ramsay LD, Craft JA, Arthur AE, Marshall DF, Lydiate DJ, Kearsey MJ (1998) The association of flowering time quantitative trait loci with duplicated regions and candidate loci in *Brassica oleracea*. Genetics 150:393-401

Bouis HE (2003) Micronutrient fortification of plants through plant breeding: can it improve nutrition in man at low cost? Proc Nutr Soc 62:403-411

Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422:433-438

Breseghello F and Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165-77

CAAS-IVF (Chinese Academy of Agricultural Science-Institute of Vegetables and Flowers) (2001) Annals of Chinese vegetable varieties. Chinese Agronomy Publishers (in Chinese)

Cao JS, Cao SC, Miao Y, Lu G (1997) Cladistic operational analysis and study on the evolution of Chinese cabbage groups (*Brassica campestris* L.). Acta Hortic Sin 24:35-42 (in Chinese)

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

Chen YP, Cao JS, Miao Y, Ye WZ (2000) Analysis of genetic polymorphism in vegetable crops of *Brassica campestris* by RAPD markers. J Zhejiang Univ (Agric Life Sci), 26:131-136 (in Chinese)

Choi SR, Kim JH, Han TH, Piao ZY, Plaha P, Koo DH, Bang JW, King GJ, Teakle GR, Allender CJ, Lydiate DJ, Beynon E, Osborn TC, Lim YP (2004) Joint meeting of the 14th Crucifer Genetics Workshop and the 4th ISHS Symposium on *Brassicas* 2004, Chungnam National University Daejeon, Korea October 24-28, pp 151

Chyi YS, Hoenecke ME, Sernyk JL (1992) A genetic map of restriction fragment length polymorphism loci for *Brassica rapa* (syn. *campestris*). Genome 25:746-757

Conner JK (2002) Genetic mechanisms of floral trait correlations in a natural population. Nature 420:407-410

Custers JBM, Cordewener JHG, Fiers MA, Maassen BTH, Van Lookeren Campagne MM, Liu CM (2001) Androgenesis in *Brassica*; a model system to study the initiation of plant embryogenesis, in: Bhojwani, S.S. and Soh, W.Y. (eds.), Current Trends in the Embryology of Angiosperms, Kluwer Acad Publ, Dordrecht, pp 451-470

Das S, Rajagopal J, Bhatia S, Srivastava PS, Lakshmikumaran M (1999) Assessment of

genetic variation within *Brassica campestris* cultivars using amplified fragment length polymorphism and random amplification of polymorphic DNA markers. J Biosci 24:433-440

De Candolle A (1886) Origin of cultivated plants, 2nd ed. Hanfner, New York pp 468

Demeke T, Adams RP, Chibbar R (1992) Potential taxonomic use of random amplified polymorphic DNA (RAPD): a case study in *Brassica*. Theor Appl Genet 84:990-994

Diederichsen A (2001) *Brassica rapa* group. In: Hanelt P (ed) Mansfelds encyclopedia of agricultural and horticultural crops (except ornamentals). Springer, Berlin Heidelberg, New York, pp 1446-1453

Dobson AJ (2002) An Introduction to Generalized Linear Models. Second Edition. Chapman & Hall/CRC, U.S.

Du S, Zhai F, Wang Y and Popkin BM (2000) Current methods for estimating dietary iron bioavailability do not work in China. J Nutr 30:193-198

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 164:1567-1587

Farnir F, Coppieters W, Arranz JJ, Berzi P, Cambisano N, Grisart B, Karim L, Marcq F, Moreau L, Mni M, Nezer C, Simon P, Vanmanshoven P, Wagenaar D, Georges M (2000) Extensive genome-wide linkage disequilibrium in cattle. Genome Res 10:220-227

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

Fujimoto R, Sugimura T, Nishio T (2006) Gene conversion from *SLG* to *SRK* resulting in self-compatibility in *Brassica rapa*. FEBS Lett 580:425-430

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

Gomez Campo C (1999) Developments in plant genetics and breeding, vol 4. Biology of *Brassica* coenospecies. Elsevier, Amsterdam Lausanne New York Oxford Shannon Singapore Tokyo

Guo JX, Zhou NY, Ma RC, Cao MQ (2002) Genetic diversity in *Brassica rapa* revealed by AFLP molecular markers. J Agric Biotechnol 10:138-143 (in Chinese)

Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in

higher plants: Present status and future prospects. Plant Mol Biol 57:461-485

Guttieri M, Bowen D, Dorsch J, Raboy V, Souza E (2004) Identification and characterization of a low phytic acid wheat. Crop Sci 44:418-424

Han HP, Sun RF, Li F, Zhang SF, Niu XK (2004) AFLP Markers linked to TuMV-resistance gene in Chinese Cabbage (*Brassica rapa* L. ssp. *pekinensis*). Sci Agric Sin 37(4):539-544 (in Chinese)

He YT, Chen BY, Fu TD, Li DR, TU JX (2003) Origins and evolution of *Brassica campestris* L. in China. Acta Genet Sin 30:1003-1012 (in Chinese)

He YT, Tu JX, Fu TD, Li DR, Chen BY (2002) Genetic diversity of germplasm resources of *Brassica campestris* L. in China by RAPD. Acta Agron Sin 28: 697-703 (in Chinese)

Hitz WD, Carlson TJ, Kerr PS, Sebastian SA (2002) Biochemical and molecular characterization of a mutation that confers decreased raffinosaccharide and phytic acid phenotype on soybean seeds. Plant Physiol 128:650-660

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

Huang JC, Corke H, Sun M (2002) Highly polymorphic AFLP markers as a complementary tool to ITS sequences in assessing genetic diversity and phylogenetic relationships of sweetpotato (*Ipomoea batatas* (L.) Lam.) and its wild relatives. Genet Resour Crop Evol 49:541-550

Jain A, Bhatia S, Banga SS, Prakash S, Lakshmikumaran M (1994) Potential use of random amplified polymorphic DNA (RAPD) technique to study the genetic diversity in Indian mustard (*Brassica juncea*) and its relationship to heterosis. Theor Appl Genet 88:116-122

James RW, Gnanasiri SP, Kevin AY (2000) Isolation of high seed inorganic P, low-phytate soybean mutants. Crop Sci 40:1601-1605

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

Jiang LJ (2001) Analysis of root swelling in turnip using AFLP markers. MsC thesis, Zhejiang University. pp 20-36 (in Chinese)

Jorde LB (2000) Linkage disequilibrium and the search for complex disease genes. Genome Res 10:1435-1444

