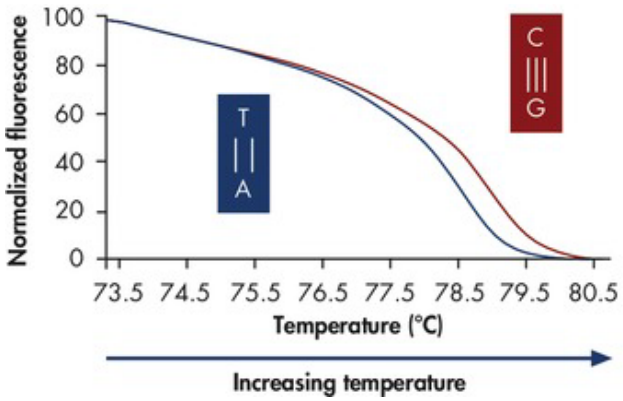
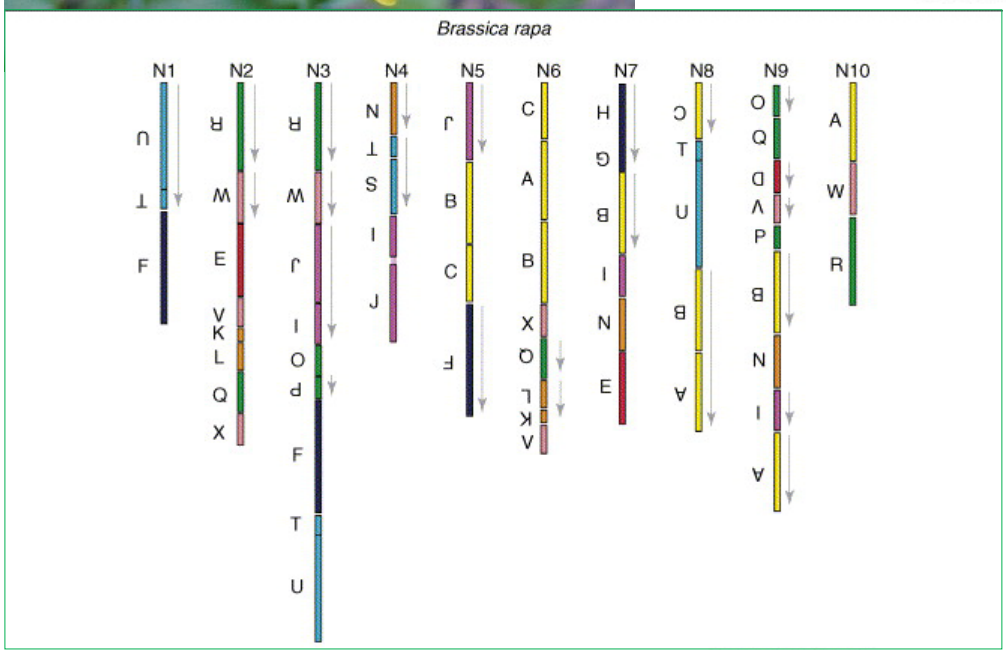
Challenges in developing flowering time related genetic markers polymorphic between Pak Choi and Chinese cabbage: two genetically related parents of a double haploid (DH) population





Name: Brian Chisunka

Chair group: Plant breeding

Wageningen University and Research Centre (WUR)

January, 2013

Challenges in developing flowering time related genetic markers polymorphic between Pak Choi and Chinese cabbage: two genetically related parents of a double haploid (DH) population

M.Sc. minor thesis

Submitted to the Department of Plant breeding of Wageningen University and Research Centre (WUR) as part of the requirement for the degree of

MASTERS of PLANT SCIENCES (MPS)

with specialization

In

PLANT PATHOLOGY AND ENTOMOLOGY

By

Brian Chisunka

WAGENINGEN UR

January 2013

Challenges in developing flowering time related genetic markers polymorphic between Pak Choi and Chinese cabbage: two genetically related parents of a double haploid (DH) population

Student: Brian Chisunka

Registration number: 78080215070

Email: brian.chisunka@wur.nl

Course code: PBR-80424

Supervisors: Dr. Ir. Guusje Bonnema

Dr. Xiao Dong

Johan Bucher

Examiners: Dr. Ir. Guusje Bonnema

Prof. Richard Visser

Table of contents

Acknowledgements

I would like to thank all the members of the *Brassica* group for welcoming me in the laboratory. In particular the group leader, Dr. Ir. Bonnema for her advice in my study and encouragements. I also thank Dr. Xiao Dong and Johan Bucher for their supervision throughout my minor thesis work. I sincerely express my gratitude to Ram Basnet for helping me in data analysis.

I also thank Prof. Richard Visser for his favourable response when I needed help. All in all, I thank all the members of the laboratory of plant breeding for the help given to me, especially technical and laboratory managers.

My minor thesis work is dedicated to my family for the moral and financial support given to me. This is specifically to my late parents, Annabella Kasakula, my dearest loving dad, Bornface Chisunka Kapansa who always wanted the best for me. I also remember my late sisters, Susan Chisunka Chishimba and Lizzy Kapansa for the memories we shared, this gives me endurance to overcome any challenge I face.

My sister, Marriane Chisunka Tembo, destiny fixed her to be my mentor and inspirational friend. This has been coupled to what I have seen and learned from my brother in-law, Solomon Tembo who has been there for me in times of need.

I also remember my employer, the Government of the Republic of Zambia (GRZ) for giving me financial support; this has made me fulfil my dream of attaining higher education. My sincere gratitude goes to the Director of SCCI, Dr. Mungoma for the confidence she has in me. I also thank the Agricultural Development Support Project (ADSP) national coordinator, Martin Liywali for facilitating my sponsorship.

Abstract

In this study a mapping population was created from DH parents, Brassica rapa subsp. pekinensis (DH CC48) and B. rapa subsp. chinensis (DH PC101). This DH population was designated as DH88. B. rapa shows a lot of intraspecific variation that is observed with different morphotypes such as turnips, heading type leaf vegetables (Chinese cabbage) and non-heading type leaf vegetables (Pak Choi). These morphotypes have different flowering time, hence are cultivated in different geographical locations. Differences in flowering time have been found to be associated with allelic variation at the FLOWERING TIME LOCUS (FLC) and FRIGIDA (FRI) in A. thaliana; colocalization studies in sytenic B. rapa have also confirmed the role of these loci in flowering time variation in populations studied in the Brassica group. Hence, in this study, marker development was undertaken for flowering time together with phenotyping of flowering time trait, in order to develop markers to be used for mapping of fQTL. Two marker types were used: single nucleotide polymorphism (SNP) markers based on Light Scanner and fluorescently labeled simple sequence repeats (SSR) markers separated on Licor gels. Genotyping of these markers showed rates of polymorphism amounting to 20% of a total of 147 markers screened. Linkage analysis showed skewness of 59% of the markers. Due to the low number of markers fitting on the chromosomes, only about 8 linkage groups were detected and fQTL analysis showed weak association of the markers with flowering time as no fQTL could be detected at $p < 0.001$. However, at $p < 0.05$, 6 markers were found to be associated with flowering time. Because of flowering time marker's low rate of polymorphism in this population, a large platform for genotyping flowering time markers is needed such as infinium ray microarray.

Key words: *Brassica rapa,* flowering time, population, markers, polymorphism, Light Scanner

INTRODUCTION

The *Brassica* genus consists of a large group of important vegetable, oil, fodder and condiment crops (Zhao et al., 2005a).These crops show a lot of variation that accounts for adaptation to different environmental conditions, which has enabled the *Brassica* to be grown in main different climatic regions. *Brassica*'s together account for about 12% of the world edible oil production (Zhao et al., 2005a). *Brassica's* are closely related to *Arabidopsis thaliana* a member of the *Brassicaceae* family. The lineage separation of *Brassica* from *Arabidopsis* occurred after a genome triplication between 7 and 9 million years ago (MYA) (Yang et al., 1999; Town et al., 2006; Wang et al., 2011). Hence, the *Brassica* share collinear blocks of genes with *Arabidopsis* and orthologous genes are arranged in syntenic pattern on their chromosomes (Wang et al 2011).

1.1 *Brassica* genome relationships

Brassica are model group of plants used to study polyploidy, this is due to the amphidiploid species that arose from hybridization between three species followed by chromosome duplication (Li et al., 2009; Zhao et al., 2005a; Wang et al., 2011, U, 193). The ancestral *Brassica* species is believed to have had eight (8) haploid chromosome number (**n = 8**). The relationship among the diploid and amphidiploid *Brassica* has been described in the "U Triangle". According to this triangle, the diploid *Brassica*, *Brassica rapa* (Genome A: **n = 10**) hybridized with *B. nigra* (Genome B: **n = 8**) and *B. oleracea* (Genome C: **n = 9**) to form amphidiploid species, *B. juncea* (Genome AB: **n = 18**) and *B. napus* (Genome AC: **n = 19**) respectively. While hybridization of the diploid species, *B. nigra* and *B. oleracea* resulted in the amphidiploids, *B. carinata* (Genome BC: **n = 17**). *Brassica*'s show a lot of genetic plasticity that is exhibited in different morphotypes (Li et al., 2009). This variation has been alluded to higher nucleotide polymorphism among *Brassica* (Wang et al., 2011).

Based on the organs consumed and morphotypes, a number of cultivar groups have been classified in *B. rapa* (Dierderichsen, 2001; Bonnema et al. 2010). Many groups of *B. rapa* are cultivated for their leaves, among them are Pak Choi and Chinese cabbage. Chinese cabbage is characterized by large leaves with wrinkled surfaces, pale green colour, white mid midrib vein and different heads of varying shapes. Pak Choi is darker, smooth green with green leaves having white midribs and does not form heads (Zhao et al., 2005a). The genetic relationship between accessions of different morphotypes showed that Chinese cabbage and Pak Choi form separate clusters, but CC and PC's are closely related. Origin is based on independent and separate domestication (Zhao et al., 2005a). These relationships among *B. rapa* subspecies shows low bootstrap values for many groupings, indicating that polymorphism contributes a small fraction to phenotypic variation (Zhao et al.,

2005a). Hence, only a few genes can account for the differences in the observed morphotypes among *B. rapa* sub species.

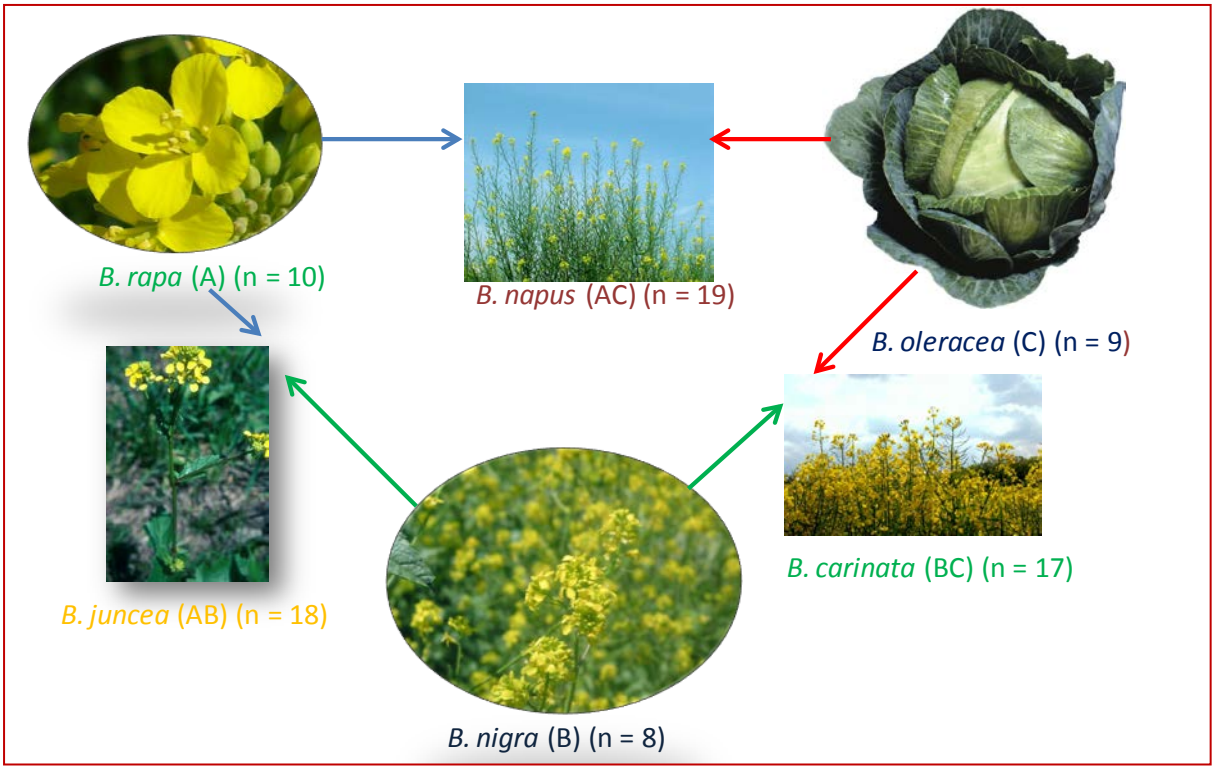


Figure 1. *Brassica* genome duplication showing the relationship between the diploids *B. rapa* (Genome A (n=10)), *B. nigra* (Genome B (n=8)), *B. oleracea* (Genome C (n=9)) with amphidiploids *B. juncea* (Genome AB (**n=18**)), *B. napus* (Genome AC (**n=19**) and *B. carinata* (Genome BC (**n=17**). N is the number of haploid chromosome (U, 1935).

1.2 Genetic maps and genome of *Brassica rapa*

B. rapa consists of leafy vegetables such as Chinese cabbage (*Brassica rapa* spp. *pekinensis*), Pak Choi (PC) (*Brassica rapa* spp. *chinensis*), turnips, oil seed turnips and sarsons (Li et al., 2009). Genetic maps make it possible to map traits and find markers linked to other traits (Li et al., 2002). The information contained in the maps can be used in crop improvement through efficient selection of agronomical traits using molecular techniques such as marker assisted selection (MAS). *Brassica* have undergone whole genome triplication (WGT) between 7 and 9 MYA (Wang et al., 2011). Many *Brassica* have characteristic mesohexaploidy, which provide an opportunity to study genome evolution in plants (Wang et al., 2011). The recently assembled *B. rapa* genome of Chiifu-401 covers over 98% gene space, which is equivalent to 283.8 Mb (Wang et al., 2009). This genome encodes about 41,174 genes designated as proteins (Wang et al., 2011). Genetic mapping of 1,427 markers in this genome has resulted in the assembly of pseudo chromosomes in *B. rapa* (Wang et al., 2011). The difference between *Arabidopsis* genome and *B. rapa* has been alluded to transposable elements in the latter that make up 39.5% of its genome (Wang et al., 2011). Transposable elements are

spaced throughout the genome, but are mostly found near the centromere. This close relationship between *Arabidopsis* and *B. rapa* enables the use of *Arabidopsis* as an out group in the study of *Brassica* (Wang et al., 2011). Hence, comparison of collinear blocks of the *B. rapa* and *Arabidopsis* is done with regards to genome triplication. It has been found that a total of 108.6 Mb of *A. thaliana* genome and 259.6 Mb of *B. rapa* genome assembly are contained in the collinear blocks (Wang et al., 2011).

Brassica species show ability to generate different morphological variants under selection by plant breeders (Teale et al., 2006). This can be attributed to large number of gene copies that can allow mutation at an increased frequency (Wang et al., 2011). In *B. rapa* it was observed that synonymous mutation was higher than other crop species, such as *Carica papaya* and *A. thaliana* when orthologue genes in these crops were compared (Yang et al., 1999; Wang et al., 2011). The generation time of papaya is longer than that of *Brassica* as papaya is a perennial crop, hence cannot undergo many cycle to observe mutation compared to *B. rapa* (Wang et al., 2011). The expansion of auxin related gene families in *Brassica* could have resulted in different morphological traits due to auxin controlling many growth and developmental processes in plants (Wang et al., 2011). Also the regulation of flowering is critical in *Brassica* as it impacts on the morphotypes generated, which also responds to different environmental cues (Wang et al., 2009; Bonnema et al., 2011).

1.3 Flowering time control in *Brassica rapa*

Genomic triplication among *Brassica* has resulted in contrast among genes controlling flowering time (Wang et al., 2009). In terms of *Brassica,* candidate genes like FLOWERING LOCUS C (FLC), CONSTANS (CO) AND FLOWERING TIME (FT) have been studied in relation to the control of flowering time (Ra et al., 1999; Zhao et al., 2010b; Wang et al., 2009; Xiao et al., submitted). It has been found that FLC genes in *Brassica* functions similar to its orthologues in *Arabidopsis,* this is through co-localization and vernalization studies aimed at elucidating flowering time (Zhao et al., 2010b). The flowering locus (FLC) has four orthologues in *B. rapa* that came about through WGT (Wang et al., 2009). The paralogues of FLC found in *B. rapa* are BrFLC1, BrFLC2, BrFLC3 and BrFLC5 (Kim et al., 2006; Yang et al., 2006; Schranz et al., 2002). These genes are crucial in the control of flowering time. Hence, temporal control of this phase is very important for agricultural production as it guarantee seed production and maturity (Mendez-Vigo et al., 2010). Therefore, the control of traits like flowering time (FT) is required in order to synchronize flowering time with seasonal changes such as light (day length) and temperature with crop growth (Mendez-Vigo et al., 2010). Also investment in the vegetative and reproductive structures by the crop is influenced by the transition to the reproductive phase, which has a pronounced effect on yield of the crop. Vernalization of

germinated seeds of *B. rapa* results in shortened generation time due early induction of flower initiation (Lin et al., 2005). In *Arabidopsis*, it has been found that the number of leaves induced during the vegetative stage determine the initiation of flowering (Zapater and Somervile, 1990; Steynen et al., 2001). Many genes that have an effect on flowering control different pathways that have an effect on flower development. These are mainly the vernalization pathway, photoperiod pathway, autonomous pathway and the gibberellin pathway (Mouradov et al., 2002; Schmitz and Amasino, 2007; Alexandre and Hennig, 2008; Seo et al., 2009). But most of the observed differences amongst *Arabidopsis* ecotypes relating to flowering time are due to allelic variation of FLC and FRIGIDA (FRI) loci (Ehrenreinch et al., 2009; Kim et al., 2005; Koornneef et al., 2004). In terms of flowering initiation in *B. rapa*, FRI gene acts upstream of FLC, thereby promoting its gene expression (Koornneef et al, 2004). FLC encodes a MADS-BOX transcription factor that represses flowering, but is negatively affected by vernalization (Sheldon et al., 1999; Michaels, and Amasino, 1999; Koornneef et al., 2004; Luo et al., 2007). BrFLC2 expression has been found to be highest in late DH lines of *B. rapa* compared to early lines (Zhao et al., 2010b). BrFLC2 expression has also been found to decrease after the onset of vernalization in both early and late DH lines, which suggests that commitment to flowering is an early development process (Zhao et al., 2010b). A total of 300 genes have been implicated in the control of flowering in *Arabidopsis* (Rae et al., 1999; Zhao et al., 2010b).

1.4 Environmental effect on flowering time quantitative trait loci (QTL) in *B. rapa*

The effect of the FLC locus in controlling flowering time under diverse environmental conditions has not been fully elucidated. In the study of different quantitative trait loci (QTL) for flowering and morphological traits using DH populations in *B. rapa*, major QTLs have been found to colocalize with BrFLC2 (Luo et al., 2007). In crosses between PC and Yellow Sarson (YS), and Rapid cycling (RC) and turnip (VT), it was found that the control of many QTL linked genes is complex due to their relation with other gene networks, which result in difficult in isolating single gene effects (Yu et al., 2003; Zhao et al., 2010b). A study involving four different DH populations of *B. rapa* generated from diverse parental lines, showed two flowering time fQTL designated as FLQTL-2 and FLQTL-6, which were detected under different environmental conditions, showing that they were marginally or not affected by these environmental conditions (Luo et al., 2007). These QTLs were depicted to be associated with flowering time related genes previously detected in *B. rapa* (Luo et al., 2007). FLQTL-2 was associated with BrFLC2 on RO2 and this QTL determines the differences in flowering time between early flowering oil type and middle maturing morphotypes in *B. rapa* (Luo et al., 2007). BrFLC2 has been found to be influenced by the environment, in particular cold temperature which reduces gene expression (Koornneef et al., 2004; Luo et al., 2007, Zhao et al., 2010b). In *Arabidopsis*,

colocalization of QTLs has revealed a modular genetic architecture were similar loci control a number of related functional processes (Koornneef et al., 2004). This modular genetic architecture implies a change in quantitative gene expression of the trait under the control of specific QTLs following a specific pattern. Linking these observation with *B. rapa*, the effect of other genes on flowering time is feasible due to the sequenced genome which has syntenic relationship with *Arabidopsis* (Wang et al., 2011). The flowering locus paralogues in *B. rapa,* BrFL2, BrFLC3 and BrFLC1 have been assigned to loci linkage groups, RO2, RO3 and RO1 respectively (Kole et. al., 2001; Schranz et al., 2002; Kim et al., 2006, Xiao et al., submitted).

