MSc Major Thesis in Plant Breeding (PBR-80436)

Temporal and Spatial Gene Expression during Turnip Formation in *Brassica rapa*

Supervisor:

Dr. Ir. Guusje Bonnema

Examiners:

Dr. Ir. Guusje Bonnema

Dr. Ir. Christian Bachem

Wageningen UR

Laboratory of Plant Breeding

January, 05, 2016

Stefan Petrasch

Abstract

Brassica shows extreme morphological diversity and similar morphotypes evolved in parallel in different species. An example are tuber forming morphotypes, which are found in several species among which the most important species are *B. rapa* (turnips) and *B. oleracea* (kohlrabi). In this work 31 turnip Selective Sweeps from one study and 13 turnip formation QTLs from two mapping populations were investigated to detect candidate genes for tuberization. The approximately 6,000 genes placed in this genomic regions were compared based on two genome wide expression analyses to identify genes with unique expression patterns for turnips. From the 175 resulting candidate genes, for 29 the transcript abundance in hypocotyl/root of three turnip accessions and two non-tuberizing *B. rapa* accessions were determined for six time points around tuber onset. For eight of those genes, the transcript abundance was also determined in leaf samples. In leafs, no expression patterns indicating involvement of genes in tuberization were found. In *B. rapa* hypocotyl/root however, several genes showed expression patterns which indicate a function in turnip formation. Those genes are involved in hormonal regulation, regulation of plant development or cell cycle regulation and are often expressed during time points which do not correspond to previous reports in literature. Furthermore regulatory mechanisms and functions described in literature for those genes suggest altered hormonal effects and micro RNA levels in turnips compared to non-tuberizing morphotypes.
If genes were tested with *B. oleracea,* no expression patterns were detected which indicate involvement of genes in tuberization.

Acknowledgment

First of all, I would like to thank Dr. ir. Guusje Bonnema for her great support during the whole process of experiments and writing the thesis and of course for making the work on this thesis possible at all. The valuable discussion and the great scientific guidance contributed a lot to this work. Furthermore I am very grateful for all the time she spend reading and correcting my manuscript.

For discussion and help during gene selection, valuable comments on my proposal presentation and above all for agreeing to review my Master thesis, I would like to thank Dr. Christian Bachem.

The support during the practical work was mainly provided by the technician Johan Bucher, who I would like to thank for his exceptional engagement and help. He provided essential tips and suggestions for the experiments and therefore contributed greatly the outcome of my work.

Furthermore, I would like to thank Dr. Theo Borm for introducing me to the genome browser, the discussion on Selective Sweeps and of course for his essential help by writing the program which the primer design was based on.

For the help and discussion during gene selection as well as recommendation of helpful online-tools, I would like to thank Dr. José Abelenda.

I am also obliged to Isolde Pereira-Bertram and Dr. Ikram Blilou for the provision of protocols, tips and help regarding the fixation and embedding of samples.

A special thank is due to Dr. Ningwen Zhang, Lin Ke and Dr. Cheng Feng for providing the valuable datasets which this work is based on.

Last but not least, I want to express my endless gratitude to my family for their great unconditional and devoted support during my whole life. Their trust in me is a great motivation.

Table of Content

1 Introduction

The genus Brassica provides some of the most important vegetable crops and displays a tremendous phenotypic diversity ranging from oil seed crops through cabbages, cauliflowers and broccolis to tubers like turnips and kohlrabies (Bonnema et al., 2011; Liu et al., 2014). This large variation makes this genus ideal for breeding, but also for research of the genetics underlying these diverse morphological characteristics.
Among the most important species of this genus is *Brassica rapa* which also shows the extreme phenotypic diversity. A hypothesis is that this species was domesticated independently several times at different geographical locations and following this several important morphotypes such as heading types and turnips have evolved (Zhao et al., 2005). Domestication and especially the artificial human selection lead to the fact that almost all organs of this species have been altered and are used for human consumption as well a feed.
Recently, the genome of this species has been sequenced and the generated data and knowledge facilitate genomic studies (Wang et al., 2011). Furthermore, genetic research in *B. rapa* benefits from the fact that this species is one of the closest crop relatives to the model plant *Arabidopsis thaliana*, which genome is the best investigated among all plants (Langercrantz, 1998; Wang et al., 2011). The genome of *B. rapa* evolved after a genome triplication and some genome remodelling/fractionation, resulting in three subgenomes distributed over the chromosomes which count is 2n=2x=20 (Wang et al., 2011).

Another very important species is *Brassica oleracea*, which shares the characteristic of large morphological diversity with *B. rapa* and even evolved analogous morphotypes to *B. rapa* such as heading and tuber-like organs (kohlrabi). The genome of *B. oleracea* (a white cabbage type) is sequenced as well and the chromosome count is 2n=2x=18 (Liu et al., 2014). The genomes of *B. rapa* and *B. oleracea* evolved both after triplication of a genome of a common ancestor with *A. thaliana.* Therefore, syntenic genomic regions can easily be identified.