Juenger T, Perez-Perez JM, Bernal S, Micol JL. (2005) Quantitative trait loci mapping of floral and leaf morphology traits in *Arabidopsis thaliana*: evidence for modular genetic architecture. Evol Dev 7:259-71

Keng H (1974) Economic plants of ancient north China as mentioned Shih Ching (Book of Poetry). Econ Bot 28:391-410

Kennard WC, Slocum MK, Figdore SS, Osborn TC (1994) Genetic analysis of morphological variation in *Brassica oleracea* using molecular markers. Theor Appl Genet 87:721-732

Kim JS, Chung TY, King GJ, Jin M, Yang TJ, Jin YM, Kim HI, Park BS (2006) A sequence-tagged linkage map of *Brassica rapa*. Genetics 174:29-39

Kole C, Teutonico R, Mengistu A, Williams PH, Osborn TC (1996) Molecular mapping of a locus controlling resistance to *Albugo candida* in *Brassica rapa*. Phytopath 86:367-369

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

Kole C, Quijada P, Michaels SD, Amasino RM, Osborn TC (2001) Evidence for homology of flowering-time genes *VFR2* from *Brassica rapa* and *FLC* from *Arabidopsis thaliana*. Theor Appl Genet 102:425-430

Kole C, Williams PH, Rimmer SR, Osborn TC (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

Koopman WJM (2005) Phylogenetic signal in AFLP data sets. Syst Biol 54:197-217

Koopman WJM, Zevenbergen MJ, Ven den Berg RG (2001) Species relationships in *Lactuca SL.* (*Lactuceae, Asteraceae*) inferred from AFLP fingerprints. Am J Bot 88:1881-1887

Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in *Arabidopsis thaliana.* Annu Rev Plant Biol 55:141-72

Kraakman ATW, Niks RE, van den Berg PMMM, Stam P, Van Eeuwijk FA (2004) Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. Genetics 168:435-446

Kraakman ATW, Martinez F, Mussiraliev B, van Eeuwijk FA, Niks RE (2006) Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. Mol Breed 17:41-58

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common

disease genes. Nat Genet 22:139-144

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

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Arizona State University, Tempe, Arizona

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

Lagercrantz U and Lydiate D (1996) Comparative genome mapping in *Brassica*. Genetics 144:1903-1910

Lagercrantz U, Kruskopf Österberg M, Lascoux M (2002) Sequence variation and haplotype structure at the putative flowering-time locus COL1 of *Brassica nigra*. Mol Biol Evol 19:1474-82

Lan TH and Paterson AH (2000) Comparative mapping of quantitative trait loci sculpting the curd of *Brassica oleracea*. Genetics 155:1927-1954.

Lander ES and Schork NJ (1994) Genetic dissection of complex traits. Science 265:2037-2048

Larson SR, Rutger JN, Young KA, Raboy V (2000) Isolation and genetic mapping of a non-lethal rice low phytic acid 1 mutation. Crop Sci 40:1397-1405

Larson SR, Young KA, Cook A, Blake TK, Raboy V (1998) Linkage mapping two mutations that reduce phytic acid content of barley grain. Theor Appl Genet 97:141-146

Li JW (1981) The origins and evolution of vegetable crops in China. Sci Agric Sin 14:90-95 (in Chinese)

Lickfett T, Matthaus B, Velasco L, Mollers C (1999) Seed yield, oil and phytate concentration in the seeds of two oilseed rape cultivars as affected by a different phosphorus supply. Eur J Agron 11:292-299

Liu HL (1984) Origin and evolution of genus *Brassica* oilseed rapes. Acta Agron Sin 10:9-18 (in Chinese)

Liu ZW, Fu TD, Tu JX, Chen BY (2005) Inheritance of seed colour and identification of

RAPD and AFLP markers linked to the seed colour gene in rapeseed (*Brassica napus* L.). Theor Appl Genet 110:303-10

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

Lu G, Cao JS, Chen H, Xiang X (2002) QTL mapping of some horticultural traits of Chinese cabbage. Sci Agric Sin 35:969-974 (in Chinese)

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

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

Ma G, Jin Y, Piao J, Kok F, Bonnema AB, Jacobsen E (2005) Phytate, calcium, iron, and zinc contents and their molar ratios in foods commonly consumed in China. J Agric Food Chem 53:10285-10290

Mackill DJ, Zhang Z, Redona ED, Colowit PM (1996) Level of polymorphism and genetic mapping of AFLP markers in rice. Genome 39:969-977

Mahmood T, Rahman MH, Stringam GR, Raney JP, Good AG (2005) Molecular markers for seed colour in *Brassica juncea.* Genome 48:755-760

Matsunomoto 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

McGrath JM and Quiros CF (1991) Inheritance of isozyme and RFLP markers in *Brassica campestris* and comparison with *B. oleracea*. Theo Appl Genet 82:668-673

Mizushima V (1980) Genome analysis in *Brassica* and allied genera. In: Tsunoda S, Hinata K, Gomez-Campo C (eds) *Brassica* crops and wild allies. Japan Scientific Societies Press, Tokyo, pp 89-105

Mollers C, Lickfett T, Matthaus B, Velasco L (1999) Influence of P-fertilizer on phytic acid content in seeds of *Brassica napus* L. and development of a NIRS calibration. Proceedings of the 10th International Rapeseed Congress, Canberra, Australia pp 26-29

Muangprom A and Osborn TC (2004) Characterization of a dwarf gene in *Brassica rapa*, including the identification of a candidate gene. Theor Appl Genet 108:1378-1384

Myburg AA and Remington, DL (2000) Protocol for high-throughput AFLP analysis using

LI-COR IR2 automated sequencer. North Carolina State Forest Biotechnology AFLP protocol. N C State Univ For Dep, Raleigh N.C.