1.5 Genetic markers used in linkage map construction

In determining the genetic contribution of quantitative traits that are influenced by polygenic expression, QTL analyses are used (Coillard et al., 2005). QTL analysis involves linking phenotypic expression of individual segregating population such as F2 with respective associated genes. QTL analysis involves the use of markers, which show genetic differences between species or genotypes (Coillard et al., 2005). Markers are broadly divided into three major categories: morphological, biochemical and molecular markers (DNA) (Colliard et al., 2005). Morphological markers are based on the phenotypic differences between genotypes such as flower colour, pod colour or hairiness. Biochemical markers on the other hand are based on differences in enzyme expression of genotypes also called isozymes (Collard et al., 2005). These are assayed by protein analysis through gel electrophoresis. Molecular or DNA markers are based on the genome sequence of genotypes, which shows differences based on polymorphism (Collard et al., 2005). Molecular markers are unlimited in number compared to both morphological and biochemical markers. Also molecular markers unlike their counterparts are not influenced by environmental conditions and the development stage of the crop (Collard et al., 2005).

Molecular markers arise due to various mutations in DNA sequence. These can be due to nucleotide substitutions (point mutations), rearrangement of the genome (insertion or deletions) or errors in replication of tandem repeats of DNA (Paterson, 1994). Molecular markers can be used in quantitative genetic diversity studies of populations and in assaying the cultivar purity of varieties (Weising et al., 1995; Jahufer et al., 2003). Molecular markers reveal differences among genotypes based on alleles for a particular marker, which are visualized through techniques such as gel electrophoresis or calometric probes through emission of energy inform of radiation or light (Collard et al., 2005). Markers that differentiate genotypes are referred to as polymorphic, while those that do not are known as monomorphic (Collard et al., 2005). Polymorphic markers can either be

dominant or codominant, this classification is based on the marker's ability to differentiate between the homozygote individual and the heterozygote (Collard et al., 2005). Generally codominant markers show genetic difference based on size, while dominant markers are based on qualitative present or absent of a particular allele. Codominant markers may show different alleles, while dominant markers only show one marker, while the other alleles are not amplified (Collard et al., 2005).

Essentially, generation of a linkage map involves three major parts. The first part involves the generation of a mapping population, followed by determination of polymorphism on the individual members of the population and finally linkage analysis of the markers (Collard et al., 2005). Production of a mapping population requires that parental lines used have enough variation due to differences in genetic distance that shows enough polymorphism for the traits of interest. A minimum of 50 individual genotypes can be used for QTL detection, but the higher the number, the more resolution and fine-tuned the QTL map generated (Collard et al., 2005).

1.5.1 Mapping population for QTL analysis

Mapping population used depend on the reproduction of the crop species used in the study. In self-pollinated crops, homozygote parents are used to generate segregating population by crossing them into F1, which then is allowed to segregate into F2, which can be used as mapping population (Collard et al., 2005). While in the case of cross pollinated species like *B. rapa*, the situation is different due to inbreeding depression and self-incompatibility. Also most cross pollinated species show polyploid which makes it difficult to generate homozygotes, such as in *B. juncea* and *B. napus* which have tetraploid chromosomes number (U, 1935). Therefore, in order to overcome these challenges, mapping populations for cross pollinated species are developed from crossing the heterozygote and haploids or with one of the parents and allowing segregating populations such as (F2) to be intersed to generate single seed descent. This, however, is time consuming requiring a minimum of 6 generations for the recombinant inbred lines (RIL) to reach some level of acceptable homozygous at the loci. Hence, the use of haploid gametes such as ovules and anthers is more time serving in crops were tissue culture techniques are amenable such as in cereals like wheat (*Triticum aestivum*)(Collard et al., 2005). Also crossing of the cultivated variety with wild relative in interspecific hybridization has been found to be useful in some cases such as that of barley (*Hordeum vulgare*) and wild barley (*Hordeum bulbosum*), which allow recovery of haploids in segregating generations such as F2. The chromosomes of the haploid gametes are doubled using tissue culture techniques resulting in homozygous lines known as double haploids (DH). DH lines are

produced mainly from anthers through chromosomal doubling using chemicals such as cochicine (Coillard et al., 2005). DH provides an efficient way of attaining homozygous at the loci, which can be used to study alleles.

1.5.2 Genetic map construction

Linkage maps designate marker position and their genetic distance on the chromosome (Collard et al., 2005). QTL mapping is based on the segregation of genes linked to markers during meiosis, which allow their analysis in their progenies (Paterson et al., 1996).This is because genes that are closely linked together will be assorted together in the progenies. In the segregating population, the frequency of recombinant genotypes is used to calculate recombination fractions, which are transformed into recombination frequencies. These are used to determine genetic distance among markers (Collard et al., 2005). The closer the markers (<10 centimorgan (cM)), the more likely they are situated on the same chromosome. Markers which are situated away from each other have recombination frequencies near 50 cM and are referred to as unlinked and maybe situated on different chromosomes (Patterson et al., 1996). The final step in the construction of a linkage map involves the coding of data on polymorphism on the markers individually, which is used to generate a linkage map. This involves the use of computer programs. Linkage between markers is calculated as a ratio of present over absent of the marker, which is expressed in logarithmic as logarithms of odds (LOD) (Collard et al., 2005). LOD values above 3 are used to construct linkage maps.

1.5.3 Simple sequence repeats (SSR) markers or microsatellite

Simple sequence repeats (SSR) are DNA tract consisting of tandem repeats (1 to 10 nucleotides) that are prone to mutation (Gao et al., 2011). SSR markers arise due to unequal crossover and slippage replication in the genome of eukaryotes (Goldstein and Schlotterer, 1999). These markers are based on DNA sequences flanking the short repetitive unique DNA sequences such as mononucleotides, dinucleotides and trinucleotides (Lowe et al., 2003). SSR markers are codominant, multiallelic and are highly reproducible, also amenable to high throughput genotyping (Lowe et al., 2003). SSR have been found in both intergenic and intragenic regions including both intron and exons (Zhang et al., 2004, Subramanian et al., 2003; Hall, 1999). Distribution of SSRs in the genome varies among species and genomic region (Lawson and Zhang, 2006). Most SSR are frequently located in intergenic regions, 5'UTRs and 3'UTRs and less in the coding regions as found in rice (*Oryza sativa*), *A. thaliana* and *B. rapa* (Fujimori et al., 2003; Lawson and Zhang et al., 2006; Lawson and Zhang, 2006; Hong et al., 2007). Expressed sequence tag (EST) derived SSR markers have the potential to reveal functional variations between genotypes compared to genomic based SSR markers (Ramchiary et al., 2011; Zhang et al., 2004). Also SSR derived from EST can easily be transferred across species, which make

them useful for comparative mapping studies. SSR markers located in the coding regions of the gene are highly conserved, hence are less polymorphic compared to those in the non-coding regions (Ramchiary et al., 2011). This can be useful in detecting gene order and conservation, which are useful means of identifying candidate genes. In *Arabidopsis*, it has been found that one SSR is detected per 8.57 kb and while in *Brassica* unigene, one SSR is detected per 7.25 Kb (Parida et al., 2010). Also nucleotide composition of SSRs varies among species (Kumpatla and Mukhopadhyay, 2005; Jiang et al., 2006; Ueno et al., 2008). Variations in SSR are due to mutation of nucleotide sequences and changes in number of repeats (Li et al., 2003). In related species, SSR variation depends on number of repeats as most of them are conserved (Vigouroux et al., 2003; Asp et al., 2007; Gao et al., 2007). The frequency of SSR increases with genome size (da Maia et al., 2009; Gao et al., 2011).

1.5.4 Single nucleotide polymorphism (SNP) markers

Single nucleotide polymorphism (SNP) represents the most abundant DNA sequence variation in an organism (Park et al., 2010). SNP are single base differences between DNA of different individuals. When identified, these can be converted into markers that can be used in genotyping (Gut, 2001; Kwok, 2001). SNP are increasingly becoming the most highly favoured molecular markers of choice compared to conventional markers such as AFLP, RFLP and SSR. Many traditional SNP genotyping methods involve the use of CAPS markers and AS-PCR (allele-specific PCR) (Liu et al., 2012). But CAPS markers can be restricted by endo nuclease site that can be inefficient and not cost effective at times (Komeczny and Ausubel, 1993; Neff et al., 2002; Thiel et al., 2004). In *Arabidopsis* more than 37,000 SNPs have been identified through comparison of two accessions, Colombia (Col-o) and Ler (Jander et al., 2002) and many more accessions now in recent times. Because of their abundance, SNPs have the potential to provide highest refined map resolution. Many tools have been developed for genotyping using SNP at high throughput with reasonable cost reduction such as the golden gate and infinium arrays. The occurrence of SNP in the genome has been found for crops like maize to be one SNP per 31 bp in the non-coding regions and 1 SNP per 124 bp in the coding regions in 18 genes (Ching et al., 2002).

1.6 QTL analysis

QTL analysis involves the association of phenotypic variation with genotypic variation (Collard et al., 2005). This is done through the phenotyping of a trait on a mapping population of individual genotypes, which are then screened with markers that underlie a particular region on the chromosome containing genes that are involved in regulation of a trait of interest. Markers are used in order to detect polymorphism among the genotypes. Markers discriminate the population into

particular groups representing different alleles. The marker data generated is used to make linkage analysis to detect the genetic distance at the loci. This data with the phenotypic data are correlated based on the significance of association between the two results. The detection of correlation at a given confidence level between phenotypic variation and markers shows the presence of QTL for a particular trait. This is used to explain the genetic variation for that particular trait and the location of the genes that are involved regulating gene expression.

1.7 Objectives of this study

It has been found that complex genetic factors control flowering time in *B. rapa* and these are not fully characterized. Hence, in this study the main aim was to develop markers that are based on flowering time genes in *B. rapa*, to construct a genetic map, to map flowering time fQTL in a *B. rapa* DH population of 88 lines, from a cross between a PC DH line and a CC DH line. This is because flowering time is a quantitative trait under the control of polygenic expression.

The following specific objectives were addressed:

1. To identify markers based on flowering time gene sequences that are polymorphic in *B. rapa* DH population derived from Chinese cabbage and Pak Choi (DH88).
2. Phenotyping of flowering time (FT) trait of DH88 population grown in the greenhouse in three blocks, with one plant per DH line per block.
3. To correlate flowering time with allelic variation for genetic markers.

2 MATERIALS AND METHODS

2.1 Plant materials and environmental conditions for growth

Parental lines of Chinese cabbage accessions DH CC-48 (CG06867) (heading) that originated from the former Soviet Union was used as the male parent in a cross with DH PC-101 (CG13926) (non-heading) as the female parent, that originated from China. The parental lines were supplied by the Dutch Genetic Resources Centre (CGN). The generated population, DH88 was grown under greenhouse conditions at Nergena, Bennekom, The Netherlands (5º 57' 57'' N, 5º 38' 37'' E). This population was grown from May to December, 2012. Pregerminated seeds were sown on filter paper at 25ºC under dark conditions for two days, then subsequently transferred to the nursery trays after two weeks. Seedlings were transplanted at 5 leaf stage to pots measuring 14 cm in diameter and later on at 45 days to pots of 21 cm in diameter. Bud pollination of DH88 lines was undertaken in order to multiply the seed. This involved selfing of the plants after the complete opening of the flower bud into a complete flower. The stigma was opened using forceps and pollen from the same plant was placed on the cut stigma. The selfed flower was labelled with the tag showing the date and the pollinator then covered with a transparent plastic.

2.2 Experimental design

A complete randomized block design was used. This had 3 replications consisting of 3 blocks, with block A having 82 lines, block B with 80 lines and block C with 77 lines. Block B and C had missing genotypes; this was caused by low seed set and low germination.

2.3 Phenotyping of flowering time (FT)

Only flowering time (FT) was measured. This involved the measurement of days to flowering time in days (d). Time to flowering was scored based on the floral bud opening into a complete flower. The average of the three blocks was taken as flowering time. The range of the result was taken as the difference between the minimum and highest days to flowering fixed at 300 days for the non-flowering plants when the data was taken. Normality test was done on the results of flowering time in order to determine how the results were distributed in the population so as to find the most accurate statistic method of analyzing the phenotype and genotype. This involved fitting the results with the residues and observing the distribution in the Cartesian plain.

2.4 Marker development for flowering time and leaf traits for *B. rapa*

A total of 147 *Arabidopsis* genes that are based on literature (Roux et al., 2006; Brachi et al., 2010; Fornara et al., 2010; Xiao et al., submiitted) which are involved in flowering time and leaf trait regulation were developed using *Arabidopsis* database (http://www.arabidopsis.org/) (TAIR). These were then blasted as queries in their homologues in *B. rapa* database (http://brassicadb.org/brad/) (BRAD) of annotated Chinese cabbage Chiifu 401 (Wang et al., 2011). The gene sequences were used to design primers using the program Primer3 (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) (Rozen and Skaletsky, 2000). Most of these markers were already mapped in DH68 population, amounting to 99 in number (Zhang, unpublished).

2.5 Polymorphism detection: high melting point resolution analysis (Light scanner)

DNA used in the detection of polymorphism was extracted from DH88 population (Mwaangi et al., 2012). Light Scanner (Idaho Technology Inc.) was used to detect polymorphism based on single nucleotide polymorphism (SNP). Screening of polymorphism on DH88 population involved pre-screening of markers on the parental lines, DHCC48 and DHPC101. Upon detection of sufficient polymorphism, genotyping of the marker on the population was done. The Light Scanner PCR reaction involved: 1 µl of 20 ng/µl of DNA, 0.1 µl of phire enzyme, 2 µl of phire enzyme buffer (X5), 1 µl of LC-green (X5), 0.4 µl of 0.2 mM of dNTP, 0.25 µl of 2.5 mM both forward and reverse primers and 5 µl of milliQ water. This amounted to the total reaction volume of 10 µl. The reaction mix was overlaid with 20 µl of mineral oil. PCR reaction was done using a 96 well 7300 Thermo cycler (Biorad USA). The PCR reaction temperature cycling protocol was: initial denaturation step of 98 ◦C for 30s, then 40 cycles of denaturation at 98 ◦C for 10s, annealing at 60 ◦C for 10s, extension at 72 ◦C for 30s. This was then followed by further extension at 72 ◦C for 30s and another denaturation at 94 ◦C for 30s to facilitate formation of heteroduplex, then cooling down to 25 ◦C for 30s and then left at infinite at 10 ◦C. After PCR, the amplicons were scanned for polymorphism on a Well Light Scanner® System (Idaho Technology Inc., USA) for higher resolution melting point (HRM) analysis. The light scanner setting was done at 70 ◦C as minimum temperature and 96 ◦C as maximum temperature with 67◦C as holding temperature. The lamp heating rate was set to automatic at ◦C/s. Analysis for SNP was done using the light scanner software (version 2) following Montgomery et al., 2007. Polymorphism was detected according to fluorescence of the introduced dye that was held between double stranded DNA and the surface, which is given out during denaturation. The melting profiles are calibrated by internal oligonucleotide and then normalized keeping the differences between the lower minimum temperatures to 1 ◦C and vice-versa for the maximum temperature differences. Polymorphism is computed according to common differences, which are visualized as fluorescence

bands versus temperature. The genotype of the alleles were scored according to the parental lines with CC48 as "a" and PC101 as "b".

2.6 Development of SSR markers linked to flowering time regulation in *B. rapa*

In order to analyse candidate genes associated with flowering time, SSR markers were developed from *Arabidopsis* (Ehrenreich et al., 2009). The coding sequence (CDS) of these candidate genes were taken from *Arabidopsis* database (above). The CDS was then used to search for orthologues in *B. rapa* database (above) by using BLAST in the region of 15 Kb upstream and downstream of the candidate gene for SSR motifs of the genes using the program FastPCR6.0 (http://primerdigital.com/fastpcr.html). The SSRs obtained consisted of different mononucleotides, dinucleotide, trinucleotide repeat types. When the gene targeted had more repeats, dinucleotides and trinucleotides were preferably chosen in this study. Primers were designed using Primer3plus as above based on the flanking sequence of the microsatellite.

The PCR reaction undertaken was 1 µl of 50 ng/µl DNA, 1 µl Dreamtaq buffer (x5), 1 µl of 2.5 mM of both forward and reverse primer, 0.4 µl of 0.2 mM of dNTP, 1U of 0.04 µl of Dreamtaq polymerize enzyme and 5.56 µl of milliQ water. The total reaction mix amounted to 10 µl. The PCR reaction was carried out using normal PCR reaction: 94 ◦C denaturation for 5 minutes, followed by 32 cycles of 94 ◦C of denaturation for 30s, 55 ◦C of annealing for 30s, 72 ◦C of extension for 30s. This was then followed by further extension step at 72 ºC for 7 minutes and kept at 10 ◦C at infinite. 0.6 µl of each amplicon was mixed with 20 µl of Blue buffer (98% formamide, 10 mM EDTA pH 8.0 and 0.1% bromophenol blue). The forward primers used in this study were radio labelled with 700 IRD and 800 IRD at their 5.' Upon mixing, the amplicons were stored in the fridge at 4◦C and denatured at 94.5 ◦C before loading after putting them on ice box. PCR amplicons were run on the gel using polyacramide gel on Li-cor sequencer for 2 hrs (**Appendix X**).

2.7 Linkage analysis

Linkage analysis for the markers was done using Join Map 4.1 (Van Ooijen, 2006). The marker order was calculated using the best fit of Chi-square test at p=0.05 and 1 degree of freedom using the Kosambi mapping function to change recombination frequencies into centimorgans (Kosambi, 1944).

2.8 QTL analysis

Data analysis to detect association between flowering time and markers was done using R Project for Statistical computing (R version 2.13.1). This was done using the non-parametric analysis of variance by means of the performing the Kruskal-Wallis Test due to lack of normal distribution of the

data. Analysis was done with p values of α=0.05, α=0.01 and α=0.001. The ranking of the markers was done based on their physical location on the *B. rapa* genetic map.

3 RESULTS

In this study using a DH population, DH 88 derived from a cross between double haploid lines, Chinese cabbage (CC) (folding type) and Pak Choi (non-folding type) was used to screen flowering time associated markers designated as candidate genes for polymorphism. This mapping population was used for genotyping of markers after pre-screening the parental lines for polymorphism. Pre-screening of DH PC101 and DH CC48, which are closely related genetically, was envisaged to show polymorphism for markers that underlie QTL for flowering time. In terms of flowering time (FT), PC is early flowering compared to CC (Zhao et al., 2005a). Therefore, it was inferred that polymorphism for these marker will be sufficient despite close relationship among the two subspecies of *B. rapa*.

Hence, in this study polymorphism and days to flowering time was analysed so that QTL analysis can be done in order to explain the genetic variation.

3.1 Phenotypic variation for flowering time

Phenotypic variation of DH88 showed transgressive segregation of the progenies with average days to flowering time of 202 d with the parental lines CC 48 and PC101 having 200 d and 164 d respectively (**Figure 2A**). Most of the individuals showed transgressive segregation towards late flowering (**Figure 2A**). Variation among the different blocks was minimal for most of the genotypes, less than 10 days to flowering time days. Only for genotype which flowered late was variation started going up. Those genotypes which did not flower after taking of the data were given days to flowering time score of 300 d. Upon plotting of the days to flowering time, the data was log transformed in order to observe the trend of the distribution of the population. This was done with respect to analysis of variance if the data was able to follow normal distribution. However, this was not the trend as can be observed from the result of log transformation of flowering time and which was identical to that of the untransformed (**Figure 2B**). The normality test of the data also showed that the distribution did not form normal distribution as the samples were not fitting on the line (**Appendix VI**).

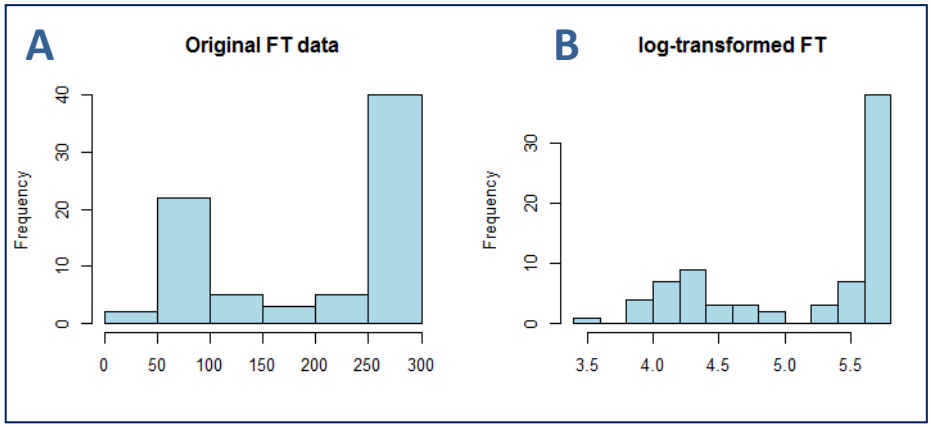


Figure 2A – B. Flowering time histogram of DH88 population, Y-axis genotype frequency (n) and X-axis showing days to flowering in days (d). Figure 1A, flowering time showing skewed distribution towards late maturing. Figure 2B. Log transformation of flowering time showing the same trend as the untransformed.

3.2 Annotation of markers designated as candidate genes in *B. rapa*

Markers underlying major QTLs in *B. rapa* that have been designated as candidate genes were genotyped after obtaining their sequences from NCBI and corresponding genes in *Arabidopsis* database (TAIR) then searched blasted in *Brassica* database. The homologous syntenic genes were then annotated according to the characterization done on *A. thaliana*, with the nomenclature of At as the beginning prefix for the gene in *A. thaliana* and Br for *B. rapa* homologue (**Appendix VII and IX**). These candidate genes show diverse functions in *A. thaliana* underlying various pathways in plants. The putative functions of the candidate genes that were found to be associated with fQTL in DH88 are detailed in table1 below.