Based on the triangle of U (Figure 1), a result of ancient and recent interbreeding of both species is the allopolyploid *B. napus* which is mainly grown for oil seeds and has a chromosome count of the sum of both species, which is 2n=4x=38 (Cheng et al., 2014; Nagaharu, 1935). Also *B. napus* includes morphotypes that form tubers (swedes/rutabaga), similar to turnip and is sequenced as well (Chalhoub et al., 2014).

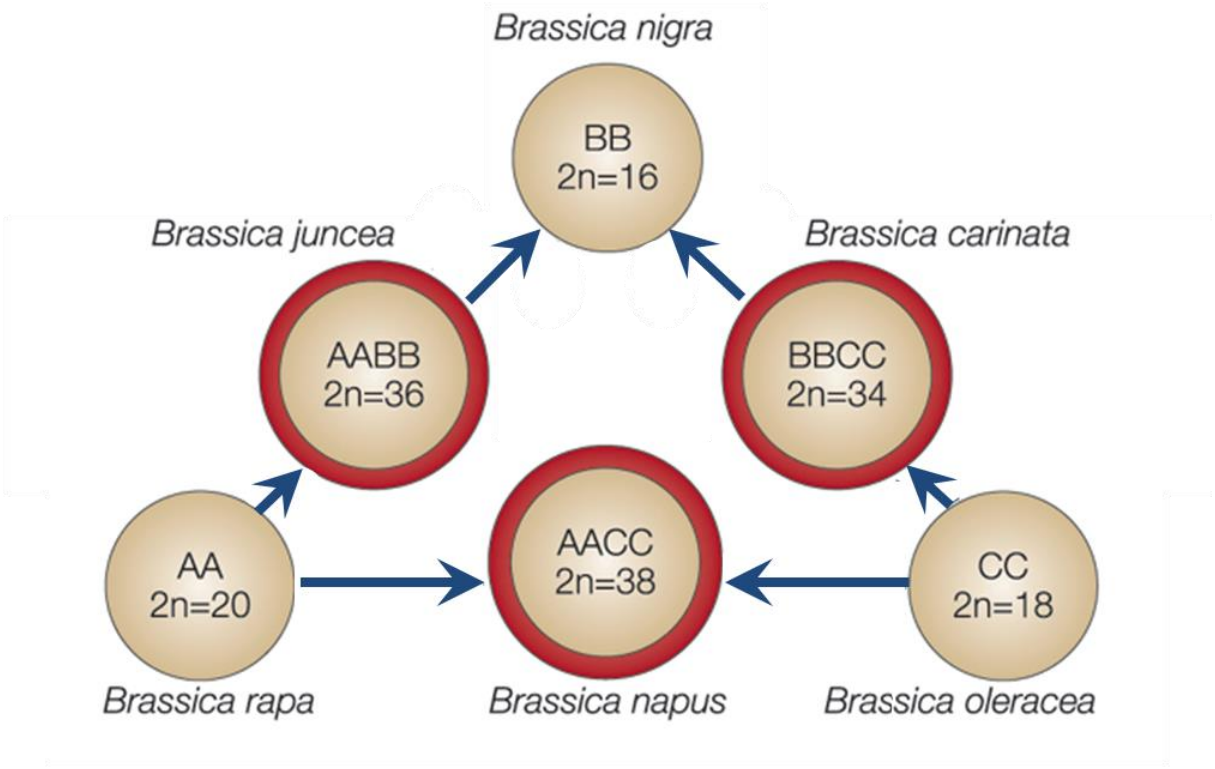

Figure 1: Genomic relationship of six Brassica species (Triangle of U)
The figure shows the three species *B. rapa, B. oleracea* and *B. nigra* with the A, B or C genome. The other three species *B. juncea, B. napus and B. carinata* are results from intraspecific hybridization events, which is also indicated by the number of chromosomes, which is the sum of the chromosome numbers of the parents. Figure modified from Stewart et al. (2003).

Several Brassica species can form tuberous storage organs which are also common in other plants and can be formed from many organs, such as roots, stems or leaves. Tuber forming plants are often domesticated by humans like in the case of potato, beet (Dohm et al., 2013), sweet potato (O'Brien, 1972), radish (Mitsui et al., 2015) and cassava (Elias et al., 2000) as they are a good source of nutrition, mainly carbohydrates. The high amount of different carbohydrates but also other nutrients, such as lipids, proteins or even minerals such as phosphorous or magnesium in the tuber is due to the major function of tubers as storage organs for the plant (Burns, 1946; Chapin et al., 1990). In almost all cases the pure storage function of tubers is combined with other functions. In potato e.g., tubers are also important for vegetative propagation (Abelenda et al., 2011). In annuals and biennials like e.g. Brassica and Beta however, tubers are essential to overbridge the time between vegetative and generative growth. They are serving as energy source for flowering and the production of seeds whereas they do not contribute to vegetative propagation as only one tuber per plant is formed (Chapin et al., 1990).
In cultivated Brassica, tuberous storage organs are formed e.g. by *B. rapa* as turnips but also by *B. oleracea* as kohlrabies (Cheng et al., 2014). In both species, the tuberization process involves swelling of stem tissue and in the case of turnip, also swelling of root tissue (Havis, 1940; Zhang et al., 2014). This makes turnip an organ composed of different proportions of stem and root tissue, whereas kohlrabies are exclusively composed of stem tissue.