Negi MS, Sabharwal V, Bhat SR; Lakshmikumaran M (2004) Utility of AFLP markers for the assessment of genetic diversity within *Brassica nigra* germplasm. Plant Breed 123:13-16

Nishioka M, Tamura K, Hayashi M, Fujimori Y, Ohkawa Y, Kuginuki Y, Harada K (2005) Mapping of QTL for bolting time in *Brassica rapa* (syn. *campestris*) under different environmental conditions. Breed Sci 55:127-133

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat Genet 30:190-193

Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Gladstone J, Goyal R, Jakobsson M, Kim S, Morozov Y, Padhukasahasram B, Plagnol V, Rosenberg NA, Shah C, Wall JD, Wang J, Zhao K, Kalbfleisch T, Schulz V, Kreitman M, Bergelson J (2005) The pattern of polymorphism in *Arabidopsis thaliana*. PLoS Biol 3:1289-1296

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

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

Nsengimana J, Baret P, Haley CS, Visscher PM (2004) Linkage disequilibrium in the domesticated pig. Genetics 166:1395-1404

Osborn TC, Kole C, Parkin IAP, Sharpe AG, Kuiper M, Lydiate DJ, Trick M (1997) Comparison of flowering time genes in *Brassica rapa*, *B. napus* and *Arabidopsis thaliana*. Genetics 146:1123-1129

Padilla G, Cartea ME, Rodrguez VM, Ordas A (2005) Genetic diversity in a germplasm collection of *Brassica rapa* subsp *rapa* L. from northwestern Spain. Euphytica 145:171-180

Padmaja KL, Arumugam N, Gupta V, Mukhopadhyay A, Sodhi YS, Pental D, Pradhan AK (2005) Mapping and tagging of seed coat colour and the identification of microsatellite markers for marker-assisted manipulation of the trait in *Brassica juncea*. Theor Appl Genet

111:8-14

Palaisa KA, Morgante M, Williams M, Rafalski A (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell 15:1795-1806

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

Payne RW and Arnold GM (2002) Genstat release 6.1. Reference Manual - Part 3: Procedure Library PL14. VSN International, Oxford, U.K.

Peng J, Slominsiki BA, Guenter W, Campbell LD, Xiong YZ (2001) The anti-nutritional factors in Chinese double-low rapeseed meal. J Chinese Cereals and Oil Association 16:6-10 (in Chinese)

Piao ZY, Deng YQ, Choi SR, Park YJ, Lim YP (2004) SCAR and CAPS mapping of *CRb*, a gene conferring resistance to *Plasmodiophora* brassicae in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Theor Appl Genet 108:1458-1465

Pilu R, Panzeri D, Gavazzi G, Rasmussen SK, Consonni G,Nielsen E (2003) Phenotypic, genetic and molecular characterization of a maize low phytic acid mutant (*lpa241*). Theor Appl Genet 107:980-987

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalaski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 3:225–238

Prakash S and Hinata K (1980) Taxonomy, cytogenetics and origin of crop *Brassica*, a review. Opera Bot 55:1-57

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

Qian W, Meng J, Li M, Frauen M, Sass O, Noack J, Jung C (2006) Introgression of genomic components from Chinese *Brassica rapa* contributes to widening the genetic diversity in rapeseed (*B. napus* L.), with emphasis on the evolution of Chinese rapeseed. Theor Appl Genet 113:49-54

Raboy V (2001) Seeds for a better future: “low phytate” grains help to overcome malnutrition and reduce pollution. Trends Plant Sci 6:458-462

Raboy V (2003) myo-Inositol-1,2,3,4,5,6-hexakisphosphate. Phytochemistry 64: 1033-1043

Raboy V and Gerbasi P (1996) Genetics of myo-inositol phosphate synthesis and accumulation. In: Biswas B, Biswas S (eds) Myoinositol phosphates, phosphoinositides and signal transduction. Plenum, New York, pp 257-285

Raboy V, Gerbasi P, Young K, Stoneberg S, Pickett G, Bauman A, Murthy P, Sheridan W, Ertl D (2000) Origin and seed phenotype of maize low phytic acid 1-1 and low phytic acid 2-1. Plant Physiol 124:355-368

Raboy V, Young KA, Dorsch JA, Cook A (2001) Genetics and breeding of seed phosphorus and phytic acid. J Plant Physiol 158:489-497

Rae AM, Howell EC, Kearsey MJ (1999) More QTL for flowering time revealed by substitution lines in *Brassica oleracea*. Heredity 83:586-596

Rasmussen SK and Hatzack F (1998) Identification of two low-phytate barley (*Hordeum vulgare* L.) grain mutants by TLC and genetic analysis. Hereditas 129:107-112

Reiner H, Holzner W, Ebermann R (1995) The development of turnip-type and oilseed-type *Brassica rapa* crops from the wild type in Europe-An overview of botanical, historical and linguistic facts.-Rapeseed Today and Tomorrow 4:1066-1069

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

Salathia N, Davis SJ, Lynn JR, Michaels SD, Amasino RM, Millar AJ (2006) *FLOWERING LOCUS C* -dependent and -independent regulation of the circadian clock by the autonomous and vernalization pathways. BMC Plant Biol 6:10

Schranz ME, Quijada P, Sung SB, Lukens L, Amasino RM, Osborn TC (2002) Characterization and effects of the replicated time gene *FLC* in *Brassica rapa* Genetics 162:1457-1468

Schranz ME, Lysak MA, Mitchell-Olds T (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends Plant Sci 11:535-542

Sebastian RL, Kearsey MJ, King GJ (2002) Identification of quantitative trait loci controlling developmental characteristics of *Brassica oleracea* L. Theor Appl Genet 104:601-609

Shi J, Wang H, Hazebroek J, Ertl DS, Harp T (2005) The maize low-phytic acid 3 encodes a myo-inositol kinase that plays a role in phytic acid biosynthesis in developing seeds. Plant

J 42:708-719

Shi J, Wang H, Wu Y, Hazebroek J, Meeley RB, Ertl DS (2003) The maize low-phytic acid mutant *lpa2* is caused by mutation in an inositol phosphate kinase gene. Plant Physiol 131:507-515

Shiba H, Kenmochi M, Sugihara M, Iwano M, Kawasaki S, Suzuki G, Watanabe M, Isogai A, Takayama S (2004) Genomic organization of S-locus region of *Brassica*. Biosci Biotechnol Biochem 67:622-626

Shukla S, Vantoai TT, Pratt RC (2004) Expression and nucleotide sequence of an INS (3) P1 synthase gene associated with low-phytate kernels in maize (*Zea mays* L.). J Agric Food Chem 52:4565-4570

Siemonsma JS, Piluek K (1993) Plant resources of south-east Asia, No. 8. Vegetables. Pudoc, Wageningen, pp 121-134

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

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

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

Song KM, Osborn TC, Williams PH (1988b) *Brassica* taxonomy based on nuclear restriction fragment length polymorphism (RFLPs). 2. Preliminary analysis of sub-species within *B. rapa* (syn. *campestris*) and *B. oleracea*. Theor Appl Genet 76:593-600

Song KM, Osborn TC, Williams PH (1990) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 3. Genome relationships in *Brassica* and related genera and the origin of *B. oleracea* and *B. rapa* (syn. *campestris*). Theor Appl Genet 76:497-506

Song KM, Suzuki JY, Slocum MK, Williams PH Osborn TC (1991) A linkage map of *Brassica rapa* base on restriction fragment length polymorphism loci. Theor Appl Genet 82:296-304