Table 1. Functional annotation of candidate genes that were associated with fQTL in DH88 at $p < 0.05$.

Homologous gene in *A. t*	**Gene family**	***A.t* Chr.**	***B. Rapa* gene**	***B. rapa* Chr.**	**Marker name**	**Putative gene function**
At4g34160	CYCD3;1	4	Bra011501	1	BraCycD3;1P1a	Regulation of cell cycle, regulation of cell proliferation, regulation of meristem growth, response to brassinosteroid stimulus, response to cyclopentone, response to cytokinin stimulus, response to sucrose stimulus, seed development, spindle assembly, stomatal and complex morphogenesis
At2g44080	ARL	2	Bra004803	3	BrARLP1a	Brassinosteroid mediated signalling pathway, cell growth and response to brassinosteroid stimulus
At3g50070	CYCD3;3	3	Bra036051	9	BrCycD3;3P3a	DNA endoreduplication, mitotic cell cycle, negative regulation of cyclin-dependent protein kinase activity, regulation of cell cycle, regulation of cell proliferation and stomatal lineage progression
At5g17690	TFL2	5	Bra013958	10	BrTFL2P2a	Vernalization response, multidimensional cell growth, negative regulation of flower development, photoperiodism, flowering, regulation of mitotic cell cycle, shoot morphogenesis and vernalization response
At5g07200	GA20OX3	5	Bra009285	10	BrGA20OX3P4b	Gibberellin biosynthesis process, oxidation-reduction process and terpenoid biosynthesis process
At5g53950	CUC2	5	Bra003023	10	BrCUC2P2a	Formation of organ boundary, leaf development and leaf morphogenesis

3.3 Polymorphism rate between DH88 parental lines

A total of 147 markers were screened on parental lines DH CC48 and DH PC101 in order to analyse polymorphism. These markers consisted of 106 SNP and 41 SSR markers. 29 markers were polymorphic between parents and formed two allelic groups upon screening the DH88 population. This represented 20% polymorphism rate between the parental lines. In terms of SNP markers, the light scanner genotyping machine (Idaho Tech., USA. Inc.) was used to detect polymorphism based on the emission of fluorescence representing the melting curves of the nucleotides of the allele derived from the parental lines. Low rate of polymorphism were detected for SNP markers based on light scanner for many flowering time markers. The marker below designated as BrSPA1PA1a_A03 showed clear polymorphism between the parents (**Figure 3A**). However, the marker BrFLC2_A02 that has been found to underlie major flowering QTL in *B. rapa* in a PC x YS (Yellow Sarson) population (Zhao et al., 2010b), did not show polymorphism among the parental lines in DH88 (**Figure 3B**).

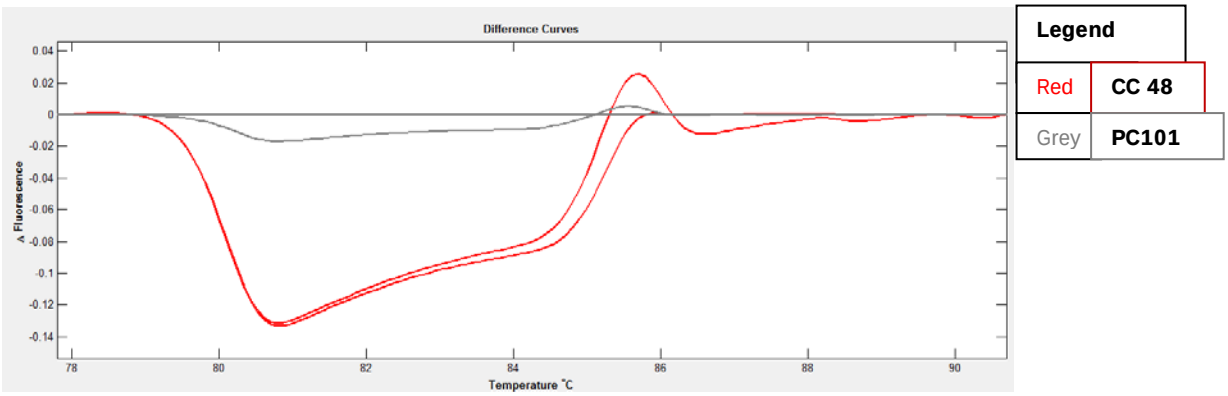


Figure 3A. The marker BrSPA1PA1a_A03 showing sufficient polymorphism between Chinese cabbage DH lines (double haploid lines) CC48 and Pak Choi (PC 101), parents of DH88 population.

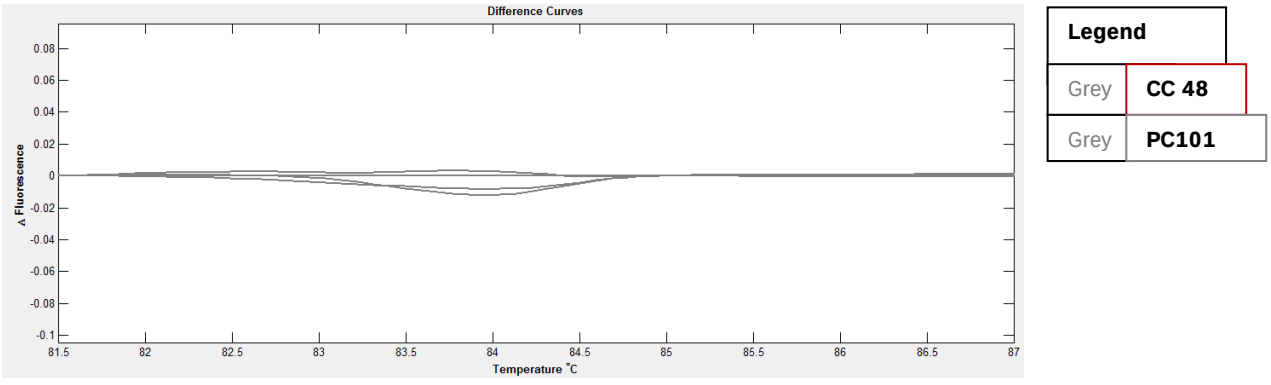


Figure 3B. BrFLC2_A02 marker underlying major flowering QTL in *B. rapa* showing low polymorphism between the parental lines, CC48 and PC 101.

3.4 Map positions of polymorphic markers in *B. rapa*

Upon screening of the markers that were deemed to be polymorphic on the parental lines, 24 SNP were able to separate individual DH lines into two distinct groups (**Figure 4**). These markers represented 34% of light scanner polymorphic markers screened on the population. Most of these markers did not show a 1:1 segregation of a DH cross, with skewness to either one of the parents **(Table 2**).

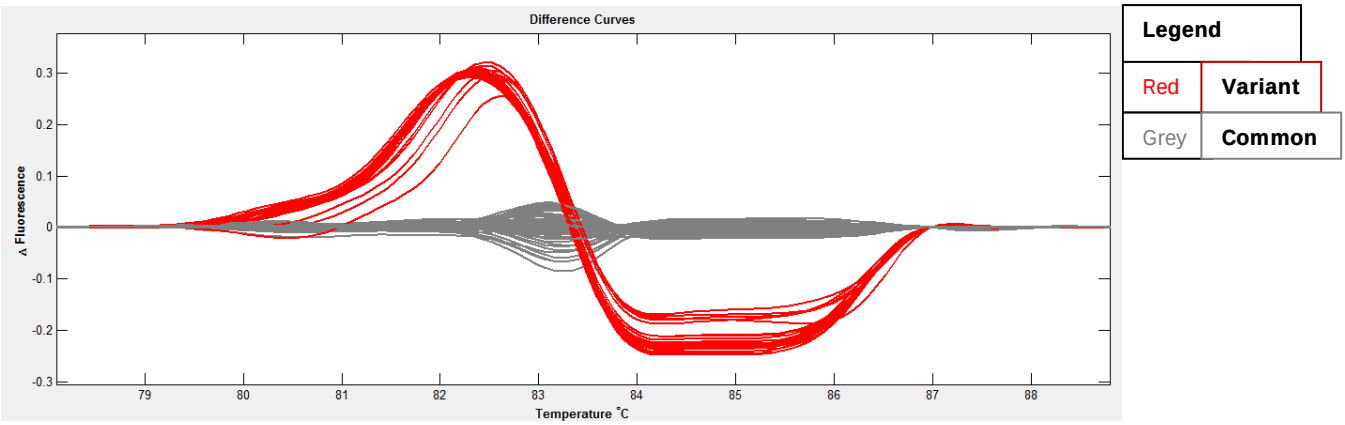


Figure 4. Flowering time marker, BrTSFP2c_A02 generated from light scanner showing sufficient polymorphism among DH88 population. The Y-axis depicting the fluorescence of nucleotide at different melting temperatures with 70 ºC being the starting point and 96 ºC the ending temperature and 67 ºC the holding temperature.

Table 2. Scored markers showing polymorphism on DH88 population to differentiate individual genotypes into two groups.

S/n	Marker		A	B	Ratio (a:b)	X2	Df	Signif.
1	BrSLY1P1b_A01	52	25	1	0.5	9.47	1	****
2	BrFLMP1b_A02.2	37	42	1	1.1	0.32	1	-
3	BrPIP2a_A02	24	56	1	2.3	12.8	1	******
4	BrTSFP2c_A02	24	56	1	2.3	12.8	1	******
5	ARF4P1b_A02	19	57	1	3.0	19	1	*******
6	AtBRAM_A03	56	23	1	0.4	13.78	1	******
7	BrFLKP2b_A03	62	15	1	0.2	28.69	1	*******
8	BrGAr3P1a_A03	56	24	1	0.4	12.8	1	******
9	BrTCTPP1b_A03	63	12	1	0.2	34.68	1	*******
10	ATVGT_A03	36	40	1	1.1	0.21	1	-
11	ARF3-A04	56	23	1	0.4	13.78	1	******
12	BrTCTPP2b_A05	39	40	1	1.0	0.01	1	-
13	BrELF5P1a_A06.1	27	51	1	1.9	7.38	1	***
14	BrGAr1P1d_A06	36	43	1	1.2	0.62	1	-
15	BrFKF1P3b_A07	29	51	1	1.8	6.05	1	**
16	BrFVEP2a_A07	30	49	1	1.6	4.57	1	**
17	BrESD4P2b_A08	32	47	1	1.5	2.85	1	*
18	BrFPAP1c_A09	58	16	1	0.3	23.84	1	*******

S/n	Marker		A	B	Ratio (a:b)	X2	Df	Signif.
19	BrFRL2P1b_A09	49	31	1	0.6	4.05	1	**
20	BrSNZP2b_A09	36	44	1	1.2	0.8	1	-
21	BrT1N6_2P1c_A09	37	43	1	1.2	0.45	1	-
22	BrFLDP1c_A10	66	8	1	0.1	45.46	1	*******
23	BrTFL2P2a_A10	39	40	1	1.0	0.01	1	-
24	BrARF4-Indel_A10	42	33	1	0.8	1.08	1	-
25	SOC1P2b_A04 SSR	44	33	1	0.8	1.57	1	-
26	FLC3P3b_A03 SSR	36	32	1	0.9	0.24	1	-
27	FLC5aP5a_A03 SSR	53	27	1	0.5	8.45	1	****
28	DDF1P2b_A09 SSR	48	28	1	0.6	5.26	1	**
29	CCA1P1e_A05 SSR	32	45	1	1.4	2.19	1	-

S/n (serial number), parental alleles (a = DHCC48 and b = DHPC101), p = 0.05 and DF (degrees of freedom, 1) Signif = Significances: * P ≤ 0.1, ** P ≤ 0.05, *** P ≤ 0.01, **** P ≤ 0.005, ****** P ≤ 0.0005, *******P ≤ 0.0001. DF (degree of freedom (1)), (a:b = CC48:P101) and X^2 = (Chi-square Test)

3.5 Flowering time markers showing polymorphisms between parents that do not segregate in DH88 population

About 70 SNP markers were generated as having polymorphism between DH88 parental lines using the light scanner genotyping machine. However, upon screening the DH progenies in the population, 46 of these markers showed clustering in one single ball like unit at the center, which represented 66% of the markers genotyped to be polymorphic on DH88 population (**Figure 5**). Hence, these markers could not be used for the analysing of linkage groups as individual genotypes were suspect in terms of the parental allele they carried. Clustering of these genotypes emanated from melting curves of the nucleotides generated, which were very close indicating the inability to effectively separate the alleles into two distinct groups (**Figure 5**). This made scoring of markers difficult.

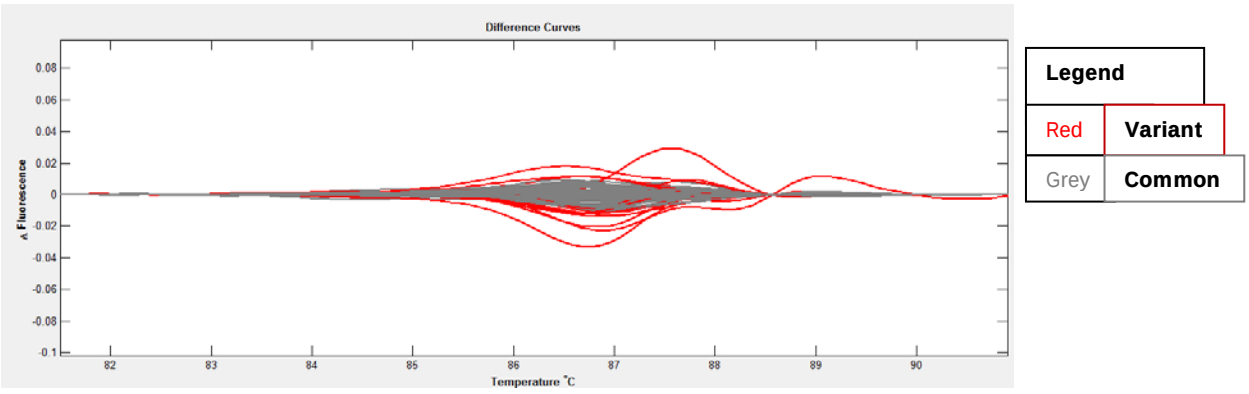


Figure 5. BrVRN1P3b_A03 marker showing clustering of genotypes making differentiating of parental alleles into two distinct groups difficult.

3.6 Flowering time simple sequence repeat (SSR) markers

A total of 41 SSR markers for flowering time were screened on parental lines of DH88 that resulted in 13 markers showing polymorphism, which represented about 32% of the total polymorphism between the parents. These markers showed polymorphism of varying sizes of parental alleles based on DNA bands. Many SSR markers genotyped on DH88 showed small differences which made it hard to score. (**Figure 6, Appendix II**). These markers upon screening on the population resulted in 5 SSR markers that could be scored very well and were being used for genotyping; the other 5 markers could not effectively separate the individual genotypes of the DH88 into the two parental alleles, CC48 and PC101, while the other 2 were not used due to technical problems resulting from poor amplification by the primers.

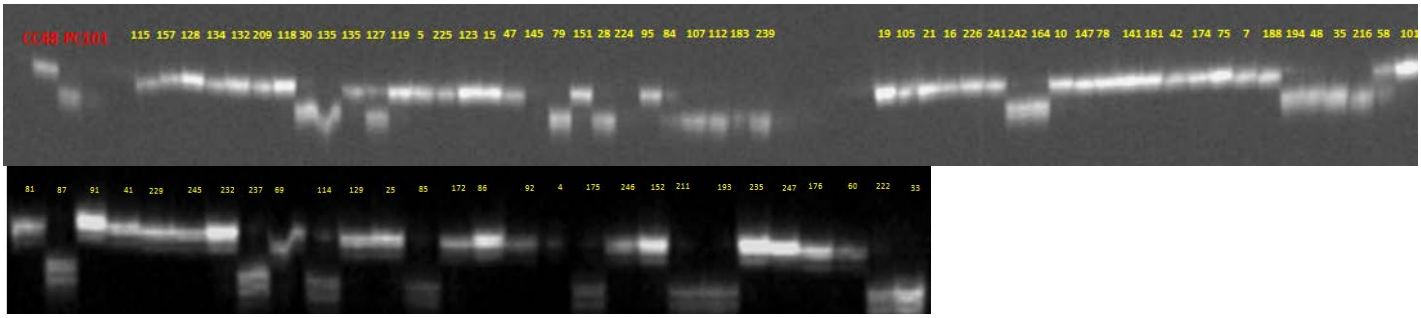
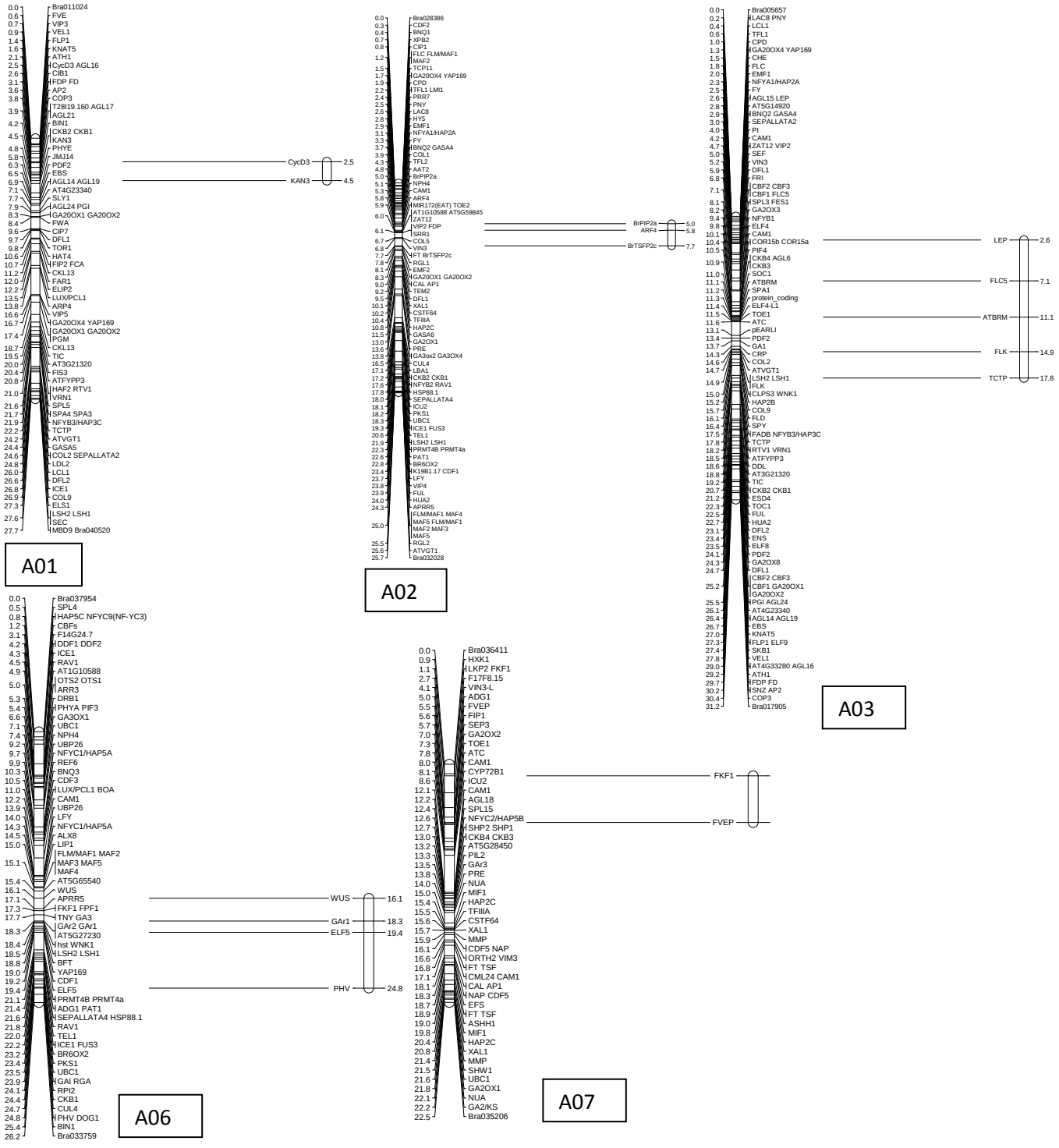


Figure 6. Simple sequence repeat (SSR) marker, BrFLC5P5a_A03, genotyped on DH88 individuals on gel electrophoresis using Li-cor machine. The parental line are on the left in red indicated as CC44 (Chinese cabbage) and PC101 (Pak Choi). The genotype identification numbers are indicated on top of the bands starting with individual 115 up to 191.The individual genotypes were scored according to the parental allele, with CC48, scored as (a) and PC101 scored as (b).

3.7 Detection of linkage groups associated with leaf and flowering time traits

Linkage analysis was undertaken using Join Map 4.1 (Van Ooijen, 2006); primary linkage maps were constructed using the Kosambi mapping function. Linkage groups were determined using LOD threshold of 2. The significance of segregation of the all the markers amounting to 85 was analysed using the Chi-square test, with p = 0.05 and 1 degree of freedom to test for the Mendelian segregation of DH population of 1:1 of parental line alleles (a:b), (a = CC48 and b = PC101). Three markers showed similarity at loci 4 (2) and 10 (1) in the total analysis of the group. All overall, the markers showed high significance of skewness in allele segregation with 59% showing deviation from the expected 1:1 ratio of DH (**Appendix II**). The range of marker ratio skewness (a:b) was from 1.6 to 8.3 in DH88 population. Hence, these markers need to be removed from further analysis or rescoring. Linkage map of the markers was based on the physical location of the markers based on previous mapping in DH68 (Zhao et al., 2010b; Xiao et al., 2012, submitted).