As tubers are very important plant organs, a lot of research is conducted to understand the formation of those organs in detail – especially in potato. However, although tuberization in Brassica is very relevant, little is understood so far about the molecular mechanisms underlying this process. The close relationship to *A. thaliana* is also not helping in understanding the process of tuberization in Brassica as Arabidopsis does not form tubers. However, there is a so called turnip mutant in *A. thaliana*, which results from ectopic expression of the LEC1 gene (Casson and Lindsey, 2006). This mutant however does not form any tubers comparable to turnips, neither does it reflect the likely more complex genetics of tuberization in *B. rapa*. In contrast, radish, with a recently sequenced genome, could help understanding this process as it is even closer related to Brassica than Arabidopsis (Mitsui et al., 2015).Nevertheless, radish has the difference to Brassica that all genotypes form tubers and there are no non-tuberizing morphotypes.
Most of the research carried out so far for the tuberization of *B. rapa* was based on anatomical investigations and classical genetic approaches. However, recently a few genomic and transcriptomic studies were performed such as an investigation of the transcriptome during early tuberization stages of turnip.

A recent morphological study reported that enlargement of tubers in a *B. rapa* DH line from a Japanese turnip accession, (VT117) starts around the 28th day after sowing and involves secondary growth of the hypocotyl and parts of the taproot combined with increasing lignification (Zhang et al., 2014).
Furthermore, several publications reported tuberization related QTLs. Lu et al. (2008) for instance detected several turnip QTLs on eight linkage groups using a population derived from a cross of the turnip cultivar 'Qishihai' and the non-heading Chinese cabbage 'Aijiaohuang'. However, these linkage groups were not assigned to chromosomes. Furthermore, for this study anonymous markers (RAPD and AFLP) were used, which makes the exact comparison of the identified QTLs with other described QTLs impossible. Kubo et al. (2010) detected five turnip QTLs on three different linkage groups in a population derived from a cross between a Chinese cabbage inbred line (Y-54) and a vegetable turnip DH line. Lou et al. (2007) detected one major QTL for turnip formation which was located on the top of linkage group A02, together with the flowering time QTL (FLQTL-2). The candidate gene for flowering time and vernalisation response in this QTL is *Br-FLC2*, a paralogue of the Flowering Locus C in *A. thaliana* (Zhao et al., 2010). The parents of this populations were the Yellow Sarson YS-143 and the Asian vegetable turnip VT-115. Considering connection of genetic control of flowering regulation and tuber formation in potato (Abelenda et al., 2013), this result may indicate a similar mechanism in *B. rapa.* Flowering time can also correlate with tuberization by the fact that early flowering reduces the time for tuber formation due to an early competition for resources. However, due to multiplications of genomes, orthologues of genes can evolve to control different developmental processes. In potato e.g., different Flowering Time orthologues control tuberization and flowering (Abelenda et al., 2013). In Brassica, *BrFLC2* however is definitely a main regulator of flowering time (Xiao et al., 2013). In summary, several QTLs for turnip formation were detected but the data is not totally comparable, especially the one of Lu et al. (2008) due to different assignment of linkage groups and presence of more linkage groups than the 10 *B. rapa* chromosomes.

A recent genomics study involving resequencing of a turnip genome and comparison to the reference genome as well as a rapid cycling genome revealed an enrichment of Class III peroxidases in the turnip unique genes within the turnip genome whereas glucosyltransferases had a lower copy number (Lin et al., 2014). Peroxidases are involved in cell wall modifications and are also suggested to be involved in the tuberization process of potato (Francoz et al., 2015; Passardi et al., 2004; Willemsen, 2014). Further investigation of these over- and underrepresented genes revealed involvement in phenylpropanoid biosynthesis pathways in *A. thaliana* (Lin et al., 2014). These results suggest an increased lignin production during tuberization, which corresponds to the previously described lignification during turnip formation (Zhang et al., 2014).