Srivastava A, Gupta V, Pental D, Pradhan AK (2001) AFLP-based genetic diversity

assessment amongst agronomically important natural and some newly synthesized lines of *Brassica juncea*. Theor Appl Genet 102:193-199

Suwabe K, Iketani H, Nunome T, Kage T, Hirai M (2002) Isolation and characterization of microsatellites in *Brassica rapa* L. Theor Appl Genet 104:1092-1098

Suwabe K, Tsukazaki, Hiketani H, Hatakeyama K, Fujimura M, Nunome T, Fukuoka H, Matsumoto S, Hirai M (2003) Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. Theor Appl Genet 107:997-1002

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* 2004, Chungnam National University Daejeon, Korea October 24-28, pp 143

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

Tadege M, Sheldon CC, Helliwell CA, Stoutjesdijk P, Dennis ES, James Peacock W (2001) Control of flowering time by *FLC* orthologues in *Brassica napus*. Plant J. 28:545-53

Tan QM (1979) The origin, distribution and domestication of Chinese cabbage. Agric Sci China 4:68-75 (in Chinese)

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

Tanhuanpää PK, Vilkki JP, Vilkki HJ (1996a) A linkage map of spring turnip rape based on RFLP and RAPD markers. Agric Food Sci Finl 5:209-217

Tanhuanpää PK, Vilkke JP, Vikki HJ (1996b) Mapping a QTL for oleic acid concentration in spring turnip rape (*Brassica rapa* ssp *oleifera*). Theor Appl Genet 92:952-956

Tenaillon MI, Sawkins MC, Long AD, Gaut RL, Doebley JF, Gaut BS (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp *mays* L.). Proc. Natl Acad Sci USA 98:9161-9166

Teutonico RA and Osborn TC (1994) Mapping of RFLP and quantitative 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 RA and Osborn TC (1995) Mapping loci controlling vernalization requirement in *Brassica rapa*. Theor Appl Genet 91:1279-1283

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

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 JG, 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 Ooijen JW and Voorrips RE (2001) JoinMap® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands

Van Ooijen JW (1992) Accuracy of mapping quantitative trait loci in autogamous species. Theor Appl Genet 84:803-811

Varshney RK, Graner A, Sorrells ME (2005) Genomics-assisted breeding for crop improvement. Trends Plant Sci 10:621-30

Vaughan JG (1977) A multidisciplinary study of the taxonomy and origin of *Brassica* crops. Bio Sci 27:35-40

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

Vos P, Hogers R, Bleeker M, Rijan 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

Vreugdenhil D, Aarts MGM, Koornneef M, Nelissen H, Ernst WHO (2004) Natural variation and QTL analysis for cationic mineral content in seeds of *Arabidopsis thaliana*. Plant Cell Environ 27:828-839

Wang HZ (2004) Strategy for rapeseed genetic improvement in China in the coming fifteen years. Chin J Oil Crop Sci 26:98-101 (in Chinese)

Wang M, Zhang FL, Meng XD, Liu XC, Zhao XY, Fan ZC (2004) A linkage map construction for Chinese cabbage based on AFLP markers using DH population. Acta Agric Boreali-Sin 19:28-33 (in Chinese)

Warwick SI and Francis A (1994) Guide to the Wild Germplasm of *Brassica* and Allied Crops, Part V: Life History and Geographical Data for Wild Species in the Tribe Brassiceae (Cruciferae). Technical Bulletin 1994, Centre for Land and Biological Resources Research, Agriculture and Agri-Food Canada.

Warwick SI, Séguin-Swartz G, Beckie H, Légère A, Nair H, Simard MJ, Mason LP, Braun L, Zhu B, Stewart N (2002) Transgenics in the environment: a canola case study Thirteenth crucifer genetics workshop. University California, Davis

Welch RM and Graham RD (2004) Breeding for micronutrients in staple food crops from a human nutrition perspective J Exp Bot 55:353-64

White PJ and Broadley MR (2005) Biofortifying crops with essential mineral elements. Trends Plant Sci 10:586-93

Wilcox JR, Premachandra PS, Young KA, Raboy V (2000) Isolation of high inorganic phosphorus, low-phytate soybean mutants. Crop Sci 40:1601-1605

Williams PH and Hill CB (1986) Rapid cycling populations of *Brassica*. Science 232:1385-1389

Yang TJ, Kim JS, Kwon SJ, Lim KB, Choi BS, Kim JA, Jin M, Park JY, Lim MH, Kim HI, Lim YP, Kang JJ, Hong JH, Kim CB, Bhak J, Bancroft I, Park BS (2006) Sequence-Level analysis of the diploidization process in the triplicated *FLOWERING LOCUS C* region of *Brassica rapa*. The Plant Cell 18:1339-1347

Yasuhisa K, Hidetoshi A, Mamiko Y (1997) RAPD marker linked to a clubroot-resistance locus in *Brassica rapa* L. Euphytica 98:149-154

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

Yu SC, Wang YJ, Zheng XY (2003) A genetic linkage map of *Brassica campestris* L. ssp. *pekinensis* (syn. *B. rapa* L. ssp. *pekinensis*). Agric Sci China 2:49-55 (in Chinese)

Yu SC, Wang YJ, Zheng XY (2003) Mapping and analysis QTL controlling heat tolerance in *Brassica campestris* L. ssp. *pekinensis*. Acta Hortic Sin 30:417-420 (in Chinese)

YU SC, Wang YJ, Zheng XY (2003) Mapping and analysis QTL controlling some morphological traits in Chinese cabbage (*Brassica campestris* L. ssp. *pekinensis*). Acta Genet Sin 30:1153-1160 (in Chinese)

Zhang FL, Aoki S, Takahata Y (2003) RAPD markers linked to microspore embryogenic ability in *Brassica* crops. Euphytica 131:207-213

Zhang LG, Wang M, Chen H (2000) Construction of RAPDs molecular genetic map of Chinese cabbage. Acta Bot Sin 42(5):485-489 (in Chinese)

Zhang XW, Gao MQ, Geng JF, Yuan YX, Jiang WS, Li GY (2001) A new cabbage F1 hybrid by isolated microspore culture - 'Yusheng 1'. Acta Hortic Sin 28:577 (in Chinese)

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

Zhu Y and Zheng Y (2001) Atlas of the traditional vegetables in China, Zhejiang Science and Technology Publishing House, China

Summary

In this thesis molecular genetic studies on *Brassica rapa* are described based on a collection of 256 accessions and 6 segregating populations. In chapter 2 and 3 the genetic variation and population structure are characterized in a set of genotypes from different geographical origins representing the various cultivar types. The association of traits with AFLP markers is analyzed in chapter 3, and the development of various mapping populations, linkage map construction and QTL analysis of phytate and diverse morphological traits are described in chapter 4 and 5.