Eight (8) linkage groups were detected designated as A01 till A10 to depict the position of the markers (**Figure 7**). The markers were grouped based on the physical location on the genetic map, and then interval mapping was done to calculate relatedness at the adjacent loci, which resulted in the formation these groups. The first linkage group was detected, designated as A01, another linkage group was AO2. Two (2) linkage groups were observed represented by a combined group A03 with one linkage group representing A06, A07 and A08. While A09 was represented two to linkage groups for markers mapped in on genetic map. No markers representing A04 and A05 were mapped in groups due to few markers assigned to these positions on the genetic map.



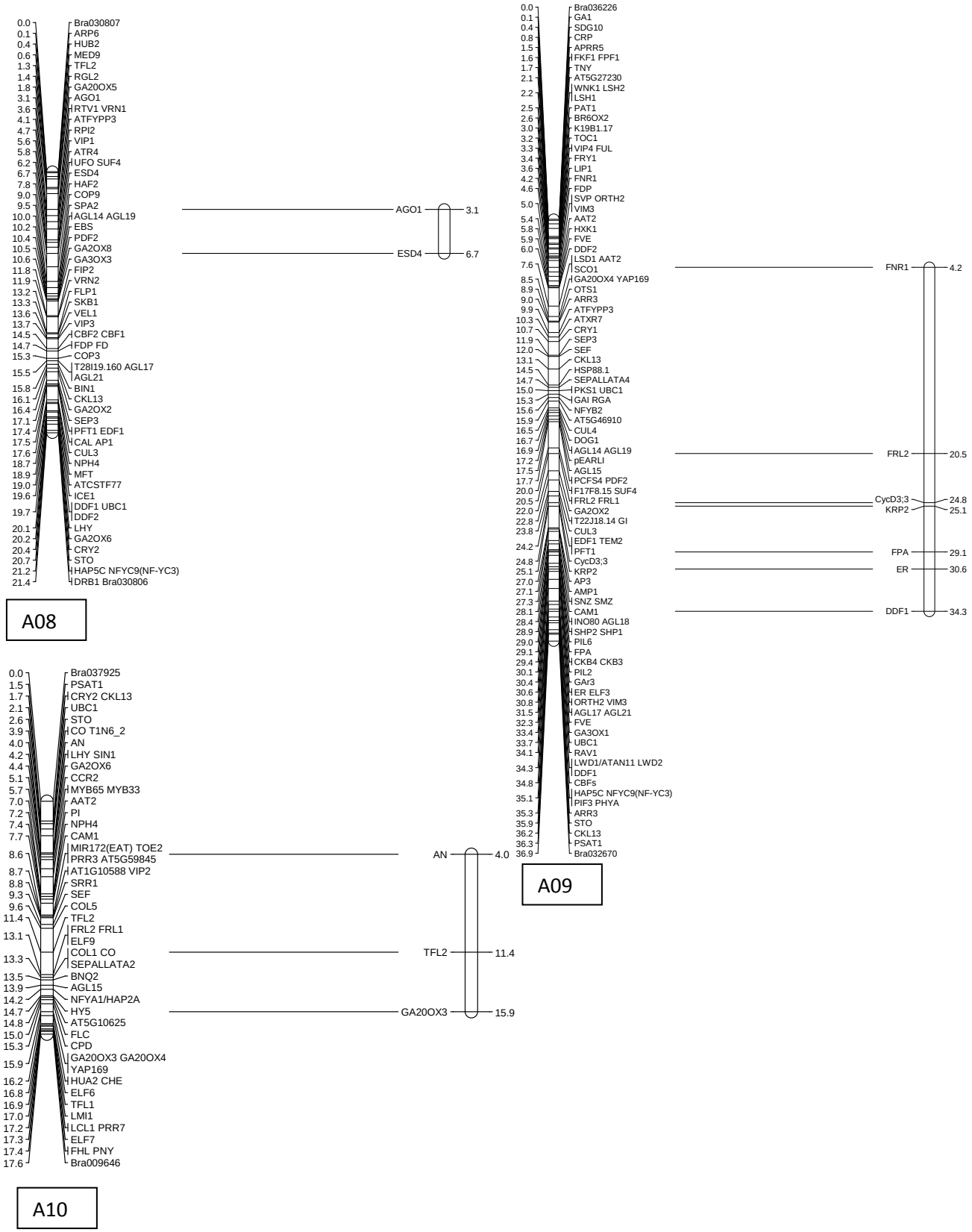


Figure 7. Markers forming linkage groups on the respective genetic map of *B. rapa* with the groups designated as A0 to represent the linkage number. A01 linkage group is for markers physical location on chromosome 1, while A02 is on chromosome 2. Chromosome 3 has two linkage groups. Chromosome A06 has 1 linkage group, same with chromosome 8. Chromosome 9 and has 2 linkage groups, A09.1 and A09.2. Chromosome 10 has one linkage group, A010. The map covered a genetic distance of 521.8 cM and an average of distance of 18.0 cM between markers.

3.8 FQTL analysis of DH88 based on flowering time and leaf trait markers

Linkage association of flowering time markers was done using Kruskal-Wallis nonparametric analysis of variance (ANOVA) based on p-values. According to this test statistic, $p < 0.05$ was used to determine the effect of the markers on flowering time. Markers were ranked with the corresponding genotype for determining association with days to flowering time based on their physical location. The total of 85 markers were used in the analysis (markers with genetic map position and without, but for all markers physical map position was known) of 80 genotypes with their corresponding flowering time for fQTL. Six markers showed an effect on flowering time at $p < 0.05$, (**Appendix V**). These markers are CycD3_1P1a_A01, ARLP1a_A03, CYcD3_3P3a_A09, GA20OX3P4b_A10, CUC2P2a_A010 and BrTL2P2a_A10 (**Appendices V**). The association of the markers with flowering time can be observed with graphical representation showing the critical negative log10 (-log10) transformed p-values at different levels and the markers on the X-axis (**Figure 8A**). FQTL were also observed at $p < 0.01$ associated with ARLP1a_A03, CYcD3_3P3a_A09, CUC2P2a_A010 and BrTL2P2a_A10 (**Figure 8B**), but non at $p < 0.001$. Hence the fQTL were not very strongly associated flowering time. Besides flowering time markers, leaf trait markers were used in order to increase the marker number. Hence, of all the detected fQTL, BrTL2P2a_A10 was the only flowering time marker, well the others were leaf trait related markers (CycD3_1P1a_A01, ARLP1a_A03, CYcD3_3P3a_A09, GA20OX3P4b_A10, and CUC2P2a_A010). BrTL2P2a_10A has been found to control vernalization response in *Arabidopsis* and negative regulation of flowering time. While the other markers are involved in pathways such as gibberellin acid biosynthesis, cell cycle control and endo duplication of chromosomes. In terms of flowering time represented by the markers underlying fQTL in DH88, which can be termed as candidate genes, genotypes having allele "a" accounted for late flowering days compared to those with allele "b". The overall average for allele "a" was 209 d and 170 d for "b". While the overall average difference between the two groups of alleles was 40 d (**Table 3**).

Table 3. Average days (d) to flowering time of candidate genes underlying fQTL in DH88 at $p < 0.05$.

Marker	Allele "a" average (d)	Allele "b" average (d)	Differences (d)
CycD3_1P1a_A01	165	254	89
ARLP1a_A03	143	245	102
CycD3_3P3a_A09	232	92	140
GA0O3P4b_A010	235	163	72
CUC2P2a_A010	227	105	122
BrTL2P2a_A010	254	158	96
Overall average	209	170	40

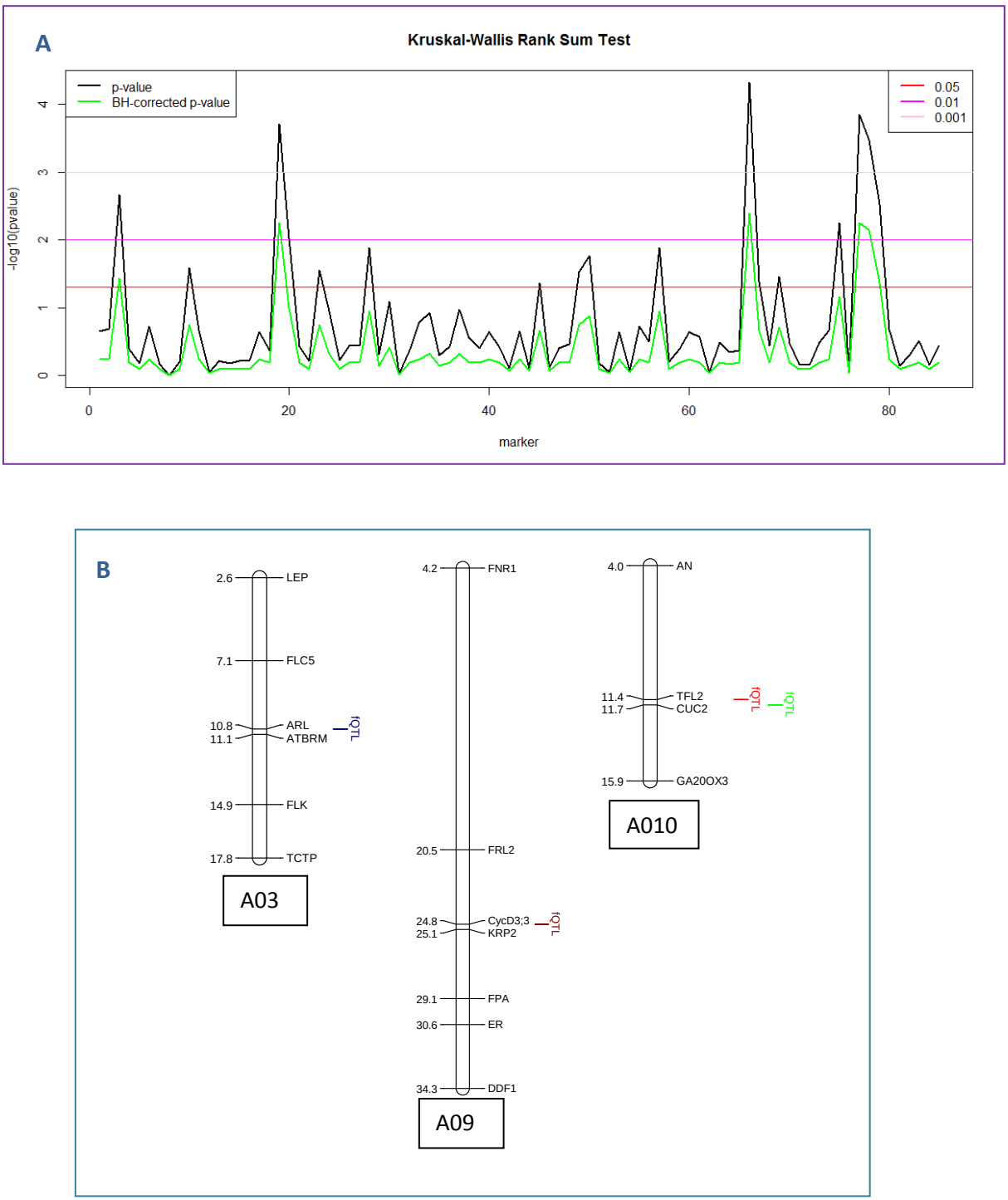


Figure 8A-B. Figure 8A. FQTL analysis using non-parametric analysis of variance (ANOVA) based on Kruskal-Wallis ranking of marker associated with flowering time (FT), which was compared based on days to flowering time (DAT) DH88. Ranking of the markers was done at $p<0.05$. The graph shows the detection fQTL at different p-values, with lines on the negative log 10 (-log10) in the Y-axis transformed p represent $p<0.05$ by red, $p<0.01$ by light red and $p<0.001$ by purple. The x-axis showing the markers groups totalling 85. The green line representing the marker association with fQTL at various corrected differences with the markers and the black for the uncorrected differences. **Figure 8B.** FQTL location on the linkage group at $p < 0.01$ showing the linkage group A03, AO9 and A010.

4 DISCUSSION

The objective of this study was to screen and map flowering time related markers in a *B. rapa* DH population and score the flowering time phenotype in this population. This was with the essence of explaining variation in flowering time using polymorphism of flowering time related markers through the detection of allelic variation in DH88 population. In order to achieve this objective, Chinese cabbage (DH CC48) and Pak Choi (DH PC101) were crossed to produce a DH population. This population was used in genotyping of SNP and SSR markers. Marker development in this population showed fundamental factors that underlie use of molecular markers in map construction. These factors had to do with genetic relationship of the parental lines and scoring of the markers using sensitive equipment.

4.1 Phenotypic variation for flowering time

Variation was observed in as far as flowering time was concerned, with most of the genotypes flowering late after the parental lines. Average flowering time was 202 d and the parental lines had 164 d and 200 d respectively for DH PC101 and DH CC48 (**Figure 2A**). Individual genotypes that did not flower after taking of the data were scored as 300d. In terms of flowering time trait, the transition from vegetative stage to reproduction is very important in annual plants (Poething, 2003). This phase accounts for the ability of the plant to adapt to new environmental conditions. Flowering time of early flowering genotypes of DH88 was the same for all the blocks, only in late maturing genotypes were small difference among the blocks, but less than 10 days for most of them. Most transgressive segregation for this phenotype was towards late maturity with only a few individuals for early maturity (**Figure 2A**). Flowering time did not follow normal distribution because most individuals flowered late; this affected the mean of the population to shift above all the parental lines (**Appendix VI**). The range of days to flowering time was from 33 d to 300 d, showing very big differences existed among the genotypes.

4.2 Polymorphism rate in *B. rapa* DH88 population

Pre-screening of DH88 parental lines DH CC48 and DH PC101 was the first step taken in the detection of polymorphism. In this regard, a total of 147 markers were screened on the parental lines resulting in 29 markers showing distinct sufficient polymorphism, which translated in 20% of polymorphism rate. This low rate of polymorphism at many flowering time loci was evident with the BrFLC2 locus that encodes a MADS-Box transcription factor (Zhao et al., 2005). BrFLC2 (**Figure 3B**) being non-polymorphic on both parents. This marker has been found to underlie many fQTL in *B. rapa* (Luo et al., 2007; Zhao et al., 2010b). Hence, it could be argued that DH CC48 and DH PC101

were related. Phylogenetically, the genetic distance between Pak Choi and Chinese cabbage is very small compared to other *B. rapa* sub species (Zhao et al., 2005b). However, in terms of flowering time, the two shows a lot of phenotypic variation with Pak Choi being early flowering than Chinese cabbage. Hence, it was on this premise that polymorphism was envisaged to be sparingly sufficient for flowering time markers underlying fQTL due to few markers that have been observed to underlie the loci close to fQTL (Luo et al., 2007; Zhao et al., 2010b, Li et al., 2007). Polymorphism rate between DH CC48 and DH PC101 is very low for the genome segment targeted in DH88 population by the flowering trait markers in this study. In the screening of the same SNP markers used in DH88 on diverse populations of *B. rapa* for polymorphism, different rates of polymorphisms were observed (Zhang, unpublished). In a cross between vegetable turnip and Rapid Cycling (VT-115 X RC-144) in population designated as RIL36, polymorphism was found to be 63% (Zhang, unpublished). In another population designated as C29-F4, which was generated from the crossing of vegetable turnip and Pak Choi (VT-115 X PC-105), polymorphism rate of 51% was observed (Zhang, unpublished). In a DH population, DH68 which was a cross between Yellow sarson and Pak Choi (YS-143 X PC175), polymorphism rate of 69% was observed, which is higher than other populations derived from other genotypes (Zhang, unpublished). Also using DH88 population, Zhang found that the polymorphism rate was low with overall rate of 20%. These results shows that the polymorphism rates of flowering time related markers and random markers show the same pattern, this is due to different rates which are independent of the population.

In this study, even when polymorphism was detected on the parental lines of DH88, many markers could not show clear polymorphism on the population. The population showed a lot of markedly low rate of polymorphism overall when genotyped with SNP markers. This low rate of polymorphism was observed with the Light Scanner machine (Idaho Technology Inc.) that uses HRM to detect the melting curves of nucleotides. The low rate of polymorphism resulted in the HRM generated failing to separate the allelic variation representing the two parental genotypes into two distinct groups. Hence, a lot of SNP markers that were deemed as suspect were not used in the analysis of fQTL **(Figure 5**). In this regard, a total of 70 markers were genotyped on DH88 population as having polymorphism, resulting in 46 of them being dropped due to clustering together of the individuals in one group. This is because these markers were not informative on the polymorphism in the population as they could represent any of the parental line. The clustering of individual markers could only be explained with regards to the nucleotide composition of the particular alleles detected by the markers. This is because Light Scanner detects SNP based on the melting curves of nucleotides that is reflected by the hydrogen bonds of the bases, with differences in guanine (G) and cytosine (C) (G≡C) and adenine (A) and thymine (T) (T=A). The higher the G and C content in the SNP

segment, the more the differences compared to that of A and T containing segment. The sure way of identifying the allele of the various genotypes that were not separated by the light scanner is to sequence the amplified regions by the markers of the genotypes and parental lines and then making a contig alignment with either of the parent in order to detect the base pair difference at the SNP. This is a direct way in that looks at the genome and depicts the mutation present in the marker, which can explains the difference between the different genotypes with the parents+. If the genotype belongs to one of the parental allele, they will be no difference at the marker region containing the SNP.

However, despite the lack of clear separation for most markers, the sensitivity of the machine was compensated by manual scoring of the markers fitting one of the parental allele despite being grouped differently. On the other hand, individuals not showing any melting profiles of the parental lines were labelled as unknown and these were not included in the analysis of the markers. This was done in order to minimise unequal segregation of the alleles so as to maintain a 1:1 of DH cross. The scoring of SNP marker using the Light Scanner requires manual calibration of sensitivity in order to balance the accuracy of the results obtained.

Using SSR markers on DH88 population did not give any significant advantage compared SNP in as far as polymorphism was concerned. This is because SSR markers also showed rates of polymorphism on the parental lines of DH88. A total of 41 markers were scored of which 5 showed clear polymorphism that was scored on the population. Available polymorphism between the parental lines was equivalent to 32%. The markers showed clear bands but also segregation was not 1:1 (**Figure 6; Appendix III**).

4.3 Leaf trait and flowering time markers showing potential for QTL mapping in DH88

In order to understand the feasibility of developing marker for mapping of leaf traits and flowering time using DH88, linkage analysis was done using Join Map 4.1. A total of 85 markers and 80 individuals were used in this analysis. Markers of both flowering time and leaf traits showed 59% significant skewness and could not fit the 1:1 segregating ratio of DH cross. The markers were skewed to one of the parental alleles, DH CC48 and DH PC101. The distortion in segregation ratio of the markers (a:b) ranged from 1.6 to 8.3 (**Table 2; Appendix II**). These distortions in segregation underlie the pitfall associated with the use of DH derived population in mapping of QTL (Cheng et al., 2009; Wu et al., 2008; Takahate and Keller, 1991). The distortion in segregation has been attributed to preferential selection of certain genotypes responsive to microspore or anther culture (Uzunova et al., 1995; Chuong et al., 1998; Foisset et al., 1996). The other reasons are genetic isolation

mechanism, presence of viability genes and occurrence of gene conversion events (Zamir and Todmor, 1986; Beavis and Grant, 1991; Nicholas and Rossignol, 1983; Nag et al, 1998). This distortion can also be reduced if genotyping of markers is also done with clear separation of the individuals in distinct groups. This may only be possible if markers are able to generate distinct groups. Skewness in markers makes it difficult to generate linkage groups as they interfere with the estimation of relatedness at adjacent loci (Lorieux et al., 1995). However, the smaller number of markers generated made it difficult to assign them to linkage groups on the genetic map. Linkage group 9 had the largest linkage group, A09.1 and a smaller one A09.2 with only two markers which were combined to form A09 (**Figure 7**). In this analysis only chromosome 4 and 5 did not have linkage groups owing to the small number of markers genotyped for these chromosomes.

FQTL analysis resulted in 6 markers associating with flowering time at $p < 0.05$ (**Figure 8A, Appendix V**). When using a strict level ($p < 0.001$), no association of markers with flowering time was detected (**Figure 8**). Looking at the genotyping of DH88 population, it is clear that more markers are needed for complete mapping of the loci needed for linkage analysis of leaf trait and flowering time QTL. This requires a large platform for marker profiling such as microarray system due to low levels of polymorphism at many loci.

The markers that were found to underlie fQTL in DH88 population are involved in various plant biological processes. These candidate genes significant at $p < 0.05$ had one flowering time candidate gene, BrTL2P2a_10, while the others (CycD3; 1P1a_A01, BrARLP_A03, CYcD3_3P3a_A09, GA20OX3P4b_A10, and CUC2P2a_A010) are leaf trait associated candidate genes (**Table 1**). BrTL2P2a_10 has been characterized to control response to vernalization and negative regulation of flowering. CycD3; 1P1a_A01 has been found to be important in the regulation of cell cycle, meristem growth and complex morphogenesis. On the other hand, BrARLP1a_A03 is involved in brassinosterioid mediated pathway and cell cycle growth. BrCyCD3; 3P3a_A09 is involved in DNA endo-reduplication, mitotic cell cycle and negative regulation of cell cyclin-dependent protein kinase activity and regulation of cell proliferation and stomatal lineage progression. BrGA20OX3P4b_A10 is involved in gibberellin biosynthesis, oxidation-reduction processes and terpenoid biosynthesis processes. BrCUC2P2a_A010 is involved in organ boundary demarcation, leaf development and morphogenesis.