However, as previously mentioned, recently also the genome wide changes of transcription were investigated (Bassetti, 2015). Furthermore, recently a genome wide expression analysis comparing different organs and the morphotypes Pak Choi, Yellow Sarson and turnip was performed by Lin Ke. The performed gene expression analysis during the first six weeks of turnip formation revealed differential expression of many genes which could be assigned to five major clusters based on the expression patterns. The two phases of turnip growth (seedling before day 21 and turnip enlargement after day 28) were also represented in the transcript information. The transcription profile showed large variation but obvious changes between day 21 and day 28. Furthermore, the results of this analysis also indicated the importance of peroxidases but also of sugar metabolism. Involvement of sugar metabolism in tuberization was also reported for potato, where even sugar rich medium can induce tuber formation in vitro (Aksenova et al., 2012; Dobranszki et al., 2008; Perl et al., 1991). This importance of sugars for tuberization was also indicated by in vitro experiments with turnips, where tuberization was induced by sucrose but also by several hormones (Nishijima et al., 2005; Peterson, 1973). Guan (2009) reported that leaves, which were excised from turnip plants grown on high sucrose formed tubers on cut surface on low sucrose, indicating storage of sugar in leaves in an amount sufficient for initiating tuberization as well as sugar as a trigger of tuberization. This furthermore indicates that sugar is a trigger for tuberization and that turnip tubers can be induced on organs that normally do not form tubers.
Guan (2009) also investigated effects of hormones on tuberization in turnips. Even though results were not absolutely clear, there was still indication that for instance Gibberellin prevents tuberization. Research of Ningwen Zhang furthermore showed positive effects of auxin on tuberization (Zhang, personal communication). However, some genotypes did not respond well to hormone application, which also indicates differences in internal hormone levels.

Another promising approach to identify important genes underlying the process of tuberization in turnip was performed by investigating selective sweeps in the turnip genomes (Lin Ke & Cheng Feng, unpublished). Cheng Feng for instance identified 31 Selective Sweeps on all chromosomes which were detected using PiHS signals. Selective sweeps are genomic regions which were favoured by artificial human selection during domestication and therefore show lower variability in the target group than other regions of the genome (Nielsen,

2005). Most likely these selective sweeps harbour genes which are essential for tuber induction, but also other turnip domestication related traits such as expansins for letting tubers grow larger. Nevertheless, due to genetic linkage, there are also some other so called hitchhiking genes, which were not under selection, placed in those selective sweeps (McVean, 2006). A similar investigation was also performed in *B. oleracea* where selective sweeps were identified for several morphotypes including kohlrabi (Hendrikse, 2015).

Although the tuberization process is described well by molecular and anatomic studies and several candidate genes or genomic regions are identified which may be involved in tuberization, still little knowledge is available on which genes are responsible for initial signals for tuberization. Furthermore, the candidate genes are only roughly characterized and investigated and no information is available on their detailed spatial expression patterns.

This work therefore aims to identify genes which are essential for turnip formation by combining and expand existing information. The data from turnip related QTLs shall be combined with the data from the expression analysis of genes during turnip formation and the information on Selective Sweeps, with the aim to choose a subset of genes for further investigation. Furthermore a similar but less comprehensive approach for kohlrabi will aim to determine if the genes responsible for tuberization in *B. oleracea* are the same or similar like those in *B. rapa* or whether the genetics underlying tuberization in this species is different. This information shall be verified via expression analysis of the selected subset of genes during growth and tuberization in a diverse set of turnip- and kohlrabi genotypes to investigate whether they share the same expression patterns during tuberization. Furthermore, the patterns will also be compared to those of non-tuber-forming Brassica morphotypes to reveal differences.

2 Material and Methods

2.1 Plant Material

The turnip gene bank accessions (see Table 1) were selected based on several criteria, one of which was the putative genetic distance, to avoid similar results due to relatedness. Therefore, the selection of turnips included European as well as Japanese (often called Asian) types. Additionally, the turnip accession VT-117 was included, as this accession has been used for genome wide expression analysis, resequencing and other relevant experiments. However, due to limited time and other resources including the amount of available seeds, the number of plants needed to be kept rather small. Therefore, three turnip acessions (VT-117, one Japanese turnip VT-012 and one European turnip VT-053) were selected as well as two *B. rapa* non-turnip accessions. As non-turnip accessions the Yellow Sarson YS-143 and the Pak choi PC-175 were used. To ensure genetic uniformity of the turnips, only doubled haploids (DH) generated from gene bank accessions were initially selected. However, due to depletion of seed stocks, it was necessary to use accessions instead of DH lines for VT-117 and YS-143. Furthermore, one kohlrabi (Bejo – F1 cultivar kolibri) and one non-kohlrabi *B. oleracea* (the Bejo – F1 kale cultivar Bonanza) were selected based on availability of seeds. As DH-lines are not easily available for these morphotypes, F1 hybrid cultivars were used, because these type of cultivars are also uniform genotypically.

Table 1: Brassica accessions used in this work

The first column indicates the Brassica species, whereas the second column indicates the morphotype. In the third column the name of the accession is indicated and whether it is a F1-hybrid. The fourth and fifth column indicate the official accession number as well as the number used in the WUR database. The origin is named in the last column. Abbreviations: No. – number.

Species	Morphotype	Name	Accession No.	WUR No.	Origin
B. rapa	Turnip	VT-012	CGN06720	BrDFS_A_101	Japan
B. rapa	Turnip	VT-053	CGN07167	BrDFS_A_145	Germany
B. rapa	Turnip	VT-117	CGN15201	pv-Br020120	Japan
B. rapa	Pak Choi	PC-175	VO2B0226	BrDFS_A_055	China
B. rapa	Yellow Sarson	YS-143	FIL500	BrDFS_A_137	USA
B. oleracea	Kohlrabi	Kolibri F1			
B. oleracea	Kale	Bonanza F1			

All plants were grown in the greenhouse. To obtain comparable results, the plants were grown under as uniform conditions as possible. Furthermore, a block design of three blocks was used to minimize influences of environmental effects. In one block all accessions were grown with four replicates per time point. Within the block, positions of plants were randomized. All plants were grown in the same soil with the same water supply. To enable uniform light and temperature, all plants were grown in the same greenhouse compartment. Each plant was grown in a pot that should provide enough space for the plant until harvest but also enable saving of space. This choice was a 9 cm pot. Seeds are sown directly in one pot to avoid transplanting as it may alter gene expression due to stress, damage roots and result in differences in planting depth.