Genetic diversity in the primary gene pool of *B. rapa* was assessed by AFLP fingerprinting (chapter 2). The relationship between accessions was revealed by the dendrogram, indicating that different cultivar groups are often region (East Asia vs. Europe) specific, emerging independently in the different geographic regions. The UPGMA dendrogram (chapter 2) and population structure (chapter 3) showed that grouping based on markers coincided often with morphotypes (cultivar groups) within a geographic region. However this clustering was not supported by significant bootstrap values, suggesting that only a few genes are involved in causing the large differences in morphologies. A number of markers differed between subpopulations thereby providing an example of linkage disequilibrium (LD) due to population structure.

Diversity collections can be used to survey the allele frequencies and to explore associations with traits (chapter 3). The STRUCTURE program indicated an association between population structure and trait values even for traits such as phytate and phosphate levels. The population structure should therefore be considered in association mapping studies in *B. rapa*. Markers that show association after correction for substructure can reliably be interpreted as being linked. Using the Composite Interval association mapping procedure, 54 markers were found significantly associated with at least one of the 9 different traits that were analyzed, 16 of which had known map positions. The observation that some of the marker-trait associations were confirmed by QTL analysis indicated that association mapping allows the detection of linkage with moderately frequent alleles.

A multiple population approach was applied in QTL analysis for phytate and morphological traits, in which the populations were developed from wide crosses and alleles of multiple parental accessions can be evaluated, revealing a large number of genomic regions harboring

allelic variation for traits of interest. For phytate and phosphate contents in seeds and leaves, and 20 morphological traits involved in flowering time, seed traits, plant height, leaf traits and turnip formation we identified a total of 54 QTL positions. Co-location of QTL for different traits was found in many cases, which might suggest pleiotropy or tight linkage. The analysis of co-location of QTL for phytate/phosphate and parental allelic effects suggested that these QTL control total phosphorus levels or are specific for the phytate pathway in *B. rapa.* Principal component analysis and co-localization of QTL for morphological traits indicated that some components of the genetic control of certain leaf and seed-related traits, flowering time and turnip formation might be based on the same genes, indicating a common genetic regulation of correlated traits underlying these QTL. The comparative analyses of QTL for certain traits, such as phytate and phosphate contents and flowering time, between the 6 populations and with *Arabidopsis thaliana* or other related *Brassica* species are discussed.

Samenvatting

In dit proefschrift wordt moleculair genetisch onderzoek beschreven aan een collectie van 256 *Brassica rapa* accessies en zes splitsende populaties. In hoofdstuk twee is de genetische variatie gekarakteriseerd van een collectie genotypes afkomstig uit verschillende gebieden, welke vaak verschillen in morfotypes, ook wel cultivar groepen genoemd, vertegenwoordigen.. In hoofdstuk drie wordt de populatie structuur van de collectie bepaald en de associatie van AFLP merkers met een aantal eigenschappen gepresenteerd. De ontwikkeling van splitsende populaties, de constructie van genetische kaarten en QTL (quantitative trait loci) analyse voor fytaat en fosfaat gehaltes en een aantal morfologische eigenschappen worden in de hoofdstukken vier en vijf beschreven.

Met behulp van AFLP fingerprints is de genetische diversiteit in twee *B. rapa* collecties bepaald (hoofdstuk 2) en gevisualiseerd met behulp van een dendrogram. Een interessante conclusie was, dat de accessies groepeerden naar geografische herkomst (Oost Azie vs Europa vs India), en dat cultivar types onafhankelijk zijn ontstaan in die regio's. Zowel de UPGMA boom (hoofdstuk 2) als de populatie structuur (bepaald met het programma STRUCTURE, zie hoofdstuk 3) gaven aan, dat de op moleculaire merkers gebaseerde groepen samenvielen met de verschillende morfotypes binnen een geografisch gebied. Het feit dat deze clusters niet gesteund worden door hoge bootstrap waardes, suggereert dat slechts een klein aantal genen verantwoordelijk is voor de grote morfologische verschillen (vergelijk Chinese kool met oliezaad of meiraapjes). Een aanzienlijk deel van de merkers had verschillende allel frequenties in de sub-groepen, en illustreerden dat associatie van merkers (linkage desequilibrium (LD)) vaak een gevolg is van populatie structuur en niet van genetische koppeling. In genenbank collecties kunnen allel frequenties bepaald worden van merkers, om de associatie van deze merkers met de onderzochte eigenschappen te detecteren. Echer in de *B. rapa* collectie was er ook een associatie tussen de subpopulaties zoals bepaald met STRUCTURE en de waardes van eigenschappen zoals bloeitijd, maar ook de fytaat en fosfaat gehaltes, waarop nooit direct is geselecteerd in veredelings-programma's. Om deze reden moet de populatie structuur meegenomen worden als parameter in associatie studies in *B. rapa*. Merkers welke nog geassocieerd zijn met eigenschappen na correctie voor populatie structuur zijn met een grotere betrouwbaarheid gekoppeld met deze eigenschappen. Met behulp van de karterings module 'Composite Interval Association Mapping'zijn 54 merkers

geïdentificeerd welke significant gekoppeld zijn met de negen geanalyseerde eigenschappen. Zestien van deze merkers hebben een kaart positie en het feit dat een aantal van deze merkers ook gekoppeld was met QTL gedetecteerd in de DH en F2 populaties, toont aan dat association mapping een goed alternatief is voor QTL analyse.

Voor QTL analyses voor fytaat-, en fosfaat gehalte en voor morfologische eigenschappen zijn een aantal populaties ontwikkeld uit kruisingen tussen verschillende morfotypes, zodat de effecten van allelen uit diverse ouders geëvalueerd konden worden. Dit leverde een groot aantal genomische gebieden op met alellische variatie voor interessante eigenschappen. In totaal zijn 54 QTL geïdentificeerd voor fytaat- en fosfaatgehalte in blad en zaad en voor 20 morfologische eigenschappen betrokken bij bloeitijd, zaad en blad kenmerken, plant hoogte, en turnip vorming. In een aantal gevallen co-localiseerden QTL voor verschillende eigenschappen wat duidt op pleiotropie of genetische koppeling. Co-lokalisatie van QTL voor zowel fytaat- als fosfaatgehalte en de richting van de ouder allelen suggereerde dat deze QTL of betrokken zijn bij de fosfaat opname (totaal fosforgehalte) of specifiek betrokken zijn bij de fytaat biosynthese. Principale componenten analyse en de co-lokalisatie van QTL voor morfologische kenmerken zoals blad en zaad gerelateerde eigenschappen, bloeitijd in verschillende seizoenen en knolvorming, wijzen erop dat de genetische regulatie gebaseerd is op dezelfde genen (met kwantitatieve effecten), ofwel op een gemeenschappelijke regulatie van de gecorreleerde eigenschappen. Voor een aantal eigenschappen, zoals fytaat- en fosfaatgehaltes en bloeitijd, zijn de QTL posities gevonden in zes *B. rapa* populaties en de *B. rapa* collectie vergeleken met de corresponderende genomische regio's in verwante *Brassica* soorten en *Arabidopsis thaliana*. Het nut van "vergelijkende genomics" (comparative genomics) en genoom syntheny voor het identificeren en kloneren van kandidaat-genen wordt besproken.