BrTL2P2a_10 has been found to be associated with negative control of gene expression of FLC that encodes a MADS-BOX transcription factor that represses flowering, this gene also is involved in negative control of flowering through epigenetic control by methylation of genes that positively control flower development (Sheldon et al., 1999; Michaels, and Amasino, 1999; Koornneef et al.,

2004; Luo et al., 2007; Ehrenreich et al., 2009). Hence it can explain the variation observed in late flowering plants difference with early flowering as early flowering plants not expressing this gene are more likely to flower early than late flowering. The other genes, are have not been categorically found to directly associate with QTL, except for BrGA20OX3P4b_A10 which has been found to underlie major leaf trait QTL in *B. rapa* (Li et al., 2007). This gene is involved in gibberellin biosynthesis in plants which account for leaf lobe differences among genotypes. The leaf trait candidate genes association with flowering QTL shows the modular effect of QTL control, with one QTL having an effect on other related process (Koornneef et al., 2004). Also the fact that flowering time is under the control of different genes underlying various pathways in plant shows the leaf trait candidate genes associated with fQTL.

The allele representing the parental lines in the fQTL identified in DH88 showed that "a" representing the male parent, DH CC48 accounted for late flowering time, with overall average of 209 d for this allele (**Table 3**). Allele "b" that represented the female parent, DH PC101 had an overall average of 170 d, this allele accounted for early flowering (**Table 3**). The results clearly shows what has been found with cultivar groups within *B. rapa* where Chinese cabbage and Pak Choi differ in flowering time, with the former been late maturing than the later (Zhao et al., 2005a). The candidate gene marker, CycD3_3P3a_A09 had the highest differences between the alleles for flowering time of 140 d, while the candidate gene GA0O3P4b_A010 had the smallest differences (**Table 3)**. The flowering time candidate gene, BrTL2P2a_A010 had the highest average for allele "a" (254 d), showing its effect on negative control of flowering time. Incidentally the highest average for allele "b" was 254 d by the candidate gene, CycD3_1P1a_A01 that is associated with leaf trait in *B. rapa*.

5 Recommendations

In view of the low polymorphism rate found in DH88 population, the following recommendations should be considered:

- A large platform is needed to genotype markers in DH88 such as microarray due to low polymorphism. This may involve the use of systems such the golden gate and infinium rays.
- Need to calibrate the Light Scanner machine to sensitivity of a particular crop based on nucleotide content in DNA.
- Exploring the genome of *B. rapa* with other markers such as SSR markers and sequence characterized amplified regions (SCAR) markers that are located in various segments of the genome.
- Manual scoring of markers when using light scanner to avoid putting markers in different genotype group which result in increased skewness.

REFERENCES

Asp, T., Frei, U., K., Didion, I., Nielsen, K., K., Lubberstedt, T. (2007). Frequency, type and distribution of EST-SSRs from three genotypes of *Lolium perenne* and their conservation across arthologous sequences of *Festuca arundinacea*, *Brachypodium distachyon and Oryza sativa.* BMC Plant Biol. 7: 35 – 47.

Beavis, W., D., and Grant, D. (1991). A linkage map based on information from 4 F2 populations of maize (*Zea mays* L.). Theor. Appl. Genet 82: 636 – 644.

Bonnema, G., Carpio, D., P., D., and Zhao, J., J. (2011). Diversity analysis and molecular taxonomy of Brassica vegetable crops. Book Chapter in Brassica vegetables, editor Jan Sadowski and Chittarananajan Kole. In the series genetic, genomics and breeding of crop plants. Series editors: C., Kole and A., G., Abbot.

Brachi, B., Faure, N., Horton, M., Flahauw, E., Vazquez, A., Nordborg, M., Bergelson, J., Cuguen, J., and Roux, F. (2010). Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLOS genetcs 6.

Cheng, Y., Geng, J., Zhang, J., Wang, Q., Ban, Q., and Hou, X. (2009).The construction of a genetic linkage mapof non-heeding Chinese cabbage (*Brassica campestris* ssp. *Chinensis* Makino. J. Genet. Genomics 36: 501 – 508.

Ching, A., Cadwell, K., S., Jung, M., Morgante, M., and Rafalski, A., J. (2002). SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbredlines. BMC Genet 3:19

Choi, S, R., Teakle, G., R., Plaha, P., et al. (2007). The reference genetic map for the multinational *Brassica rapa* genome sequence project. Theoretical and Applied Genetics 115, 777 – 792.

Chuong, P., V., Deslauriers, C.,Kott, L., S., and Beversdorf, W., D. (1988). Effects of donor genotype and sampling on microspore culture of *Brassica napus*. Can. J. Bot 66: 1653 – 1657.

Cloutier, S., **Cappadoria, M., and Laundry, B., S.** (1995). Study of microspore culture responsive in oilseed rape (*Brassica napus* L.) bvy comparative mapping of F2 population and two microspore-derived populations. Theor. Appl. Genet 91: 841 – 847.

da Maia, L., C., Palmieri, D., A., de Souza, V., Q., Kopp, M., M., de Carvalho, F., I., F., and de Oliviera, A., C. (2009). Tandem distribution of gene transcripts in three plant families. Genet. Mol. Biol. 32: 822 – 833.

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

Engelmann K., and Parugganan, M. (2006). The molecular evolutionary ecology of plant development: flowering time in Arabidopsis thaliana. Advances in botanical Research 44, 507 – 526.

Ehrenreich, M., I., Hanzawa, Y., Chou, L., Roe, L., J., Kover, X., P., and Purugganan, P., M. (2009). Candidate gene association mapping of Arabidopsis flowering time. Genetics 185: 325 – 335.

Foisset, N., Delourme, R., Barret, P. (1996). Molecular-mapping analysis in *Brasssica napus* using isozymes, RAPD, RFLP markers on double haploid progeny. Theor. Appl. Genet 93: 1017 – 1027.

Fornara, F., de Montaiga, A., Coupland, G. (2010). Snapshot: control of flowering in *Arabidopsis.* Cell 141: 550-550.e552.

Fujimori, S., Washito, T., Higo, K., Ohitomo, Y., Kikuchi, S., and Tomita, M. (2003). A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription. Febs Lett 554: 17 – 22.

Gao, C., Tang, Z., Yin, J., An, Z., Fu, D., and Li, J.,(2011). Characterization and comparison of gene-based simple sequence repeats across Brassica species. Mol. Genet. Genomics 286: 161 - 170.

Gao, H., Cai, S., L., Yan, B., L., Chen, B., Y., and Yu, F. (2009). Discrepancy variation of repeats in eukaryotic genomes. Bios Res. 42: 365 – 375.

Goldstein, D., B., and Schlotterer, C. (1999). Microsatellite: evolution and application. Oxford university Press, Oxford.

Gut, I., G. (2001). Automation in genotyping of single nucleotide polymorphism. Human Mutat 17: 475 – 492.

Hall, T. (1999). Bio Edit: a user friendly biological sequence alignment editor and analysis program for Window 95/98 NT. Nucl. Acids Symp. Ser 41: 95 – 98.

Hong, C., P., Piao, Z., Y., Kang, T., W., Bhak, J., Park, B., S., Edwards, D., and Lim, Y., P. (2007). Genomic distribution of simple sequence repeats in *Brassica rapa.* Mol. Cells 23: 349 – 356.

Jander, G., Norris, S., R., Rounsely, S., D., Bush, D., F., Levin, I., M., and Last, R., L. (2002). *Arabidopsis* map-based cloning in the post genome era. Plant Physiol 129: 440 – 450.

Jiang, D., Zhong, G., Y., and Hong, Q., B., C. (2006). Analysis of microsatellite in citrus unigene. Yichuan Xue Bao 33: 345 – 353.

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

Kole, C., Quijada P., Michaels, S. D., Amasino, R., M., Osborn, T., C. (2001). Evidence for homology of flowering-time genes VFR2 from *Brassica rapa* and FLC from *Arabidopsis thaliana*. Theoretical and Applied genetics 102, 425 – 430.

Komeczyny, A., and Ausubel, F., M. (1993). Procedure for mapping Arabidopsis mutations co-dominant ecotype-specific PCR based markerd. Plant J. 4: 403 – 410.

Koornneef, M., Alonso-Blanco, C., and Vreugdenhil, D. (2004). Maturally occuring genetic variation in *Arabidopsis thaliana*. Plant Biology 55, 141 – 172.

Kosambi, D., D. (1944). The estimation of map distance from recombination values. Ann. Eugen 12: 172 – 175.

Kumpatla, S., P., and Mukhopadhyay, S. (2005). Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. Genome 48: 985 – 998.

Kwok, P., Y. (2001). Methods forgenotyping single nucleotide polymorphism. Annu Rev. Genomics Hum. Genet 2: 235 – 258.

Lagercrantz, U. Osterberg, M., Lascoux, M. (2002). Sequence variation and haplotype structure at the putative flowering-time locus COL1 of *Brassica nigra*. Molecular Biology Evolution 19, 1474 – 1482.

Lawson, M., J., and Zhang, L. (2006). Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. Genome Biol. 7: R14.

Li, F., Kitashiba, H., Inaba K., and Nashio, T. (2009). A *Brassica rapa* linkage Map of EST-based SNP markers for identification of candidate genes controlling flowering time and leaf morphological triats. DNA Research 16, 311 – 323.

Li, Y., C., Fahima, T., Roder, M., S., Kirzhner, V., M., Beiles, A., Korol, A., B., and Nevo, E., (2003). Genetic effects on microsatellite diversity in wild emmer wheat (*Triticum diccoides*) at Yehudiyya microsatellite, Israel. Heredity 90: 150 – 156.

Liu et al., (2012). An improved allele-specific PCR primer design methods for SNP marker analysis and its application. Plant Methods 8: 34.

Luo, P., Zhao, J., Kim, J., S., Shen, S., Del Carpio, D., P., Song, X., Jin, M., Vreugdenhil., D., wang, X., Koornneef, and bonnema, G. (2007). Quantitative trait loci for flowering time and morphological traits in multiple populations of Brassica rapa. Journal of experimental Botany 58, 4005 – 4016.

Marrinez-Zapater, J M., and Somerville, C., and R. (1990). Effect of light quality and vernalization on late flowering mutants of Arabidopsis thaliana. Plant Physiology 92, 770 -776.

Mendez-Vigo, B., Teresa de Andres, M., Ramiro, M., Martinez-Zapater, J., M., and Alonso-Blanco, C. (2010). Temporal analysis of natural variation for the rate of leaf production and its relationship with flowering initiation in *Arabidopsis thaliana*. Journal of experimental Botany 61, 1611 – 1623.

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

Montgomery, J. Wittwer, C., T., Palais, R., Zhou, L. (2007). Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. Nature Protocols 2: 59 – 66.

Mouradov, A. cremer, F., and Coupland, G. (2002). Control of flowering time: interacting pathways as the basis for diversity. The Plant Cell 14, (Suppl.) S111 – S130.

Nag, D., K., White, M., A., and Petes, T., D. (1989). Palindromic sequences in heteroduplex DNA inhibit mismatch repair in yeast. Nature 340: 318 – 320.

Neff, M.M., Turk, E., and Kalishman, M. (2002). Web-based primer design for single nucleotide polymorphism analysis. Trends Genet 18: 613 – 615.

Nicolas, A., and Rossignol, J., L. (1983). Gene conversion: point mutation heterozygosities lower hetroduplex formation. EMBO J. 2: 2265 – 2270.

Okazaki, K., Sakamoto, K., Kikuchi, R., Saito, A., Togashi, E., Kugimuki, Y. Matsumoto, S., and Harai, M. (2007). Mapping and characterization of FLC homologs and QTL analysis of flowering time in *Brassica oleracea.* Theoretical and Applied Genetics 114, 595 – 608.

Parida, S., K., Yadava, D., K., and Mohapatra, T. (2010). Microsatellite in Brassica unigenes relative abundance, marker design, and use in comparative physical mapping and genome analysis. Genome 53: 55 – 67.

Poethig, R., S. (2003). Phase change and the regulation of development timing in plants. Science 301, 334 – 336.

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

Ramchiary, N., Yu, G., Wang, T., Liang, C., Wang, N., Wang, Z.,Choi, S., R., Lim, Y., P., and Piao, Z., Y. (2001). Mapping Quantitative Trait Loci for leaf and heading-related traits in Chinese cabbage (*Brassica rapa* L. Ssp. *pekinensis*). Hort. Environ. Biotechnol 52: 494 – 501.

Roux, F., Touzet, P., Cuguen, J., and Le Corre, V. (2006). How to be early flowering: an evolutionary perspective. Trend in Plant Science 11: 375 – 381.

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

Seo, E., Lee, H., Jeon, J., Park, H., Kim, J., Nohc, Y., S., Lee, I. (2009).Croostalk between cold response and flowering *Arabidopsis* ismediated through the floering-time gene SOC1 and its upstream negative control FLC. The Palnt cell 21, 3185 – 3195.

Sheldon, C., C., Burn, J., E., Perez, P., P., Metzger, J., Edwards, J., A., Peacock, W., J., Dennis, E., S. (1999). The FLC MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. The Plant Cell 11, 445 – 458.

Steynen, Q., J., Bolokoski, D., A., Schultz, E., A. (2001). Alteration in flowering time causes accelerated or decreased progression through vegetative phase. Canadian Journal of Botany 79, 657 – 665.

Subramanian, S., Madgura, V., M., George, R., Kumar, S., Pandit, M., W., and Singh, L. (2003). SSRD: simple sequence repeats database of the human genome. Comp. Funct. Genome 4: 342 – 345.

Suwabe, K., Iketami, H., nunome, T., Kage, T., and Harai, M. (2002). Isolation and characterization of microsatellites in *Brassica rapa*. Theoretical and applied Genetics 104, 1092 – 1098.

Takahate, Y., and Keller, W., A. (1991). High frequency embryogenesis and plant regeneration in isolated microspore culture of *Brassica oleracea* L. Plant Sci 74: 235 – 242.

Teale, W., D., Paponov, I., A., and Palme, K. (2006). Auxin in action: signalling, transport and the control of plant growth and development. Nature Reviews Molecular Cell Biology 7, 847 – 859.

Thiel, T., Kota, R. grosse, I., Stein, N., and Graner, A. (2004). SNP2: a SNP and INDEL analysis tool for CAPS marker development. Nucleic Acids Res 32: e5.

Town, C., D., et al., (2006). Comparative genomics of *Brassica oleracea* and Arabidopsis thaliana reveal gene loss fragmentation and dispersal after polyploidy. The Plant cell 18, 1348 – 1349.

Ueno, S., Taguchi, Y., and Tsumura, Y. (2008). Microsatellite markers derived from *Quercus mongolica* var. *crispula* (Fagaceae) inner bark expressed sequence tags. Genes genet. Syst 83: 179 -187.

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

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

Van Oijen, J., W. (2006). JointMap software for the calculation of genetic linkage maps (Kyazma BV, Wageningen, The Netherlands) Version 4.1.

Vigouroux , Y. Jaqueth, J., S., Matsuoka, Y. Smith, O., S., Beavis, W., D., Smith, J., S., and Doebley, J. (2002). Rate and pattern of mutation at microsatellite loci in maize. Mol. Boil. Evol 199: 1251 – 1260.

Vogouroux, Y. Matsuoka, Y., and Doebley, J. (2003). Directional evolution for microsatellite size in maize. Mol. Biol. Evol 20: 1480 – 1483.

Wang, X., wang, H., Sun, R., Wu, J., Liu, S., bai., Y., Mun, J-H, Bancroft, F., Cheng, F., Huang, S., et al. (2011). The genome of the mesopolyploid crop *Brassica rapa*. Nature genetics 43, 1035 - 1040.

Wu, J. Yuan, Y., X., Zhang, X., W., Sun, R., F., Koornneef, M., Aarts, M., G., M., and Wang, X., W. (2008). Mapping QTL for mineral accumulation and shoot dry biomass under different Zn nutritional conditions in Chinese cabbage (*Brassica rapa* L. Ssp. *Pekinensis*). Plant Soil 310: 25 – 40.

Xiao, D., Basnet, R., Zhang, N., Carpio, D., Lin, K., Cheng, F., Hou, X., Wang, X., and Bonnema, G. (2012). Genetic architecture of flowering time regulation in a *Brassica rapa* double haploid population. (Unpublished).

Yang, Y., W., Lai., K., N., Tai, P., Y., and Li, W., H. (1999) rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between Brassica and other angiosperm lineage. Journal of Molecular Evolution 48, 597 – 604.

Yu, S., C., wang, and Y., J., Zheng, X., Y. (2003). Mapping and analysis QTL controlling some morphological traits in Chinese cabbage (*Brassica camprestis* spp. *pekinensis*). *Acta genetica Sinia* 30, 1153 – 1160.

Yuan, Y., X., Wu, J., Sun, R., F., Zhang, X. F., Xu, D., H., Bonnema, G., Wang, X., W. (2009). A naturally occuring splicing site mutation in the *Brassica rapa* FLC1 gene is associated with variation in flowering time. Journal of experimental Botany 60, 1299 -1308.

Zamir, D., and Tadmor, Y. (1986) Unequal segregation of nuclear genes in plants. Bot. Gaz 147: 355 – 358.

Zhang, L., D., Yuan, D., J., Yu, S., W., Li, Z., G., Cao, Y., F., Miao, Z., Q., Qian, H., M., and tang, K., X. (2004). Preference of simple sequence repeats incoding and noncoding regions of *Arabidopsis thaliana*. Bioinformatics 20: 1081 – 1086.

Zhang, L., D., Zuo, K., J., Zhang, F., Cao, Y., F., Wang, J., Zhang, Y., D., Sun, X., F., Tang, K., X. (2006). Conservation of noncoding microsatellites in plants: implication for gene regulation. BMC genomics 7: 233 – 336.

Zhang, N. Screening of diffrent population of *B.rapa* with flowering time SNP markers for polymorphim. Unpublished.

Zhao, J., J., wang, X., W., Deng, B., Luo, P., Wu, J., Sun, R., F., Xu, Z., Y., Vromans, J., koornneef, M., Bonnema, G. (2005a). Genetic relationships within the *Brassica rapa* as inferred from AFLP fingerprints. Genetics 110: 1301 -131.

Zhao, J., Kulkarim, V., Del Capio, D., P., Buchner, J., and Bonnema, G. (2010b). BRFLC2 (FLOWERING LOCUS C) as a candidate gene for vernalization responseQTL in *Brassica rapa.* J. Exp. Bot 61: 1817 – 1825.