At each time point plant material for RNA isolation was collected from nine plants per accession. The material from three plants (from the same block) was bulked to eliminate

outlier effects to a certain extent, resulting in three biological repeats (one per block) per accession and time point. In *B. rapa* the time points were seven days apart, as these time points are commonly used and turnip formation occurs around week three and four (Zhang et al., 2014). However, depending on the start of tuber formation, two more flexible time points were placed within the seven days before the first tubers showing radial swelling were harvested. Those time points are 3 (time point 3.6) and 5 (time point 3.3) days prior to the first regular time point where tuberization was observed (see Figure 2). Via delayed sowing of those intermediate time points it was possible to harvest them together with the regular time point where tubers were first visible. The last harvest took place six weeks after sowing in PC-175, VT-117 and VT-053. Due to no tuber formation in VT-012, the harvest of the last time point including the intermediate time points of this genotype and the control YS-143 was delayed to week 8. The first five time points (week one to week five) were however harvested weekly.

In *B. oleracea* the time points were 14 days apart, as kohlrabi develops slower than turnip and stem swelling starts around seven weeks after germination (Selman and Kulasegaram, 1967). Weekly harvesting therefore would result in too many plants that need to be grown. Similar to *B. rapa*, two time points were placed within the 14 days before the first swollen tubers were harvested. Those time points are 6 (time point 3.6) and 10 days (time point 3.3) prior to the first regular time point where tubers are observed. The last harvest of kohlrabi took place 8 weeks after sowing, as already after 60 days, the tuber diameter is usually already above 3 cm (Li, 2015). Therefore, there were eight time points in total in *B. rapa* and six time points in *B. oleracea* to represent several developmental stages.

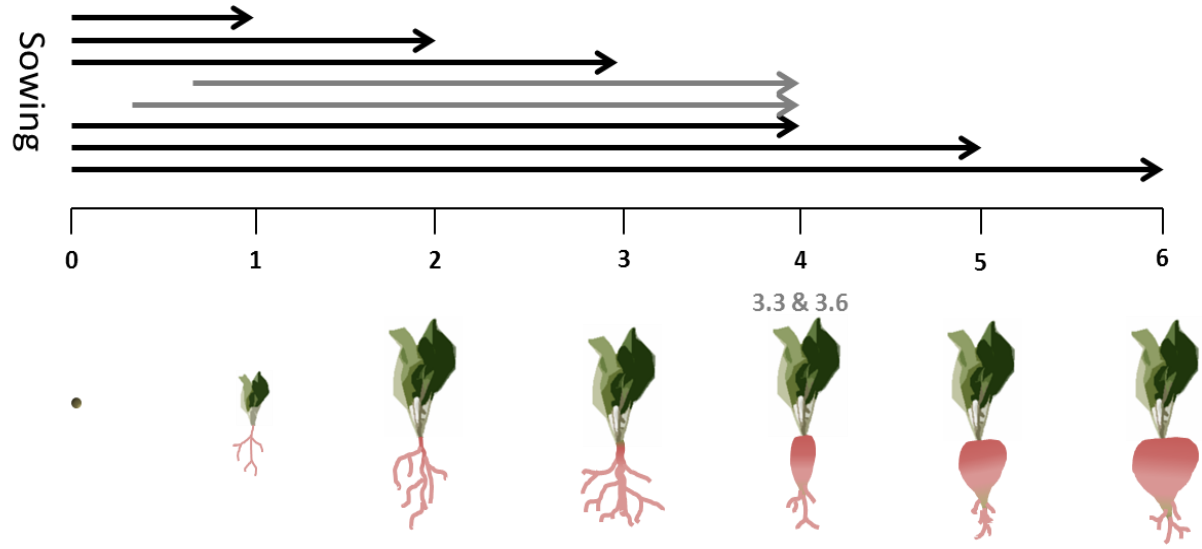

Figure 2: Illustration of experimental time points
The scale and the arrows indicate time points in weeks after sowing when samples were collected. The grey numbers as well as the grey arrows correspond to the flexible intermediate time points. Illustrations below the scale illustrate schematically the developmental stage of turnips in time with an example initiation of swelling in week four after sowing.