芸薹类作物植酸含量和形态性状的遗传分析

赵建军

芸薹属植物与模式植物拟南芥（*Arabidopsis thaliana*）同属十字花科。芸薹属植物包括三个基本的二倍体种芸薹（*Brassica rapa*: AA, n=10）、甘蓝（*Brassica oleracea*: CC, n=9）、黑芥（*Brassica nigra*: BB, n=8）和三个复合种甘蓝型油菜（*Brassica napus*: AACC, n=19）、芥菜（*Brassica juncea*: AABB, n=18）、埃塞俄比亚芥（*Brassica carinata*: BBCC, n=17）。芸薹类作物（*B. rapa*）是重要的蔬菜和油用作物，在长期的自然选择和人工选择条件下，形成了多个亚种、变种及生态型，种内遗传变异广泛，品种类型丰富，亲缘关系复杂，在起源、分类研究方面存在争议。针对以上状况，有必要从分子水平上系统探讨该物种的遗传变异，了解其复杂的遗传背景。这不仅为澄清物种起源演化关系、确立合理的分类地位奠定基础，还可以为有效地发掘和利用丰富的基因资源提供理论依据。

在构建遗传图谱的基础上，芸薹类作物重要性状的基因定位研究取得了一定进展。然而，QTL（quantitative trait loci）的定位研究还不全面，许多形态生理性状的 QTL 没有完全被检测出来,不能完全满足育种的需要。另外，随着人民生活水平的提高，蔬菜的营养品质或微量营养成分日益受到关注。例如，植酸等抗营养因子的存在一定程度上限制了一些矿物营养成分的有效利用。研究植酸等抗营养因子以及微量营养成分在二倍体芸薹类作物中的积累代谢机制和遗传途径，将有助于了解该物种中微量营养成分在人类日常膳食中的作用，以提高营养元素的有效吸收。同时为进一步研究油菜籽饼粕的营养成分奠定理论基础。

基于上述情况，本研究以 256 份不同类型芸薹类种质为材料，从 DNA 水平揭示其遗传多样性和群体结构，探讨不同基因型间的遗传关系（第二章；第三章）；评价叶片和种子中有效磷和植酸含量的自然变异特点，为开展有效磷和植酸在芸薹类作物中积累代谢和遗传研究提供信息，并进一步探讨关联作图（Association Mapping）在该物种中的应用（第三章）；创建遗传信息丰富的遗传研究分离群体，构建芸薹类作物遗传连锁图谱框架图。开展重要农艺性状、有效磷和植酸含量的 QTL 定位与遗传效应分析，为芸薹类作物基因组学研究和将来分子标记辅助选择育种提供理论依据（第四章；第五章）。

利用荧光 AFLP 技术对芸薹类作物核心种质库多样性分析的结果表明：芸薹类作物

主要分为由不同亚类群聚类组成的两大类群－类群Ⅰ、类群Ⅱ。材料起源与聚合分类（第二章）和群体结构分析（第三章）结果基本一致，芸薹类作物不同的形态类型通常与同一地理起源（东亚与欧洲）的其它形态类型密切相关，这种相关程度甚至大于与不同来源的相同形态类型之间的关系，暗示芸薹类作物在两个地区分别具有独立的起源和（或）经历了长期独立的驯化和育种选择。然而，低的 bootstrap 值说明各个亚类群内与亚类群间的遗传变异幅度相当，可能只有少数基因参与了明显不同形态性状的表达调控。不同亚类群间具有大量不同分子标记的事实提供了由于群体结构造成连锁不平衡（linkage disequilibrium, LD）的一个例证。

现有的自然群体可以用来研究基因位点频率，并确定标记与目标性状的关联关系（第三章）。STRUCTURE 程序分析表明，群体结构与包含植酸和无机磷含量在内的表型性状之间存在关联关系，是芸薹类作物关联作图中的重要考虑因素。理论上，通过校正群体结构进行“性状-标记”关联分析，检测到的连锁关系是遗传连锁产生的。采用复合区间作图方法（Composite Interval Mapping Procedure），共检测到与至少 1 个目标性状（共 9 个）相关联的 54 个分子标记。其中，已经确定图谱位置的为 16 个标记。通过传统 QTL 分析方法，论文进一步验证了“性状-标记”关联分析的部分结果。考虑适当的基因位点频率，“性状-标记”关联分析是可行的。

采用不同来源的多个栽培亚种类型为材料，进行广泛杂交并制备大批 F2 和 DH（Double Haploid）分离群体。应用多群体（6 个）QTL 分析策略，对芸薹类作物种子和叶片的植酸和有效磷磷含量以及 20 个形态性状（开花时间、种子相关性状、株高、叶片性状和芜菁直根膨大性状）进行了 QTL 定位及遗传效应分析。在 8 个连锁群上，共检测到 54 个 QTL。部分不同性状的 QTL 分布在同一连锁群的相近区段，在一定程度上反映了这些性状的遗传相关性，相关关系的基础是不同的 QTL 的紧密连锁或一因多效。控制植酸/有效磷的 QTL 在连锁群上的分布相近性以及效应方向分析表明，芸薹类作物中检测出的 QTL 与总磷水平或植酸生物合成途径特异相关。结合目标形态性状 QTL 的分布相近性与主成分分析，我们发现某些相关性状具有共同的遗传调控方式，相同基因可能同时参与多个性状（如开花时间和芜菁直根膨大性状；某些叶片性状和种子相关性状）的表达调控。最后利用 6 个芸薹类作物群体检测 QTL，并进一步比较讨论了与拟南芥或其它相关芸薹属物种相同性状（如植酸和有效磷含量，开花时间）QTL。

Curriculum Vitae

Jianjun Zhao was born on the 8th March, 1971 in Hebei, China. He studied at Hebei Agricultural University (China) and obtained his BSc degree with Agriculture as major in 1993 and his MSc degree with Crop Genetics and Breeding as major in 1998. After his studies he worked as a wheat breeder in Hebei Agricultural University. He started his PhD program in the Chinese Academy of Agricultural Sciences (CAAS) in 2001. In 2002, he got the chance to participate in the Joint Sandwich PhD program between CAAS and Wageningen University (WU), funded by WU INREF funds and the Asian Facility Program. He will defend his PhD thesis on 22 January 2007 in Wageningen to obtain his PhD degree.