APPENDICES

Appendices I: Primers of flowering time markers and leaf traits

Maker name	Forward primer sequence(5'-3')	Reverse primer sequence(5'-3')
BrAGL15P2a	CGAACGCTTTTGAGATACGAC	CACACAAAAGTGGTGGATGG
BrFDPP1a	GCCAGCCTTCTCCATATAACC	TCGGCTCGTAATGTTATCAAAC
BrAP2P1d	TATGCGGCTCAAGGTAAAGTG	GTTCTCAGCCACCGGAAATAG
BrAP3P1b	GCTCTTGCTCAGTTTGTTTCCT	TAGCAAGCGATTCCAGAGATG
BrAPRR5P1a	CGCGTTGAGAGTTTTGCTAGT	TCAGACAGAAGCCAACTGACC
BrARP4P1a	AAGTGAGGCGATGGATGTG	AACGCATGTTCCCAAATGTT
BrAtBRMP1d	CAGCCAATACAACTGTGTGCTT	GCAATCCAACCTGAATGACC
BrATVGT1P1a	GGCTGCATTGCTACTCTATGTG	CAGAGGTGAGAAAGCGAATGT
BrCALP1g	CTTTTGAGAGACAGACGAACTGG	GACGGAGTCGTACATAAGTTGGTT
BrCALP2a	AGATACTTGAACGCTATGAGAGG	CTTTAGGGCTCATTGCTTG
BrCBFsP2b	GATGGTTAGGTTCCCGATGA	CTTCTTCAGAAAACCGCTCAA
BrCCA1P3b	TCTTTTCAGACGACCAATGAAG	CTATTCTCGTCTGGAGCATGTG
BrCDF1P2e	CCTACCAAGATTATTCCCTGTCC	GGTGGCGTGCTGGTAGTTAT
BrSLY1P1b	TGAAGAAAAAGAGGATGCGG	GAAAGTTGTGTTGGCTTTGTTG
BrCOP1f	TGGTCAATACAATCAACCTACTCAAC	TCTCTATCTTCCCTTTGTGCGT
BrCRY1P1c	TTGGCGATTCTACCGAGTTT	TCTGCTATCTCCTCCAGAGTTTC
BrCRY2P1a	TGAAGTCTCTTGGCTCCCAC	GCATAGAGATCCCACGTTCC
BrCRY2P2c	CAAGATGGTGGATGAAGCAA	TCTCCCACGGCTCATACAAC
BrCRY2P3a	TGGTCTTTGCCCGTCTTC	TTCTGACACCAGCGAGCA
BrDRB1P1d	ATGAACTACGCCATTCCACTCT	TGCTTTTGGAGTCTTGATGGT
BrEBSP1b	TGTGAAGCTGGAGAGTGTGTG	AATACTTACACAGCAACTCGGTCA
BrELF3P1c	CCATGTAACCCAAATGTCCTC	TCACTCCAGTTTCATCACCATC
BrELF5P1a	TGGAATGTACGGTGGTCATCT	AGTCTTTGGCTTTGGTTTGGT
BrELF8P1b	GAACAAGCTCCTCTCGATCTCT	CCTTGTGCTTAGTCTCGGTTTT
BrESD4P2b	GCATGGCTAAATGACGAGGT	CCAACTGTCTCTGCGTAGTCC
BrFDPP1a	GACCAGAAGATGACGCTGATG	TGGCTTCTTTTCTTCCCAAAC
BrFDPP2a	ACCTCCACCTTCGTCTTCCT	TCTTCCCAAAGCAACCAAAC
BrFES1P2c	TGGAAATGGAAAGAGGGAGTC	GATGGAACAGATAACGCTTGCT
BrFKF1P2d	CGTTTGGGGGTACTGTTGAG	AGCCAAGGTCCAACACAAAC
BrFKF1P3c	ATATACCGACGTGGAGGGAG	GCTGAGAAGCCGAGTCAATC
BrFLC2-2(XSA02)	CAAGCGAATTGAGAACAAAA	GAAGCGAAAAGGGCAAATAA
BrFLDP1c	CAGGTATGTATGAGCCACAAGG	GTCTCTGAACAAATCCGCAAG
BrFLKP2b	CTCCTGAGATGCGTCAACCT	TTACAGTTGCTCCACTAAGACGAC
BrFLMP1b	GTCTCCGCCTCTGGAAAACT	AACACTGAGCACGTATCACCAC
BrFLMP2d	CTTCAGTCAAAAGCTCCGAAGT	TTTCTTTGAGGCTATCAACAAGC
BrFPAP1c	CCTCTACACACTTTGCGTCTTC	TACTTCGCAAAACCTAGCAACC
BrFPF1P2a	TAGTGGAGAACCCTAACCAGTCA	AAGTACCTCTCCCAGCCAAGA
BrFRIP2c	TCTCAAGAGCAAAGGCATCAC	CATTCCAACACGAAGTTAGCAG
BrFRL1P2a	GGCGTTGTTGTTTCTTCACTTAC	CGTAACCTTTCAACACAGAGACC

Maker name	Forward primer sequence(5'-3')	Reverse primer sequence(5'-3')
BrFRL2P1b	CCTCTCTCCAGTCCTCACTCTCT	GATGGCGTCAAGAACCAAAG
BrFRY1P1b	AGGATACCGTGAGAAGATTTGG	TCGTTAGCAACTATAATCCCTGTGT
BrFYP1a	TCTTATCACTGGCTCCCAAAG	ACATACTTTAGCGTGCCTCCA
BrGAr3P1a	AAGGAGTCAAGTTCCCCAAGA	GGAGATGACTTGCTTTGACTATCC
BrGAr3P2b	CCTGCTTGTAATCCCTTTGG	AGTCTTCTTGAGACCATCGACAT
BrGIP3b	ATGGTCCTGAAGTGGAGTGG	TGATTGTCTTGATTCTCCTCTGC
BrHUA2P1b	GAGGTATGATGGTCGGTTGTTT	GCCTCATTTTCCAACTTTCGT
BrLKP2P1a	TGAGATTCCGGTCTAACGATG	TGGTGGACTTCCCTTAACACTC
BrMAF4P2a	GGCAGAGCTAATGATGGAGTTT	GCCAAAACCTTGTTCTCTTCC
BrMFTP2b	ATCACGATTGCAGAGAGCAGT	GAGTTAATCGCTGTCCTCGGT
BrPFT1P3b	TCGTCGGCTGGACAAAGT	CTGTCCACCAGAATGACTTCC
BrPHYAP1g	TGAACCTTGGGATTGTGCT	GTCCCCTCCATCATCTTCCT
BrPHYEP1b	ACCTAGCGTCACCGTTTTTCT	CATCGGAGGAAGACATTTTGA
BrPHYEP2c	CTGGGATTATGGACCTTGTGA	ATCCTTATCCTTGGGGTGATG
BrPIP2a	TGGCAGGAGAAAATGTTGGT	TGCCCATTTTGATCTCTCATC
BrFVEP2a	CCGGAAACCTAAGATCAAGGAG	ACACTCTGCTTCGTCTTCTTCC
BrFVEP3a	AAGAAAACGAAGCAGGAGGAG	GACGCTGTCGATTCTTGTAGG
BrRGAP2c	CCCAGTAGCCAAGACAAGGTT	GGTTGTTATCAGCGGTCGAG
BrSLY1P1b	CGAGTAACAAGAAGATGAAAAAGACC	TCGAGAGGAGAGAAAGGGAGA
BrSNZP2b	GTTACGTTTTACCGGCGAAC	CTGCATGAGCAGTGTCAAATC
BrSPA1P1a	AGTTTTGACGCTGAAGAGGAAC	ACGCACCTGAACTACACCATC
BrSPL5P2b	GGCAAAAAAGCACAAGGACT	GAAGATGCCTTTGCATGAACT
BrSVPP1c	GCCCTGTTGAGCAAAGAGATT	CGTTGTTTTGCTTCTTACCTTTC
BrT1N6_2P1c	GCCAAGTAAACTGCCCAAATAC	ACCGTTAGAAACCGCAACC
BrTCTPP1b	GCCTATCATCAAACCCCATC	GAACCCCTTCTTGTCGTAAGTG
BrTCTPP2b	GAAGTTGAAGGAAAGGTCTGTTGT	GAAGGCTGCTTGATCTTCCTC
BrTFL2P1c	AAGACGGAACCACAAGTAACAAC	CGTTACTTCCTTCCCATCAGAC
BrTFL2P2a	AGAAGAAAACCCAGGTGGTGA	CTTTCCTCTTTCCCTCCTCCT
BrTSFP2c	AAAGGATCAAGCCAAATGTCC	GAGCTGTCGGAACAATACGAG
BrVEL1P1e	GCAACCACAGTATCAACCAAGA	ATCTCATCCACTCCAAGCTCTC
BrVIM3P3b	AATGCTTTGTGAAGTGGACTGA	CAGTCGTAAATGCTTCATCTGG
BrVIP4P1b	AGACGATGAGGAAGAAGACGG	ACGATGCGTAGGAGCTTTTCT
BrVRN1P1b	TCTCATGGACACCGCAC	GTTCAAAGGAAACACCCAAC
BrVRN1P3b	TCTGCATCACACAACAACCAC	CTGCATTAACAGCCCTCTCTCT
BrVRN2P4b	CTTCCTAGATGCTTGAGTTACAAT	CATAGCAAAACAAGAGAGTTGG

Appendix II: Genotype frequency of markers not confirming to 1:1 segregation

S/n	Locus	A	B	-	X2	Df	Significance	Classification
1	ATGRF5P1b_A01	46	29	5	3.85	1	**	[a:b]
2	BrSLY1P1b_A01	52	25	3	9.47	1	****	[a:b]
3	BrPIP2a_A02	24	56	0	12.8	1	******	[a:b]

S/n	Locus	A	B	-	X2	Df	Significance	Classification
4	BrTSFP2c_A02	24	56	0	12.8	1	******	[a:b]
5	CUC2P1b_A02	45	26	9	5.08	1	**	[a:b]
6	GA20OX3P1a_A02	46	30	4	3.37	1	*	[a:b]
7	JAGP1a_A02	53	21	6	13.84	1	******	[a:b]
8	JAGP1b_A02	63	9	8	40.5	1	*******	[a:b]
9	JAGP1c_A02	6	72	2	55.85	1	*******	[a:b]
10	KNAT2P1a_A02	73	5	2	59.28	1	*******	[a:b]
11	PIN1P1a_A02	31	48	1	3.66	1	*	[a:b]
12	ARF4P1b_A02	19	57	4	19	1	*******	[a:b]
13	ARLP1a_A03	28	48	4	5.26	1	**	[a:b]
14	AtBRAM_A03	56	23	1	13.78	1	******	[a:b]
15	BrFLKP2b_A03	62	15	3	28.69	1	*******	[a:b]
16	BrGAr3P1a_A03	56	24	0	12.8	1	******	[a:b]
17	BrTCTPP1b_A03	63	12	5	34.68	1	*******	[a:b]
18	Ing2P1a_A03	69	9	2	46.15	1	*******	[a:b]
19	KNAT1P1a_A03	70	7	3	51.55	1	*******	[a:b]
20	ROT4P1a_A03	62	11	7	35.63	1	*******	[a:b]
21	GA20OX3P2a_A03	59	19	2	20.51	1	*******	[a:b]
22	ATGRF2P2b_A03	30	49	1	4.57	1	**	[a:b]
23	LEPP1a_A03	56	24	0	12.8	1	******	[a:b]
24	YAB1P1a_A03	27	52	1	7.91	1	****	[a:b]
25	SEP1a_A04	57	16	7	23.03	1	*******	[a:b]
26	ARF3-A04	56	23	1	13.78	1	******	[a:b]
27	AGO1P1a_A05	19	61	0	22.05	1	*******	[a:b]
28	BrELF5P1a_A06.1	27	51	2	7.38	1	***	[a:b]
29	CNAP1a_A06	55	21	4	15.21	1	*******	[a:b]
30	WUSP1b_A06	29	48	3	4.69	1	**	[a:b]
31	AGO7P1a_A07	31	46	3	2.92	1	*	[a:b]
32	AGO7P2a_A07	27	52	1	7.91	1	****	[a:b]
33	BrFKF1P3b_A07	29	51	0	6.05	1	**	[a:b]
34	BrFVEP2a_A07	30	49	1	4.57	1	**	[a:b]
35	YAB1P3a_A07	16	64	0	28.8	1	*******	[a:b]
36	AGO1P2a_A08	48	31	1	3.66	1	*	[a:b]
37	BrESD4P2b_A08	32	47	1	2.85	1	*	[a:b]
38	cow1P3a_A08	17	57	6	21.62	1	*******	[a:b]
39	AtEXP10P1b_A09	68	11	1	41.13	1	*******	[a:b]
40	BrFPAP1c_A09	58	16	6	23.84	1	*******	[a:b]
41	BrFRL2P1b_A09	49	31	0	4.05	1	**	[a:b]
42	CycD3;3P3a_A09	62	16	2	27.13	1	*******	[a:b]
43	ERP1b_A09	46	27	7	4.95	1	**	[a:b]
44	KAN2P4a_A09	14	57	9	26.04	1	*******	[a:b]
45	KNAT6P2a_A09	64	11	5	37.45	1	*******	[a:b]
46	KRP2P2a_A09	53	26	1	9.23	1	****	[a:b]
47	BrFLDP1c_A10	66	8	6	45.46	1	*******	[a:b]

S/n	Locus	A	B	-	X2	Df	Significance	Classification
48	CUC2P2a_A10	64	15	1	30.39	1	*******	[a:b]
49	FLC5aP5a_A03	53	27	0	8.45	1	****	[a:b]
50	DDF1P2b_A09	48	28	4	5.26	1	,	[a:b]

Significances: * p ≤0.1, ** P≤0.05, *** P≤0.01, ****≤0.005, ****** P≤0.0005, *******P≤0.0001. DF (degree of freedom (1)), (a:b = CC48:P101) and X^2 = (Chi-square Test)

Appendix III: SSR markers genotyped on DH88population

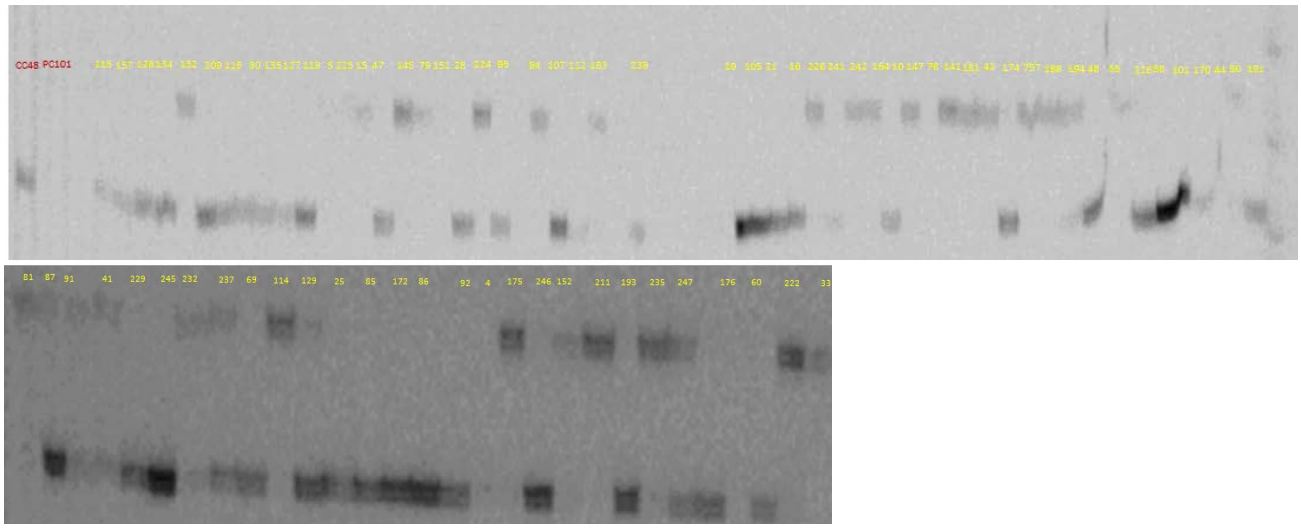


Figure 9. Simple sequence repeat (SSR) marker, BrSOC1P2b_A04, genotyping of DH88 individuals on gel electrophoresis using Li-cor machine. The parental line are on left in red indicated as CC44 (Chinese cabbage) and PC101 (Pak Choi). The genotypes identification numbers are indicated on top of the band starting with individual 115 up to 33. The individual genotypes were scored according to the parental allele, with CC48, scored as (a) and PC101 scored as (b).

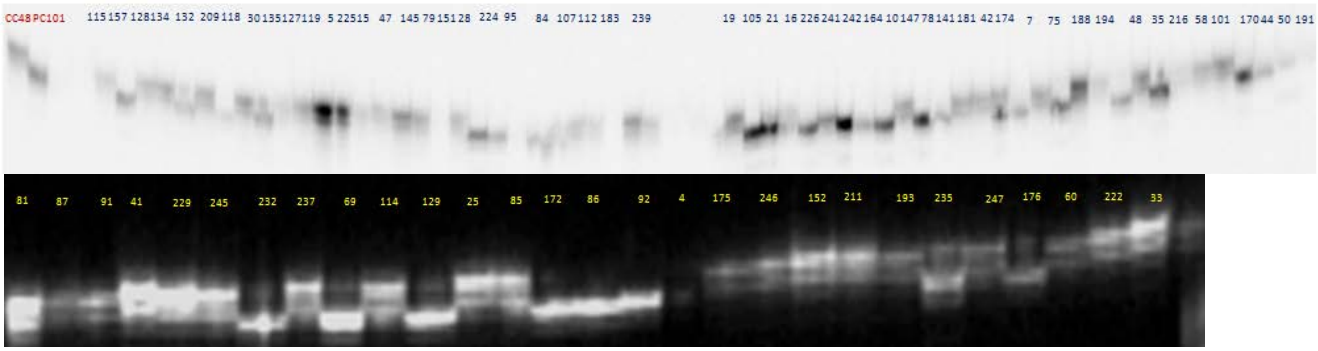


Figure 10. Simple sequence repeat (SSR) marker, BrDDF1b2_A09, genotyping of DH88 individuals on gel electrophoresis using Li-cor machine. The parental line are on left in red indicated as CC44 (Chinese cabbage) and PC101 (Pak Choi). The genotypes identification numbers are indicated on top of the band starting with individual 115 up to 33. The individual genotypes were scored according to the parental allele, with CC48, scored as (a) and PC101 scored as (b).

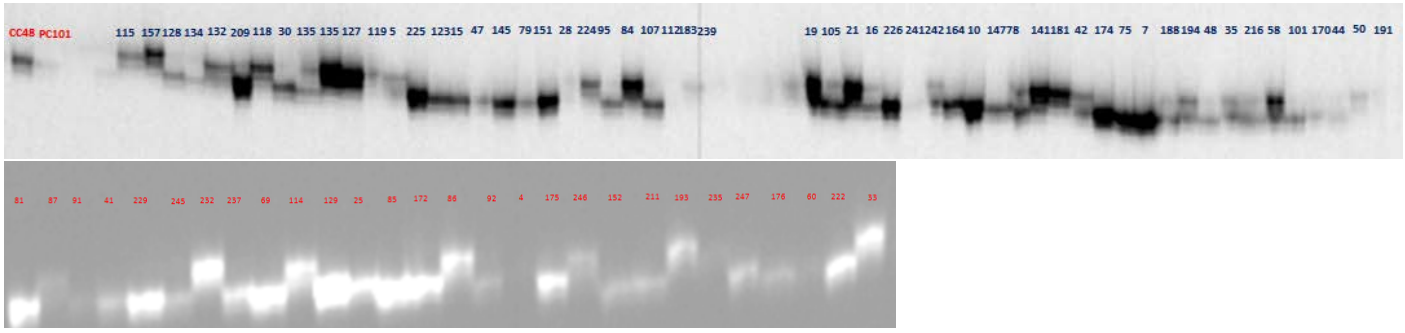


Figure 11. Simple sequence repeat (SSR) marker, BrCCA1P1e_A05, genotyping of DH88 individuals on gel electrophoresis using Li-cor machine. The parental line are on left in red indicated as CC44 (Chinese cabbage) and PC101 (Pak Choi). The genotypes identification numbers are indicated on top of the band starting with individual 115 up to 191. The individual genotypes were scored according to the parental allele, with CC48, scored as (a) and PC101 scored as (b).

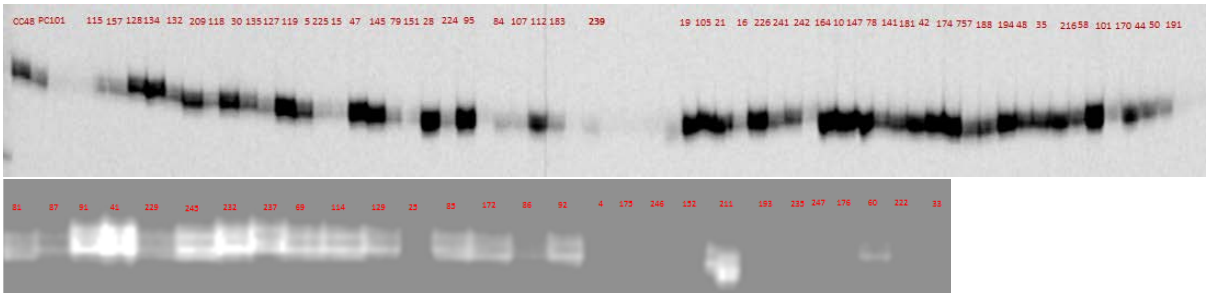


Figure 12. Simple sequence repeat (SSR) marker, FLC3P3b_A03, genotyping of DH88 individuals on gel electrophoresis using Li-cor machine. The parental line are on left in red indicated as CC44 (Chinese cabbage) and PC101 (Pak Choi). The genotypes identification numbers are indicated on top of the band starting with individual 115 up to 191. The individual genotypes were scored according to the parental allele, with CC48, scored as (a) and PC101 scored as (b).