The plant material was frozen in liquid nitrogen directly after harvest to enable isolation of RNA later on. The plant parts which were harvested are the tubers (without apical meristem)

or the corresponding root/hypocotyl parts for *B. rapa* and similar proportions of roots as well as the whole stem (not including shoot apical meristem and leaves) for *B. oleracea*. In the following parts of the work, those samples will be called 'root' samples, although this is technically not correct as always stem tissue is included. The inclusion of the whole stem and some roots was necessary for *B. oleracea* as tubers are formed in the upper part (above cotyledons) of the stem. Only harvesting this part would therefore not enable a reliable comparison with *B. rapa*. Leaves were harvested from both species in the same manner to enable profiling expression of genes in this plant organ, as some genes are expressed in other plant parts and the transcripts or signal molecules are transported to the target region afterwards (Citovsky and Zambryski, 2000). Two leaves were harvested per plant – the oldest not yet senescing leaf as well as a young, not fully expanded leaf. As the tubers or the stem parts which likely develop tubers later on were harvested, the harvesting method was destructive. Therefore for isolation on RNA, 72 plants (allowing harvesting of 72/9=8 time points) were grown per *B. rapa* accession and 63 plants were grown per *B. oleracea* hybrid. To deal with potential problems, some more plants were grown. All seeds were sown at the same time except from the seeds which are necessary for the two intermediate time points, which were sown three and five days later.

For in situ hybridization, it is also necessary to harvest tuber or hypocotyls/roots and the corresponding tissues of young plants, respectively. Therefore, additional plants were grown. For each accession and time point, three biological repeats are preferred. The time points were the same as those of harvest for the purpose of RNA isolation to enable comparison of results. Thus, for each *B. rapa* accession, 24 plants were grown, whereas for *B. oleracea*, 21 plants per hybrid were grown. Also for these experiments, some more plants were grown to ensure availability of a sufficient number of plants. As fixed tissue can be stored for several months in 70 % ethanol at 4 ºC (comment of Jeff Long on Jackson, 1992), samples were harvested and fixed at the same time points as for RNA isolation and then stored until embedding. All seeds were sown at the same time as for the purpose of RNA isolation.

2.2 Initial Data Analysis and Identification of Candidate Genes

The available data which this project is based on is the genome wide gene expression analysis during the first weeks of turnip formation in *B. rapa*, data on selective sweeps in *B. rapa* (turnip) and *B. oleracea* (kohlrabi) as well as data on tuberization related QTLs in turnip. Furthermore, a genome wide gene expression dataset comparing four different organs and the germination stage in turnip, pak choi and yellow sarson was used.

The data of the expression analysis was generated by Ningwen Zhang with a microarray experiment as described by Bassetti (2015). RNA extracted from turnips or the corresponding hypocotyl/root parts at weekly time points and two biological repeats was hybridized in a self-self-design. The microarray data were then analysed by Niccolo Bassetti. This investigation revealed the expression patterns of around 61,000 probes (representing about 40,000 genes) during turnip formation in the DH turnip accession VT-117. Furthermore, the analysis revealed several potentially important genes involved in the tuber development. These genes

were selected based on promising expression patterns but also on predicted function of the gene product.

Data of selective sweeps is available for both turnips and kohlrabi. From Cheng Feng (IVF-CAAS) data were based on a large set (199 *B. rapa*, including 54 *B. rapa ssp. rapa turnips)*. Based on these data sets, at least 31 selective sweeps of varying size on all ten chromosomes were identified by piHS. Data for selective sweeps based on the same principle was also generated by Yonina Hendrikse for kohlrabi using a population of 16 kohlrabi accessions (eight from Rijk Zwaan and eight from Bejo) and 105 non kohlrabi *B. oleracea* accessions. In this investigation 37 selective sweeps specific for kohlrabi were identified on all chromosomes except for chromosome 6.
Some promising genes based on gene ontology were already selected in both datasets but in this project, still all genes in the selective sweeps were taken into account.

QTLs for turnip formation were reported in several publications (Kubo et al., 2010; Lou et al., 2007; Lu et al., 2008) and unpublished data by Ningwen Zhang is available as well. In this work however only the QTLs detected by Ningwen Zhang were used. One set of QTLs was based on a RIL population derived from a cross of the vegetable turnip accession VT-115 and a Wutacai accession. The other set of QTLs was based on a cross of the vegetable turnip accession VT-115 and the rapid cycling accession RC-144.Combining those two dataset, turnip QTLs are located on nine chromosomes (1, 2, 3, 4, 6, 7, 8, 9, 10).
However, as the reported QTLs are usually larger than the selective sweeps, only those QTLs were completely analysed which did not overlap with an piHS based selective sweep of Cheng Feng. If a QTL was overlapping with a Selective Sweep, only the overlap region was investigated. This decision was based on the need to reduce the number of genes for closer investigation.