Publications

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

Zhao J, Paulo MJ, Jamar D, Lou P, Van Eeuwijk F, Bonnema AB, Koornneef M, Vreugdenhil D (2006) Genetic variation for phytate, phosphate and several agronomic traits in *Brassica rapa*: an association mapping approach. Submitted.

Zhao J, Jamar D, Lou P, Song X, Li Y, Wu J, Wang X, Bonnema AB, Koornneef M, Vreugdenhil D (2007) QTL analysis of phytate and phosphate content in seeds and leaves of *Brassica rapa*. To be submitted.

Lou P*****, **Zhao J***, Song X, Pino Del Carpio D, Shen S, Vreugdenhil D, Wang X, Koornneef M, Bonnema AB (2007) Mapping quantitative trait loci for morphological traits in multiple populations of *Brassica rapa*. To be submitted.

Lou P, **Zhao J**, He H, Pino Del Carpio D, Verkerk R, Custers J, Koornneef M, Bonnema AB (2006) Quantitative Trait Loci for Glucosinolate Accumulation in *Brassica rapa* Leaves. Submitted.

Wu J, **Zhao J**, Song X, Li Y, Li X, Sun R, Koornneef M, Aarts M, Wang X (2007) Mapping QTLs for mineral accumulation and shoot dry biomass yield under different Zn nutritional conditions in heading Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). Submitted.

Zhang X, Wu J, **Zhao J**, Song X, Li Y, Zhang Y, Xu D, Sun R, Yuan Y, Xie C, Wang X (2006) Identification of QTLs related to bolting in *Brassica rapa* ssp. *pekinensis*. Agricultural

Sciences in China 5:265-271

Zhao J, Xu Z, Fang X (2003) A review of progress in study on crop breeding for phytic acid. Journal of Oil Crops 25:94-98 (in Chinese)

Zhao J, Xu Z, Fang X, Luo L (2002) The preliminary analysis on the variation for the inorganic phosphorus contents in seeds of rapeseed (*Brassica napus*). Journal of Oil Crops 24:79-81(in Chinese)

Acknowledgements

It is my greatest pleasure to achieve my PhD degree, which is one of my dreams in life. During the past four years of this study, many people supported and encouraged me to complete the thesis. I would like to appreciate all those whom were involved in my studying and will remember them forever.

Firstly, I would like to express my special gratitude to my promotor, Prof. Maarten Koornneef. Your excellent guidance, decisive support and continuous encouragement have helped me to complete this study. From the first day of my coming to Wageningen in 2002, you supervised me kindly and introduced me to the field of scientific research. I enjoyed and learned a lot from each discussion with you. You always give clear and valuable comments and suggestions for the design of experiments and writing of each of the chapters. I learned from you not only scientific knowledge but also scientific attitude, which will promote me to be a real scientist in the future. Deep gratitude is also for you and your family for the kind hospitality and friendship during my stay in Wageningen.

I am deeply indebted to my co-promotor Dr. Guusje Bonnema who gave me the opportunity to join the PhD program between the Chinese Academy Agricultural Sciences (CAAS) and Wageningen University (WU). Your excellent guidance and patience in explaining the basic knowledge of experiments, the usage of scientific terms and the making of presentations have helped me to understand correctly and complete timely my project. I will never forget that you show me how to perform PCR analysis, how to phenotype plants in greenhouse, how to present the results and how to communicate with other colleagues. You did your best to help me to apply fellowship programs of the Netherlands ministry of Agriculture (IAC) and Nuffic, which have been very helpful for me to finish the thesis. During the final parts of the thesis, you spend much time even your holiday time to look through each chapter, give nice suggestions and make the Dutch summary. I will benefit from all your help in my future research career. I am also thankful for your effort to arrange my visa applications and make me live comfortably in Wageningen. My family and I would like to express our highly thanks to you and your family.

I would like to express my sincere gratitude to Dr. Dick Vreugdenhil, my co-promotor, for

your support, advice and guidance during the period of my Ph.D. When I first arrived in Wageningen, you picked me up from the airport, introduced the Netherlands to me and helped me to be familiar with a new environment soon. I highly appreciate your great efforts for arranging plant materials, designing experiments and improving the thesis writing. During the laboratory work in plant physiology, you always gave valuable comments and positive encouragements. I would like to give my best wishes to you and your family.

Respect and thanks should also go to my Chinese supervisor Prof. Zeyong Xu in CAAS. You always supported and encouraged me to participate and finish the joint PhD program. In my mind, you are like a father teaching me how to do research and behave well. I would like to thank Prof. Rifei Sun, whom supported me to start my sandwich PhD study and gave assistance during my working in China. Prof. Xiaowu Wang, you offered me good experimental environments and intensive supervision when I was in your lab. Your ideas to solve scientific questions are always clear and efficient, which will affect me in my future research.

Dr. Ping Lou, my best friend and colleague, you always give help for data analysis, experimental techniques, research information and for the design of the thesis cover picture. Dr. Joao Paulo, I am very glad with your kind help for the association analyses and great contribution to chapter 3. Prof. Fred van Eeuwijk, I also express my sincere gratitude for your useful suggestions and comments on chapter 3. Dr. Jaap Vromans, thanks for your valuable help to deal with AFLP data of chapter 2; I enjoyed a lot the cooperation that we had.

I am very thankful Prof. Evert Jacobsen for your encouragement and suggestions for my studying, writing and presentation skills. You always give solutions to problems. Your optimistic attitude on understanding life will affect me. Thanks also are expressed to Dr Mark Aarts for your valuable help, discussions and suggestions. Special thanks should go to Dr. Jan Custers for your kind help for microspore culture. I am also grateful to Prof. Richard Visser for your kind support of my stay finalizing the thesis in plant breeding and my last one month extension. I would also like to thank Dr. Roeland Voorrips for suggestions on improving genetic maps.

Many thanks should be delivered to people from CAAS for making efforts to initiate the

cooperation between CAAS and WU, especially Prof. Dongyu Qu and Prof. Qingguang Lu.