Appendices IV: SNP markers from light scanner analysis showing polymorphism on DH88

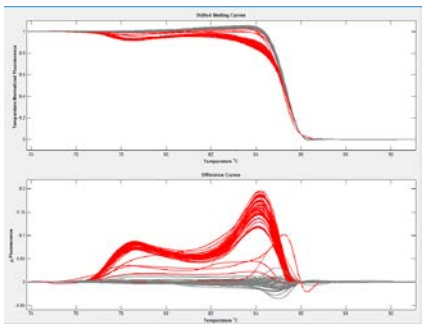


Figure 13. Melting curves profile of BrATGT_A03 marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

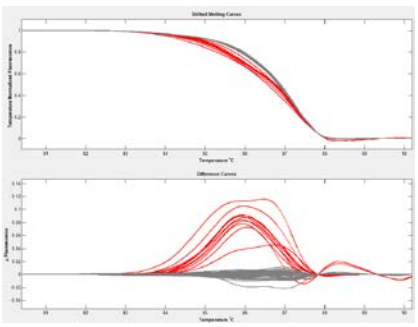


Figure 14. Melting curves profile of BrAtBRAM_A03 marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. In the Y-axis fluorescence and temperature in degree ºC

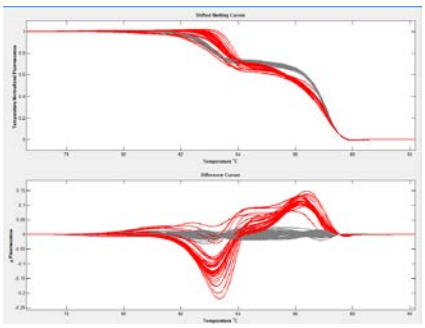


Figure 15. Melting curves profile of BrESD4P2b_A08 marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the common allele "a", for PC101 and grey for alleles of variant "b", forCC48. In the Y-axis fluorescence and temperature in degree ºC

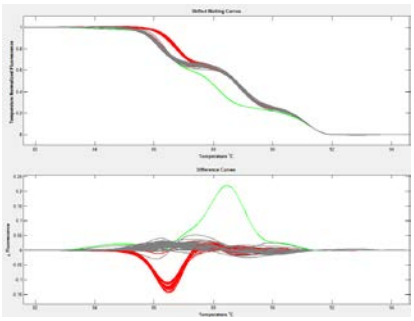


Figure 16. Melting curves profile of BrELF5P1a_A06marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

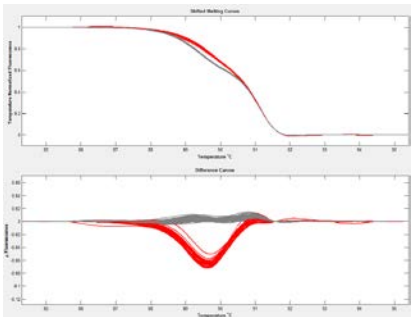


Figure 17. Melting curves profile of BrFKF1P3b marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

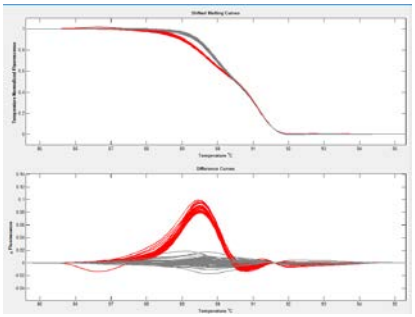


Figure 18. Melting curves profile of BrFKF1P3c_A07 marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

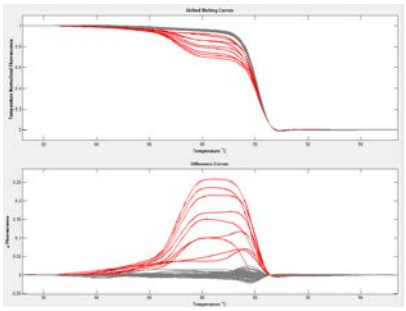


Figure 19. Melting curves profile of BrFLDP1c_A10marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

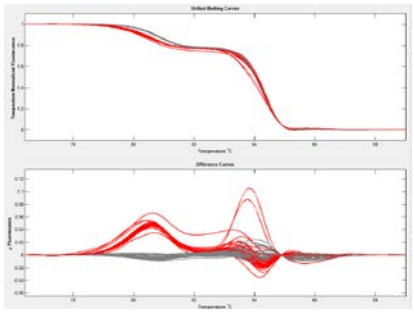


Figure 20. Melting curves profile of BrFLMP1b_A02 marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

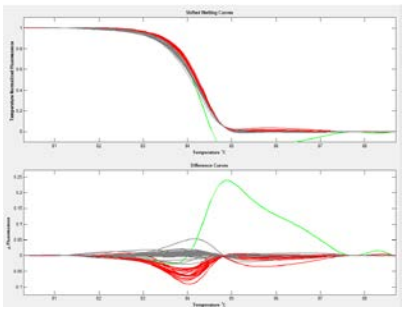


Figure 21. Melting curves profile of BrFPAP1c_A09marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "b", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

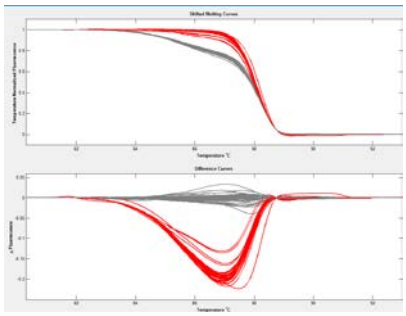


Figure 22. Melting curves profile of BrFRL1P2a_A10marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

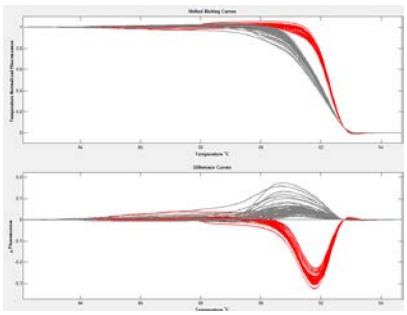


Figure 23. Melting curves profile of BrFRL2P1b_A09marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "b", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

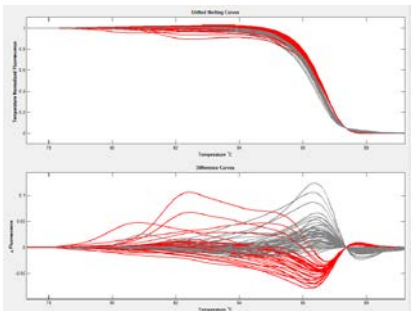


Figure 24. Melting curves profile of BrFVEP2a_A07 marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

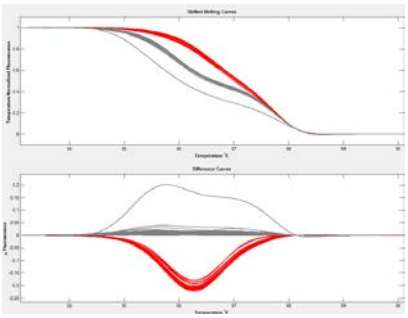


Figure 25. Melting curves profile BrGAr3P1a_A03marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", forCC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

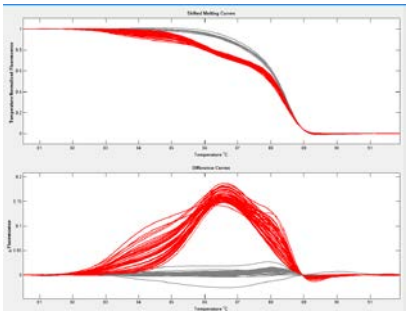


Figure 26. Melting curves profile BrGAr1P1d marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

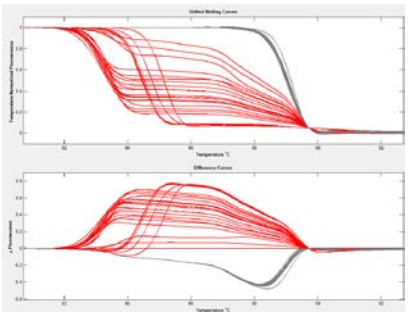


Figure 27. Melting curves profile BrGAr1P1d marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

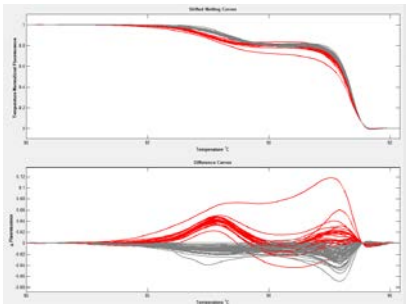


Figure 28. Melting curves profile BrSLY1P1b_A01marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

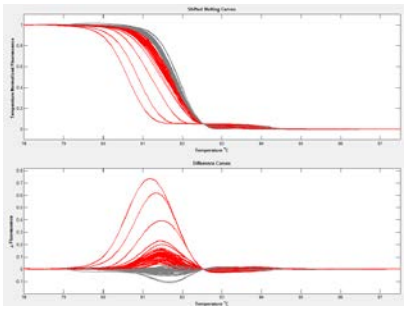


Figure 29. Melting curves profile BrSNZP2b_A09marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

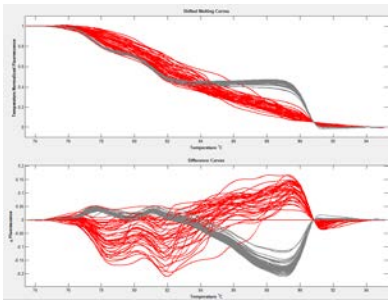


Figure 30. Melting curves profile BrT1N6_2P1cmarker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101 The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

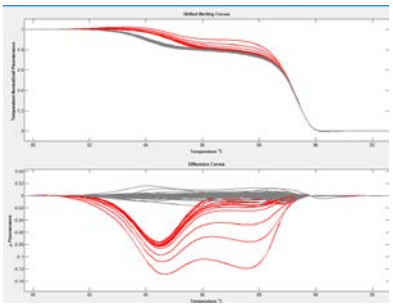


Figure 31. Melting curves profile BrTCTPP1b_A03marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

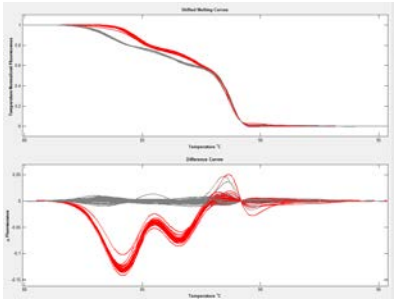


Figure 32. Melting curves profile BrTCTPP3b marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

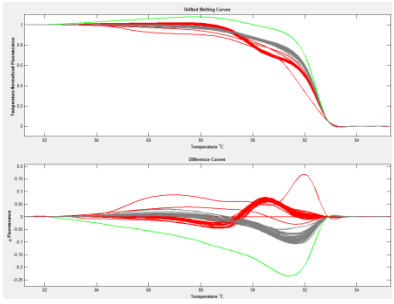


Figure 33. Melting curves profile BrTFL2P2a_A10marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

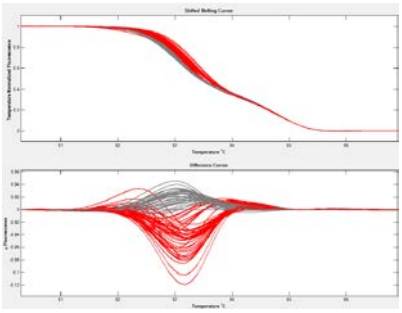


Figure 34. Melting curves profile BrTFLP2c marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

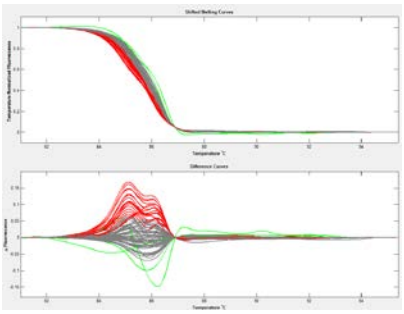


Figure 35 Melting curves profile BrARF4P1b_A02marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "a", for CC48 and grey for common allele "b", for PC101. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

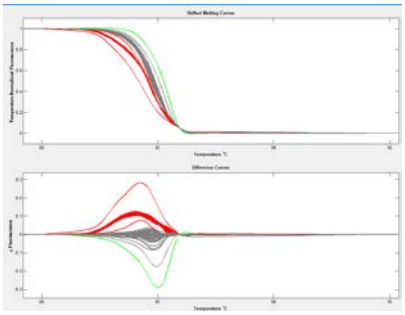


Figure 36. Melting curves profile BrARF3_A04marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

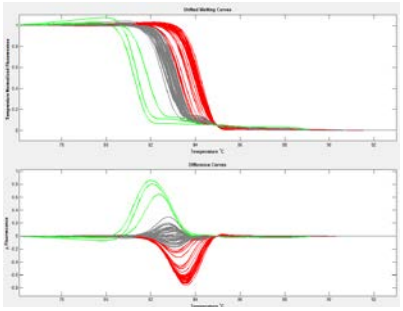


Figure 37. Melting curves profile BrARF4Indel_A10marker on DH88 individuals separating them in distinct groups after computing the difference below. Red colour representing the alleles of variant "b", for PC101 and grey for common allele "a", for CC48. The Y-axis representing the melting curves of nucleotides as fluorescence and the X-axis as melting temperature in degrees ºC.

Appendix V: fQTL pValues for analysis of significance

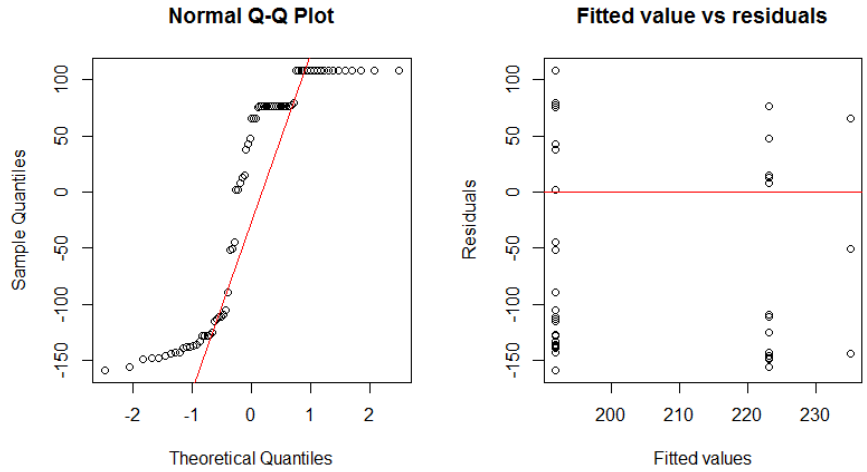
Marker	Pvalue	Pvalue adjust
ATGRF5P1b_A01	0.221849	0.575275
BrSLY1P1b_A01	0.203339	0.575275
CycD3_1P1a_A01	0.002166	0.03682*
HST1P1b_A01	0.401627	0.642355
KAN2P1b_A01	0.663542	0.801245
KAN3P1b_A01	0.188073	0.575275
BrFLMP1b_A02.2	0.670384	0.801245
BrPIP2a_A02	0.990368	0.990368
BrTSFP2c_A02	0.629191	0.798228
CUC2P1b_A02	0.025743	0.177927
CYCD1_1P1a_A02	0.215661	0.575275
GA20OX3P1a_A02	0.886666	0.92224
JAGP1a_A02	0.607579	0.795234
JAGP1b_A02	0.661065	0.801245
JAGP1c_A02	0.60812	0.795234

KNAT2P1a_A02	0.583761	0.795234
PIN1P1a_A02	0.228314	0.575275
ARF4P1b_A02	0.421743	0.642355
ARLP1a_A03	0.000197	0.00559**
AtBRAM_A03	0.009416	0.10004
BrFLKP2b_A03	0.368521	0.642355
BrGAr3P1a_A03	0.605492	0.795234
BrTCTPP1b_A03	0.02811	0.177927
Ing2P1a_A03	0.112724	0.479076
KNAT1P1a_A03	0.597142	0.795234
ROT4P1a_A03	0.364653	0.642355
GA20OX3P2a_A03	0.358883	0.642355
cow1P2a_A03	0.013142	0.11263
ATGRF2P2b_A03	0.486029	0.712284
ATVGT_A03	0.081922	0.386853
LEPP1a_A03	0.944478	0.955722
YAB1P1a_A03	0.423199	0.642355
CLFP1b_A04	0.166473	0.575275
SEP1a_A04	0.118664	0.480306
ARF3_A04	0.495807	0.714298
AGO1P1a_A05	0.369257	0.642355
ROT4P2a_A05	0.105636	0.472583
BrTCTPP2b_A05	0.272613	0.642355
BrELF5P1a_A06.1	0.39445	0.642355
CNAP1a_A06	0.23011	0.575275
CycB2_3P1a_A06	0.377911	0.642355
PHVP1A_A06	0.776756	0.852262
REVP2b_A06	0.223127	0.575275
SEP2a_A06	0.782076	0.852262
WUSP1b_A06	0.043759	0.218793
BrGAr1P1d_A06	0.766749	0.852262
AGO7P1a_A07	0.390267	0.642355
AGO7P2a_A07	0.339986	0.642355
BrFKF1P3b_A07	0.029306	0.177927
BrFVEP2a_A07	0.017304	0.133713
KAN1P2a_A07	0.663542	0.801245
TCP15P2a_A07	0.879681	0.92224
YAB1P3a_A07	0.225338	0.575275
AGO1P2a_A08	0.830387	0.893455
AGO1P2b_A08	0.18595	0.575275
BrESD4P2b_A08	0.31426	0.642355
cow1P3a_A08	0.013251	0.11263
CycB2_3P2a_A08	0.620567	0.798228
CycB3_1P2a_A08	0.404157	0.642355
PPD2P2a_A08	0.227827	0.575275
YAB2P3a_A08	0.265145	0.642355

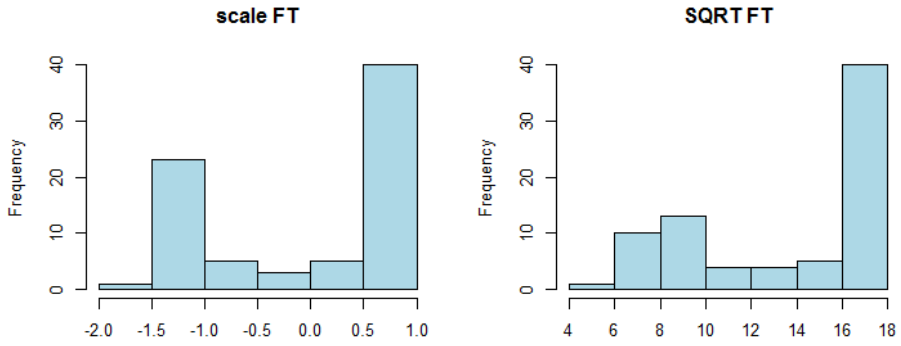
AtEXP10P1b_A09	0.891225	0.92224
BrFPAP1c_A09	0.324319	0.642355
BrFRL2P1b_A09	0.450141	0.671263
BrSNZP2b_A09	0.422117	0.642355
CycD3_3P3a_A09	4.78E-05	0.004064**
ERP1a_A09	0.040514	0.215233
ERP1b_A09	0.363964	0.642355
FNR1P2b_A09	0.034587	0.195992
KAN2P3a_A09	0.333604	0.642355
KAN2P4a_A09	0.678702	0.801245
KNAT6P2a_A09	0.698335	0.807002
KRP2P2a_A09	0.324537	0.642355
BrT1N6_2P1c_A09	0.213921	0.575275
ANP1b_A10	0.005634	0.068417
BrFLDP1c_A10	0.90054	0.92224
BrTFL2P2a_A10	0.000144	0.00559**
CUC2P2a_A10	0.000339	0.007208**
GA20OX3P4b_A10	0.003023	0.04283*
BrARF4.Indel_A10	0.212601	0.575275
SOC1P2b.SSR_A04	0.71206	0.807002
FLC3P3b.SSR_A03	0.50439	0.714552
FLC5aP5a.SSR_A03	0.306236	0.642355
DDF1P2b.SSR_A09	0.705177	0.807002
CCA1P1e.SSR_A05	0.362438	0.642355

*Significant *p < 0.05, ** p < 0.01, and ***p < 0.001

Appendix VI: Normality check of sample



Appendix VII: Flowering time phenotypic expression of DH88



Flowering time data

Genotype	Flowering Time	D.a.T - Days after Transplanting			
Pedigree	**D.a.T**	**Pedigree**	**D.a.T**	**Pedigree**	**D.a.T**
DH88 241	33	DH88 241	33	DH88 241	33
DH88 134	49	DH88 134	49	DH88 134	49
DH88 242	50	DH88 242	62	DH88 242	56
DH88 041	54	DH88 041	50	DH88 041	58
DH88 025	52	DH88 025	52	DH88 025	54
DH88 157	56	DH88 157	52	DH88 157	58
DH88 079	54	DH88 079	54	DH88 079	54
DH88 245	54	DH88 245	54	DH88 245	54
DH88 129	58	DH88 129	62	DH88 129	58
DH88 176	69	DH88 176	58	DH88 176	69
DH88 118	66	DH88 118	66	DH88 118	60
DH88 232	66	DH88 232	73	DH88 232	61
DH88 237	69	DH88 237	61	DH88 237	62
DH88 019	62	DH88 019	61	DH88 019	69
DH88 164	66	DH88 164	66	DH88 164	90
DH88 183	66	DH88 183	86	DH88 183	88
DH88 170	66	DH88 170	88	DH88 170	76
DH88 209	76	DH88 209	87	DH88 209	69
PC101	123	PC 101	69	PC101	300
DH88 128	78	DH88 128	72	DH88 128	108
DH88 114	72	DH88 114	89	DH88 114	76
DH88 095	89	DH88 095	87	DH88 095	72
DH88 132	76	DH88 132	75	DH88 132	75
DH88 224	75	DH88 224	88	DH88 224	76
DH88 119	185	DH88 119	76	DH88 119	76
DH88 147	76	DH88 147	121	DH88 147	102
DH88 135	175	DH88 135	88	DH88 135	79
DH88 188	87	DH88 188	103	DH88 188	82
DH88 021	108	DH88 021	82	DH88 021	104

Genotype	Flowering Time	D.a.T - Days after Transplanting			
Pedigree	**D.a.T**	**Pedigree**	**D.a.T**	**Pedigree**	**D.a.T**
DH88 044	92	DH88 044	202	DH88 044	300
DH88 229	121	DH88 229	92	DH88 229	92
DH88 160	99	DH88 160	300	DH88 160	111
DH88 115	103	DH88 115	300	DH88 115	300
DH88 216	108	DH88 216	300	DH88 216	300
DH88 021	108	DH88 021	82	DH88 021	104
DH88 048	111	DH88 048	300	DH88 048	300
CC048	111	CC048	188	CC48	300
DH88 086	300	DH88 086	121	DH88 086	300
DH88105	175	DH88105	133	DH88 105	111
DH88-84	213	DH88-84	300	DH88-84	300
DH88 145	178	DH88 145	199	DH88 145	175
DH88 087	182	DH88 087	213	DH88 087	188
DH88 235	213	DH88 235	300	DH88 235	300
DH88 84	213	DH88 84	300	DH88 84	300
DH88 112	207	DH88 112	300	DH88 112	300
DH88 247	207	DH88 247	207	DH88 247	300
DH88 175	207	DH88 175	300	DH88 175	185
DH88 194	202	DH88 194	300	DH88 194	300
DH88 193	199	DH88 193	199	DH88 193	185
DH88 246	189	DH88 246	300	DH88 246	199
DH88 105	175	DH88 105	133	DH88 105	111
DH88 135	175	DH88 135	88	DH88 135	79
DH88 069	300	DH88 069	175	DH88 069	300
DH88 226	300	DH88 226	300	DH88 226	300
DH88 030	300	DH88 030	300	DH88 030	300
DH88 141	300	DH88 141	300	DH88 141	300
DH88 078	300	DH88 078	300	DH88 078	300
DH88 010	300	DH88 010	300	DH88 010	300
DH88 005	300	DH88 005	300	DH88 005	300
DH88 127	300	DH88 127	300	DH88 127	300
DH88 092	300	DH88 092	300	DH88 092	300
DH88 086	300	DH88 086	300	DH88 086	300
DH88 172	300	DH88 172	300	DH88 172	300
DH88 085	300	DH88 085	300	DH88 085	300
DH88 075	300	DH88 075	300	DH88 075	300
DH88 174	300	DH88 174	300	DH88 174	300
DH88 042	300	DH88 042	300	DH88 042	300
DH88 181	300	DH88 181	300	DH88 181	300
DH88 47	300	DH88 47	300	DH88 47	300
DH88 015	300	DH88 015	300	DH88 015	300
DH88 123	300	DH88 123	300	DH88 123	300
DH88 225	300	DH88 225	300	DH88 225	300