Combining data of selective sweeps and QTLs, the total number of genes covered by those regions is approximately 6,000 genes (15 % of all genes). This number does not include genes located in regions of QTLs that did overlap with selective sweeps and also does not include genes located in selective sweeps detected by other signals.
The previously mentioned 6,000 genes were then closer investigated to define whether they might be promising candidates for turnip formation. The first step was an investigation of the paralogues with the help of the genome wide expression analysis dataset of Ningwen Zhang and Niccolo Bassetti. Usually due to the genome triplication, every gene should have two paralogues. Therefore, it may give hints to see whether only one of those three paralogues is situated in the selective sweeps or QTLs and whether this one is expressed differently than the others. If there was a clear difference (>10 %) between expression of a gene in selective sweeps or QTLs and its paralogues which are not in those regions, a gene was considered as potentially promising. For genes with no paralogues, this step was skipped.
The next step conducted with the promising genes was the comparison of expression in turnip and the two other morphotypes using the dataset of Lin Ke. Here, genes were considered more promising when their expression in turnip was clearly different (>10 %) from the expression in the other two morphotypes in at least one organ or the germination stage. However, a

difference in the expression in flowers was considered less promising than differences in other organs as flowers occur after initiation of tuberization.
The last theoretical step to reduce the number candidate genes was a slightly more detailed investigation to determine whether those genes have the potential to be upstream signals, e.g. by analysing interaction and association of the gene products with products of other genes. This information was obtained from databases, such as STRING (Szklarczyk et al., 2015) but also from literature on Brassica, Arabidopsis or other plants. However, final gene selection also took intuition into account suggestions of colleagues working on potato that selected genes, which are involved in the process of tuberization in potato.

2.2.1 Identification of Reference Gene

A list of reference gene candidates was created by the following steps to obtain a list of genes which are stably expressed over all time points and among different morphotypes and tissues. The first two steps were based on the genome wide expression analysis dataset which compares the three different morphotypes, whereas the third step was based on the genome wide expression analysis of hypocotyl/root or turnip tissue during the six first weeks after germination of the Japanese vegetable turnip VT-117.
For the first step, the maximum expression difference within an organ or germination stage and between the three morphotypes was calculated. In the second step the maximum expression difference between organs and germination stage as well as between the three morphotypes was calculated. The third step was the calculation of the maximum difference of expression in the Japanese vegetable turnip VT-117 during the first six weeks of growth after germination.
These three values were averaged for each gene and divided by the average of the six expression values of the corresponding gene during the first six weeks of growth of turnip to relate the stability of expression to the amount of expression, as a reference gene should ideally be rather high expressed. Based on the last mentioned calculation genes were ranked from small to large values and compared to the publication of (Xiao et al., 2012) to find the rankings of genes which were used as reference genes in *B. rapa* before.

2.2.2 Primer Design

Due to the large number of genes, primer design was partially automated by Dr. Theo Borm. Sequences of the genes were blasted against the *B. rapa* genome to find the paralogues and to identify regions in the exons which were rather unique for this gene. For *B. oleracea*, the *B. rapa* gene was blasted to the genome using the same method.
For primer design, those regions of the genes were used which are rather unique for this specific gene. Optimally one exon was chosen as target for primer design to reduce the possibility of differential splicing of the amplified sequence. This exon sequence was used as input for Primer3Plus (Untergasser et al., 2007) with the special settings qPCR. From the proposed primer pairs, this pair was chosen which both amplifies the *B. rapa* gene but also the most similar *B. oleracea* gene. Furthermore, only the pair of primers was chosen which was indicated to only amplify this one gene but no paralogues in each species. This indication

was retrieved from a blast with the algorithm blastn against the genome of the two species via the website ensemblPlants (Cunningham et al., 2015). For list of the final primers see Appendix Table 1. For the reference gene Cyp (Bra037296) the primers described in Xiao et al. (2012) were used.
Primer specificity was experimentally tested via RT-qPCR melting curves and in cases where RT-qPCR melting curved indicated unspecific binding also via agarose gel electrophoresis.

2.3 Experiments

As soon as the putatively essential genes were identified, the information was used to further characterize the genes by analysing them based on newly grown plants. The experiments were based on transcriptomics and molecular biology.

2.3.1 RNA Isolation, DNase treatment, quantity & quality control and cDNA synthesis

RNA was isolated from frozen and manually grinded tissue using the RNeasy® Mini Kit of the company Qiagen (Milden, Germany) following the instructions of the Quick-Start Protocol. For lysis of leaf samples, RLT buffer was used, whereas for lysis of 'root' samples, RLC buffer was used. RNA was eluted in 30 µl RNase-free water with a repetition of the elution step using the previous eluate. Remaining DNA was eliminated via DNase digestion using the Kit DNase I, Amplification Grade (Invitrogen) following the manufacturer's instruction with the following modifications: Instead of the recommended reaction mix, 1 µl DNase I Reaction Buffer and 1 µl DNase I were added to the approximately 30 µl RNA solution and the reaction was stopped using 1 µl 25mM EDTA. Presence of DNA in the RNA samples was tested via RT-qPCR using primers of β-Actin. If amplification resulted in a higher cq-value than 33, the DNase digestion was repeated.
The amount of RNA in the solution was quantified using a Nanodrop 1000 spectrometer (Thermo Scientific) and quality of isolated RNA was checked via gel electrophoresis using a 1 % agarose gel. RNA was then stored at -80 ºC until cDNA synthesis.
After quantification and quality check, RNA was conversed to cDNA using the iScript cDNA Synthesis Kit (Bio-Rad, USA) following the manufacturer's instruction. If the RNA concentration was too low to obtain a total concentration of 1 µg RNA in the reaction mix, a total concentration of 0.5 µg was used. After cDNA synthesis, the reaction mix was diluted 80 times, or 40 times for samples, where only 0.5 µg RNA were used for cDNA synthesis. The diluted cDNA was stored at either -20 ºC or 4 ºC depending on the frequency of use.