I would like to thank Dunia Pino Del Carpio, Nurmi P. D. Pangesti, Jiangling Xiong, Xiaofei Song for phenotypic scoring of turnip characteristics, testing of SSR markers and constructing the genetic maps. Thanks also are expressed to staff in Plant Science Experimental Centre of Wageningen University, the Netherlands, and Yanling Liu in the Institute of Vegetables and Flowers (Beijing), Chinese Academy of Agriculture Science, China, for taking care of the plants and assistances.

I would also like to express my deep gratitude to my colleagues in the laboratories of genetics, plant physiology and plant breeding for their very kind help. Corrie and Diaan, you are so kind to help me to know lab rules, create Double Haploid populations, test phytate and phosphate levels and solve the technical problems. Petra and Fien, I learned AFLP technique from you.

I would like to thank other members of plant breeding, specially, Johan Bucher, Marcos Malosetti and Ralph van Berloo for your useful discussion during lab meeting or personal communications. I would like also to thank other members of genetic laboratory, Joost Keurentjes, Judith van de Mortel, Artak Ghandilyan, Wessel van Leeuwen, Sangita Talukdar, Ana Assuncao, Hedayat Bagheri and Liana Tsiatsiani for your kind support and valuable discussion. Very special thanks to Annie Marchal, Letty Dijker and Mariame Gada, the secretaries of Plant Breeding, to Aafke van der Kooi and Corrie Eekelder, the secretaries of Genetics, and to Trees van Nieuwamerongen and Wytske Nijenhuis, the secretaries of Plant Physiology, for your administrative help.

I would like to thank Hans, Farhad, Pauel and Alireza who shared a same office with me in plant physiology and in plant breeding. We had nice conversations and I received your kind help during my stay in the Netherlands.

And I appreciate the time spent together with my Chinese friends in WU and CAAS. Thanks all of you: Shujun Zhou, Xi Wan, Zhanguo Bai, Jungen Kang, Hongju He, Miqia Wang, Jian Wu, Jun Guo, Jun Du, Chengwei Li, Ningwen Zhang, Ke Lin, Yuling Bai, Zifu Yan, Junming Li, Zhongkui Sun, Jifeng Tang, Aiguo Zhu, Yehong Liang, Yongyan Qi, Liying Yan, Ying Li, Guangcun Li, Haishan Li, Feng Ji, Yanguo Zhang, Baojun Yang, Donghui Xu, Junrong Huang,

Xiaobin Wang, Dongsheng Yin, Daidi Che, Wenbin Wang, Jiayou Zhong and Qi Jing.

This research was financed by WU INREF funds, the Asian Facility Program, the fellowship program of the Netherlands ministry of Agriculture and the fellowship from an exchange student program (Huygens) between Chinese Scholarship Council and Nuffic in the Netherlands.

I will always keep in my mind a lot of support that I received from my colleagues in my home university, Hebei agricultural university, especially Prof. Shuxing Shen and Prof. Yuxing Zhang, who encouraged me to finish my study abroad.

Finally, my deepest gratitude goes to my mother, my father, my mother-in-law, my father-in-law, my wife Bi Jie, my son Zhao Penghao, my brother Zhao Jianqing and my sister Zhao Jianmin for their love and support.

Jianjun Zhao
Wageningen, the Netherlands
January, 2007

Education Statement of the Graduate School

Experimental Plant Sciences

Issued to: **Jianjun Zhao**
Date: **22 January 2007**
Group: **Laboratory of Genetics, Wageningen University**

1) Start-up phase	*date*
► **First presentation of your project**	
Screening and constructing of the molecular marker linkage map for the low phytic acid trait in *Brassica rapa*	January 16, 2003
► **Writing or rewriting a project proposal**	
► **Writing a review or book chapter**	
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	1.5 credits*

2) Scientific Exposure	*date*
► **EPS PhD student days**	
PhD day, Utrecht	March 27, 2003
PhD day, Wageningen	September 19, 2006
PhD day, Paris	June 09, 2006
► **EPS theme symposia**	
EPS Theme 4 symposium, Wageningen	December 20, 2002
EPS Theme 3 symposium, Utrecht	November 24, 2005
► **NWO Lunteren days and other National Platforms**	
NWO Lunteren days	April 07 -0 8, 2003
Vegetable molecular breeding symposium in China	June 17 - 18, 2004
NWO Lunteren days	April 03 - 04, 2006
► **Seminars (series), workshops and symposia**	
Seminar Frontiers in Plant Science, Wageningen (2x)	2003
INREF workshop on project related to micronutrients, Wageningen	November 13 - 16, 2002
WU-CAAS Autumn Workshop, China	November 10 - 15, 2003
Seminar, biulding 107, Wageningen	October 4, 2005
Flying seminar, Wageningen	October 24, 2005
Symposium on Quality: from Soil to Healthy People, Wageningen	November 16, 2006
► **Seminar plus**	
► **International symposia and congresses**	
The 14th Crucifer Genetics Workshop and the 4th ISHS Symposium on Brassicas, South Korea	October 24 - 28, 2004
The 14th Crucifer Genetics Workshop, the Netherlands	October 01 - 04, 2006
► **Presentations**	
Oral presentation in workshop on micronutrinents, Wageningen	November 14, 2002
Oral presentation in Symposium on Quality	November 16, 2006
Oral presentation in NWO Lunteren days	April 03 - 04, 2006
Poster presentation in Crucifer Genetics Workshop, South Korea	October 24 - 28, 2004
Poster presentation in autumn school (in CAAS), China	November 10 - 15, 2003
Poster presentation in Crucifer Genetics Workshop, the Netherlands	October 01 - 04, 2006
► **IAB interview**	September 18, 2006
► **Excursions**	
Subtotal Scientific Exposure	17.3 credits*

3) In-Depth Studies	*date*
► **EPS courses or other PhD courses**	
Summer school 'Natural Variation', Wageningen	April 22 - 25, 2003
Techniques for Writing and Presenting a Scientific Paper	February, 2006
► **Journal club**	
Literature discussion in genetic group meeting	2002-2003
► **Individual research training**	
Biological chemical experiments training, China	January, 2002
Microspore culture training in Bioscience lab, the Netherlands	August, 2003
Subtotal In-Depth Studies	12.0 credits*

4) Personal development	*date*
► **Skill training courses**	
How to write scientific proposal (in graduate school of CAAS), China	2002
► **Organisation of PhD students day, course or conference**	
4th International ISHS Symposium on Edible Alliaceae in China	April 21 - 25, 2005
► **Membership of Board, Committee or PhD council**	
Subtotal Personal Development	4.5 credits*

TOTAL NUMBER OF CREDIT POINTS*	**35.3**

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

Printed by: Printparters Ipskamp