Genotype	Flowering Time	D.a.T - Days after Transplanting			
Pedigree	**D.a.T**	**Pedigree**	**D.a.T**	**Pedigree**	**D.a.T**
DH88 152	300	DH88 152	300	DH88 152	300
DH88 007	300	DH88 007	300	DH88 007	300
DH88 028	300	DH88 028	300	DH88 028	300
DH88 151	300	DH88 151	300	DH88 151	300
DH88 211	300	DH88 211	300	DH88 211	300
DH88 101	300	DH88 101	300	DH88 101	300
DH88 058	300	DH88 058	300	DH88 058	300
DH88 035	300	DH88 035	300	DH88 035	300
DH88 107	300	DH88 107	300	DH88 107	300
DH88 033	300	DH88 033	300	DH88 033	300
DH88 222	300	DH88 222	300	DH88 222	300
DH88 060	300	DH88 060	300	DH88 060	300
DH88 161	300	DH88 161	300	DH88 161	300
DH88 050	300	DH88 050	300	DH88 050	300
DH88 239	300	DH88 239	300	DH88 239	300

Appendix VIII: Annotation of leaf trait markers as candidate genes in *B. rapa* based on *A. thaliana* characterization of the homologous genes

Homologous gene in *A. t*	Gene family	*A.t* Chr.	*B. Rapa* gene	*B. rapa* Chr.	Marker name	Putative gene function
At5g60450	ARF4	5	Bra020243	10	BrARF4-Indel	Response to hormone stimulus, regulation of transcription and DNA dependent
At1g12610	DDF1	1	Bra026963	9	BrDDF1P2b	Regulation of transcription and DNA-dependent
At5g02840	CCA1	5	Bra019545	5	BrCCA1P1e	Regulating circadian rhythms
At2g24765	ARF3	2	Bra032062	4	BrARF3	Intracellular protein transport and GTPase mediated signal transduction
At2g45660	SOC1	2	Bra039324	4	BrSOC1P2b	Flower development, maintenance of inflorescence, regulation of transcription and response to gibberellin stimulus
At5g10140	FLC	5	Bra028599	3	BrFLC3P3b	Carpel development, negative regulation of flowering development, regulation response to temperature stimulus and vernalization response
At5g10140	FLC	5	Bra028599	3	BrFLC5aP5a	Carpel development, negative regulation of flowering development, regulation response to temperature stimulus and vernalization response
At5g60450	ARF4	5	Bra002479	2	ARF4P1b	Abaxial cell fate specification, regulation of transcription, DNA dependent, response to hormone stimulus and vegetative phase change
At3g05040	HST1	3	Bra040117	1	BrHST1P1b	Leaf morphogenesis, negative regulation of flower development, polarity specification of adaxial/ abaxial axis, pre-mRNA export from the nucleus and shoot development
At1g32240	KAN2	1	Bra023254	1	BrKAN2P1b	Carpel development polarity specification of adaxial/ abaxial axis, regulation of transcription and DNA-dependent
At4g34160	CYCD3;1	4	Bra011501	1	BraCycD3;1P1a	Regulation of cell cycle, regulation of cell proliferation, regulation of meristem growth, response to brassinosteroid stimulus, response to cyclopentone, response to cytokinin stimulus, response to sucrose stimulus, seed development, spindle

Homologous gene in *A. t*	Gene family	*A.t* Chr.	*B. Rapa* gene	*B. rapa* Chr.	Marker name	Putative gene function
						assembly, stomatal and complex morphogenesis
At4g17695	KAN3	4	Bra040176	1	BrKAN3P1b	Regulation of transcription and DNA-dependent
At3G13960	ATGRF5	3	Bra021521	1	BrGRF5P1b	Floral organ formation, histone modification, leaf development, leaf morphogenesis, regulation of floral development and regulation of transcription
At5g53950	CUC2	5	Bra003023	2	BrCUC2P1b	Formation of organ boundary, leaf development, leaf morphogenesis, primary shoot apical meristem specification, regulation of timing organ formation, regulation of transcription, DNA-dependent and secondary shoot formation
At1g68480	JAG	1	Bra004312	2	BrJAGP1a	Abaxial cell fate specification, anther development, carpel development, flower development, leaf morphogenesis, meristem development, petal formation, sepal formation, specification of floral organ identity and stamen development
At1g70510	KNAT2	1	Bra007920	2	BrKNAT2P1a	Chromosome segregation, cytokinin signalling pathway, determination of bilateral symmetry, floral whorl development and flower morphogenesis
At5g07200	GA20OX3	5	Bra009285	2	BrGA20OX3P1a	Gibberellin biosynthesis process, oxidation-reduction process and terpenoid biosynthetic process
At4g00220	NUB	4	Bra037322	2	BrJAGP1c	Organ boundary specification between lateral organs and meristem, and xylem development
At1g68480	JAG	1	Bra004312	2	BrJAGP1b	Abaxial cell fate specification, anther development, carpel development, flower development, leaf morphogenesis, meristem development, petal formation, sepal formation, specification of floral organ identity and stamen development
At1g70210	CYCD1;1	1	Bra007901	2	BrCYCD1;P1a	DNA endoduplication, G1 phase of mitotic cell cycle, stomatal lineage progression

Homologous gene in *A. t*	**Gene family**	***A.t* Chr.**	***B. Rapa* gene**	***B. rapa* Chr.**	**Marker name**	**Putative gene function**
At1g73590	PIN1	1	Bra015983	2	BrPIN1P1a	Anthocyanin accumulation in tissue in response to UV light, auxin polar transport, cell wall macromolecule metabolism, cotyledon development, leaf formation, leaf morphogenesis, leaf shaping and nucleotide biosynthesis
At4g37740	GRF2	4	Bra011781	3	BrCRF2P2b	Embryo development ending in seed dormancy, leaf development, leaf morphogenesis, regulation of transcription and DNA-dependent
At4g08150	KNAT1	4	Bra00638	3	BrKNAT1P1a	Cell fate commitment, cell fate specification, determination of bilateral symmetry, floral whorl development, flower morphogenesis, meristem Initiation, organ development, organ morphogenesis, pattern specification process, polarity specification adaxial/abaxial axis and regulation of transcription
At2g36985	ROT4	2	Bra005225	3	BrROT4P1a	Regulation of cell proliferation and shoot development
At5g07200	GA20OX3	5	Bra009285	3	BrGA20OX3P2a	Gibberellin biosynthetic process, oxidation-reduction process and terpenoid biosynthesis process
At2g44080	ARL	2	Bra004803	3	BrARLP1a	Brassinosteroid mediated signalling pathway, cell growth and response to brassinosteroid stimulus
At1G54390	ING2	1	Bra040776	3	BrIng2P1a	Methylated histone residue binding
At4g34580	COW1	4	Bra011544	3	BrCOW1P2a	Developmental growth, transport and trichoblast differentiation
At5g13910	LEP	5	Bra008793	3	BrLEPP1a	Regulation of transcription, DNA-dependent and response to gibberellin stimulus
At2g45190	YAB1	2	Bra040322	3	BrYAB1P1a	Abaxial cell fate specification, cell fate commitment, chromatin assembly or disassembly, development ending in seed dormancy, inflorescences meristem growth, meristem structure organization, polarity specification of adaxal/abaxial axis, regulation of flower development, specification of organ position, stomatal complex

Homologous gene in *A. t*	Gene family	*A.t* Chr.	*B. Rapa* gene	*B. rapa* Chr.	Marker name	Putative gene function
						position and vegetative to reproductive phase transition of meristem
At2g27100	SE	2	Bra012034	4	BrSEP1a	RNA splicing, chromatin modification, mRNA spicing via spliceosome, positive regulation of transcription, DNA-dependent meristem development and regulation of mitotic cycle
At2g23380	CLF	2	Bra032169	4	BrCLFP1b	DNA mediated transformation, DNA methylation, gene silencing by RNA, histone methylation, leaf morphogenesis, epigenetic regulation of gene expression y genetic imprinting, vegetative to reproductive phase transition of meristem and vernalization
At1g48410	AGO1	1	Bra032254	5	BraGO1P1a	RNA interference, adaxial/abaxial pattern specification, auxin metabolic process, auxin polar transport, cell adhesion, cell division, cell wall organization, cell-cell signalling, chromatin silencing and leaf vascular tissue pattern formation
At2g36985	ROT4	2	Bra005225	5	BrROT4P2a	Regulation of cell proliferation and shoot development
At1g30490	PHV	1	Bra032394	6	BrPHVP1a	Adaxial/axial specification, cell adhesion, cell wall organisation, determination of bilateral symmetry, embryonic pattern specification, meristem initiation and meristem maintenance
At5g60690	REV	5	Bra002458	6	BrREVP2b	Cell differentiation, cell growth, cell wall organization, leaf morphogenesis, meristem initiation, polarity specification of adaxial/abaxial axis and radial pattern formation
At2g27100	SE	2	Bra012034	6	BrSEP2a	Regulation of mitotic cycle, RNA splicing, chromatin modification, mRNA splicing via spliceosome, positive regulation of transcription and regulation of adaxial/abaxial pattern formation
At1g52150	CNA	1	Bra030088	6	BrCNAP1a	Determination of bilateral symmetry, determination of dorsal identity, leaf morphogenesis, meristem initiation and polarity specification of adaxial/abaxial axis
At2g17950	WUS	2	Bra039894	6	BrWUSP1b	Carpel development, plant type cell wall modification and regulation of transcription
At1g20610	CycB;3	1	Bra025831	6	BrCyCB2;3P1a	Regulation of cell cycle and regulation of cyclin dependent protein kinase activity

Homologous gene in *A. t*	**Gene family**	***A.t* Chr.**	***B. Rapa* gene**	***B. rapa* Chr.**	**Marker name**	**Putative gene function**
At1g69440	AGO7	1	Bra003999	7	BraGO7P1a	Gene silencing by mRNA, production of siRNA involved in RNA interference and vegetative phase change
At5g16560	KAN1	5	Bra008613	7	BrKAN1P2a	Abaxial cell fate specification, adaxial\abaxial axis specification, carpel development, organ morphogenesis, polarity specification of adaxial\abaxial axis and radial pattern formation
At2g45190	YAB1	2	Bra040322	7	BrYAB1P3a	Abaxial cell fate determination, cell fate commitment, stomatal complex morphogenesis and polarity assembly
At1g69690	TCP15	1	Bra004407	7	BrTCP15P2a	Cell proliferation, circadian rhythm and leaf morphogenesis
At1g48410	AGO1	1	Bra032254	8	BraAGO1P2a	Leaf proximal/distal pattern formation, vegetative phase change and RNA interference
At1g09700	HYL1	1	Bra019999	8	BraHYL1P2a	Leaf vascular tissue pattern formation, response to cytokinin stimulus and response to auxin stimulus
At1g08465	YAB2	1	Bra018624	8	BraYAB2P3a	Abaxial cell fate specification
At4g34580	COW1	4	Bra011544	8	BrCow1P3a	Cell tip growth, development and cell differentiation
At1g16330	CYCB3;1	1	Bra026065	8	BrCycB3;1P2a	Regulation of cell cycle and stomatal lineage progression
At4g14720	PPD2	4	Bra036885	8	BrPPD2P2a	Leaf development and oxidation reduction process
At1g20610	CycB;3	1	Bra025831	8	BrCycB2;3P2a	Regulation of cell cycle and regulation of cyclin-dependent protein kinase activity
At1g26770	AtEXP10	1	Bra024686	9	BrEXP10P1b	Plant type cell wall loosening, plant type cell wall modification involved in multidimensional cell growth and cell organization
At2g26330	ER	2	Bra007759	9	BrERP1a	Asymmetrical cell division, leaf morphogenesis, plant type cell wall organization,

Homologous gene in *A. t*	Gene family	*A.t* Chr.	*B. Rapa* gene	*B. rapa* Chr.	Marker name	Putative gene function
						polarity specification of adaxial/abaxial axis, regulation of anthocyanin biosynthesis process, regulation of organ morphogenesis and stomatal complex morphogenesis
At1g32240	KAN2	1	Bra023254	9	BrKAN2P3a	Carpel development, specification of adaxial/abaxial axis, regulation of transcription and DNA-dependent
At1g23380	KNAT6	1	Bra024593	9	BrKNAT6P2a	Determination of bilateral symmetry, regulation of transcription and DNA-dependent
At5g66190	FNR1	5	Bra012057	9	BrFNR1P2b	Glucosinolate biosynthesis, oxidation reduction process, pentose-phosphate shunt and photosynthesis
At3g50070	CYCD3;3	3	Bra036051	9	BrCycD3;3P3a	DNA endoreduplication, mitotic cell cycle, negative regulation of cyclin-dependent protein kinase activity, regulation of cell cycle, regulation of cell proliferation and stomatal lineage progression
At3g50630	KRP2	3	Bra036095	9	BrKRP2P2a	Cell cycle arrest, mitotic cell cycle, negative regulation of cyclin-dependent protein kinase activity and stomatal lineage progression
At1g01510	AN	1	Bra033258	10	BrANP1b	Leaf morphogenesis, monopolar cell growth, oxidation-reduction process, regulation of cell shape, regulation of epidermal cell differentiation and regulation of epidermal cell division.
At5g53950	CUC2	5	Bra003023	10	BrCUC2P2a	Formation of organ boundary, leaf development and leaf morphogenesis
At5g07200	GA20OX3	5	Bra009285	10	BrGA20OX3P4b	Gibberellin biosynthesis process, oxidation-reduction process and terpenoid biosynthesis process

Appendix IX: Annotation of flowering markers as candidate genes in *B. rapa* based on *A. thaliana* characterization of the syntenic genes

Homologous gene in *A. t*	Gene family	*A.t* Chr.	*B. Rapa* gene	*B. rapa* Chr.	Marker name	Putative gene function
At4g24210	SLY1	4	1	Bra013776	BrSLY1P1b	Floral organ morphogenesis and gibberellic acid mediated signalling pathway
At1g77080	FLM	1	2	Bra031888	BrFLMP1b	Anthocyanin accumulation in tissues in response to ultra violet (UV) light, carpel development, negative regulation of flowering development and vernalization response
At5g20240	PI	5	2	Bra002285	BrPIP2a	Ovule development, regulation of transcription and DNA-dependent
At4g20370	TSF	4	2	Bra015710	BrTSFP2c	Photoperiodism, and positive regulation of flowering development
At2g46020	BRM	2	3	Bra039296	BrAtBRMP1d	Adenosine triphosphate (ATP)-dependent chromatin remodelling, organ boundary specification between lateral organs and the meristem, regulation of gene expression, epigenetic, regulation of transcription and DNA-dependent
At3g03090	VGT1	3	3	Bra032018	BrATVGT1P1a	Fructose transport, glucose transport, positive regulation of flower development, and transmembrane transport
At3g04610	FLK	3	3	Bra001111	BrFLKP2b	Positive regulation of flowering development
At3g63010	GAr3	3	3	Bra007722	BrGAr3P1a	Floral organ morphogenesis, gibberellin acid mediated pathway, response to external stimulus and response to gibberellin acid
At3g16640	TCTP	3	3	Bra021187	BrTCTPP1b	Regulation of mitotic cycle, auxin homeostasis, auxin mediated signalling pathway, cell proliferation, mitosis, pollen tube growth, positive regulation of microtubule depolymerization, post-embryonic development, regulation of cell growth and regulation of stomatal closure
At5g62640	ELF5	5	6	Bra010108	BrELF5P1a	RNA processing, photoperiodism and flowering

Homologous gene in *A. t*	**Gene family**	***A.t* Chr.**	***B. Rapa* gene**	***B. rapa* Chr.**	**Marker name**	**Putative gene function**
At3g16640	TCTP	3	5	Bra022172	BrTCTPP2b	Involved in auxin homeostasis, auxin mediated signalling pathway, cell proliferation, pollen tube growth, positive regulation of microtubule depolymerization, post-embryonic development, regulation of cell growth and regulation of stomatal growth
At5g27320	GAr1	5	6	Bra009970	BrGAr1P1d	Floral organ morphogenesis, gibberellic acid mediated signalling pathway, positive regulation of gibberrelic acid signalling pathway, and raffinose family oligosaccharide biosynthesis process
At1g68050	FKF1	1	7	Bra004278	BrFKF1P3c	Circadian rhythm, positive regulation of flower development, regulation of gene expression, regulation of transcription, DNA dependent, response to blue light and ubiquitin-dependent protein catabolic process.
At2g19520	FVE	2	7	Bra040681	BrFVEP2a	DNA repair, flower development, leaf morphogenesis, response to UV-B, trichome morphogenesis and unidimensional cell growth
At4g15880	ESD4	4	8	Bra012758	BrESD4P2b	Glucuroxylan metabolic process, mRNA export from nucleus, maintenance of meristem identity, positive regulation of flower development, proteolysis, regulation of flower development and xylan biosynthesis process
At2g43410	FPA	2	9	Bra007458	BrFPAP1c	Chromatin silencing of small RNA, embryo development, embryo sac development, mRNA polyadenylation, positive regulation of flower development and vegetative to reproductive phase transition
At1g31814	FRL2	1	9	Bra023207	BrFRL2P1b	Vegetative to reproductive phase transition of meristem
At2g39250	SNZ	2	9	-	BrSNZP2b	Regulation of transcription, DNA-dependent
At1g01640	T1N6_2	1	9	Bra033246	BrT1N6_2P1c	Floral organ abscission, flower development, flower morphogenesis and proximal/distal pattern formation

Homologous gene in *A. t*	**Gene family**	***A.t* Chr.**	***B. Rapa* gene**	***B. rapa* Chr.**	**Marker name**	**Putative gene function**
At3g10390	FLD		10	Bra001357	BrFLDP1c	Oxidation reduction process
At5g17690	TFL2	5	10	Bra013958	BrTFL2P2a	Vernalization response, multidimensional cell growth, negative regulation of flower development, photoperiodism, flowering, regulation of mitotic cell cycle, shoot morphogenesis and vernalization response

Appendix X: Li-cor protocol

Before starting: clean the gel making items with water, then with ethanol and dry them with a new paper. Check glass plates carefully. Check if the 94 ◦C heater near the window is switched on. Put 20 µl volumes of formamide loading buffer to 0.6 µl of the PCR product, close the seal again and mix on centrifuge at 1000g maximum speed.

Place the two spacers on the large glass plate (black-plate) and leave some space on the bottom of the plate (this is easier for the cleaning of the glass plates at the end) lace the other smaller plate (front-plate) on top of the back plate (which have the spacers and do not move the spacers, the writing on the front plate is now readable).

Place the two black holders at each side and make sure that they make contact with the glass plate on the bottom and the top.

First screw the middle nut and then the bottom nut (no to tight).

Make the gel solution: 20 ml Long Ranger solution

15 µl TEMED (add in the small fume hood)

150 µl APS (add in the small fume hood, smell like rotten eggs)

Pour the gel and let the capillary do the work (maybe gently tap on the glass with your finger)

Place the top spacer in the gel and lock it with the black/ transparent holder (screw the top nuts not to tight).

Let the gel polymerize for $1^1/_2$ hour

Place the gel assembly support in the transparent holder on the lab table, so that the gel is standing vertical and put a paper comb 1 mm in the gel (from now on be very careful not to move the comb).

Screw the top compartment at the top of the back plate.

Place a silver paper at the back of the plate.

Place the gel inside the LICOR machine

Make a new buffer: 100 ml 10X buffer (TBE)

900 ml distilled water (green tap)

First fill the top compartment to the maximum that is indicated with the buffer.

The fill the bottom compartment with the remaining buffer

Run the calibration program, don't change the parameters only fill in the name of the folder and the file name.

When running the program make sure that the indicated values left and right are having not more than 50 units of difference. Let the gel heat up to 45 ◦C.

Heat the samples in the 94 ◦C heater for 1 minute and directly place in the cooling block from the fridge.

Open the door of the LICOR machine.

Clean the spacer between the comb by spraying with a syringe (this is now filled with ureum which will interfere with the loding of the gel).

Load 0.5 µl of IRD 800 product.

Load 0.5 µl of the IRD 800 size marker (at least 3 size markers should be included one at each side and one in the middle).

Close the top compartment and the bottom compartment and then close the door of the LICOR machine and press enter when the machine indicate this. The same should be done for the 700 IRD labeled products.

Close the top compartment and then close the door of the LICOR machine and press enter when the machine indicate this.

Let the gel run for 3 to 4 hours

The LICOR will shut down automatically.