2.3.2 Quantitative Real Time PCR

RT-qPCR was performed following the description of Basnet et al., (2013) with the cDNA prepared as described above. However, the reaction mix was different, as 10 µl mix were composed of 5 µl iQ™ SYBR Green Supermix reagent (Bio-Rad; Hercules, California, USA) mixed with 4 µl cDNA sample (80x diluted), 0.4 µl filter-sterilized MQ-water and 0.3 µl forwards as well as backward primer.
As mentioned above, Cyp was used as reference gene for all samples and due to time and financial limitation, no technical repeats were prepared.
The following PCR-program was used:

Temperature	Time in min	
95 ºC	3:00	
95 ºC	0:15	40x
60 ºC	0:30	
95 ºC	0:10	
65 ºC	0:05	
95 ºC	0:05	

The data recorded by the cycler and implemented in the CFX Manager file was loaded to Excel to calculate ΔCq as well as the Relative Expression (= $2^{-\Delta Cq}$).

2.3.3 Fixation and Embedding of Plant Material

Spatial expression of genes or localization of gene-products can be investigated by In-Situ Hybridization (ISH). These experiments can reveal/validate transcript abundance, localization and temporal patterns. Several methods are described in the literature, ranging from methods without embedding (Borlido et al., 2002) to different fixation and embedding methods (Braissant and Wahli, 1998; McFadden et al., 1988; Meyerowitz, 1987; Urieli-Shoval et al., 1992). For this project, embedding is likely to be the best choice due to long storability of the end product. As embedding requires proper fixation of the material, fixatives containing paraformaldehyde or similar substances are the best choice, as they enable cross linking of proteins and good maintenance of the amount and localizations of RNA.

The procedure followed for fixation and embedding of the samples is roughly those described by the Dolf Weijers Lab (Appendix B) which is based on the protocol of Jackson (1991) and adapted by Saiga et al (2008). Due to the large size of tissues, several steps were adapted as follows. Turnips or corresponding root/hypocotyl parts were cut into three pieces (lower root part, middle part and upper part including cotyledons) and also cut on one side to enable better penetration of fixative. Older stages of tubers were cut to one approximately 1 cm thick slice containing the central vessels. They were then immediately placed in ice cold fixative where they remained until the vacuum infiltration. During the Ethanol series, 70 % Ethanol was renewed after one to two nights to enable better storage for several weeks at 5 ºC. To enable better Xylene infiltration, samples were left in 100 % Xylene plus some added paraplast chips for at least 24 hours. Furthermore, to enable good infiltration with paraplast, samples were left in molten paraplast for six days with paraplast exchanges twice a day for the first three days and one more exchange at the day prior to embedding. Samples were embedded in molds which are usually used for Technovit® embedding. Larger samples however were embedded in small petri dishes.

RNA and DNA maintenance was tested with Acridine Orange staining. Samples were covered with some drops of Acridine Orange Solution and were investigated under UV-light immediately to 10 minutes after staining. The Acridine Orange solution was composed as follows: 6 µg/ml Acridine Orange, 0.09 M citric acid, 0.02 M Na_2HPO_4.

3 Results

3.1 Plant Growth and Development

All accessions germinated fast and uniform within a few days. Kohlrabi however, needed more time for germination and germinated more or less after a week. Seedling growth was fast and rather uniform. After two weeks, it however became clear that plants of the DH line VT-053 are generally smaller and appear weaker than e.g. plants of the same age of the DH line VT-012 or the accession VT-117. Nevertheless VT-053 had clear tuber onset below the soil surface in week 4 and therefore formed tubers at the same time as VT-117 which however formed tubers above the ground. VT-012 however, although always slightly more vigorous than VT-117 didn't form any tubers during the eight weeks of observation. However, the roots thickened to some extent but always clearly remained in the shape of a root. Furthermore, the morphology of the plants of the DH VT-012 was less uniform compared to what was observed in the other DH-lines or even the accession VT-117. Plants strongly differed in colour and some plants had a high density of trichomes, whereas others completely lacked trichomes (see Figure 3 A and B). Plants of DH PC-175 (Pak Choi) and the accession YS-143 (Yellow Sarson) developed fast and uniform and Yellow Sarson started flowering already in week 6 to 7.

Kohlrabi and Kale also both developed fast and uniform. Kohlrabi stems clearly started to swell in week 6 and usually tubers were set above the fourth leaf (counted upwards from the cotyledons). Therefore, in contrast to *B. rapa*, plants of *B. oleracea* at early time points (before week 4) did not yet form the tissue which later on forms tubers.

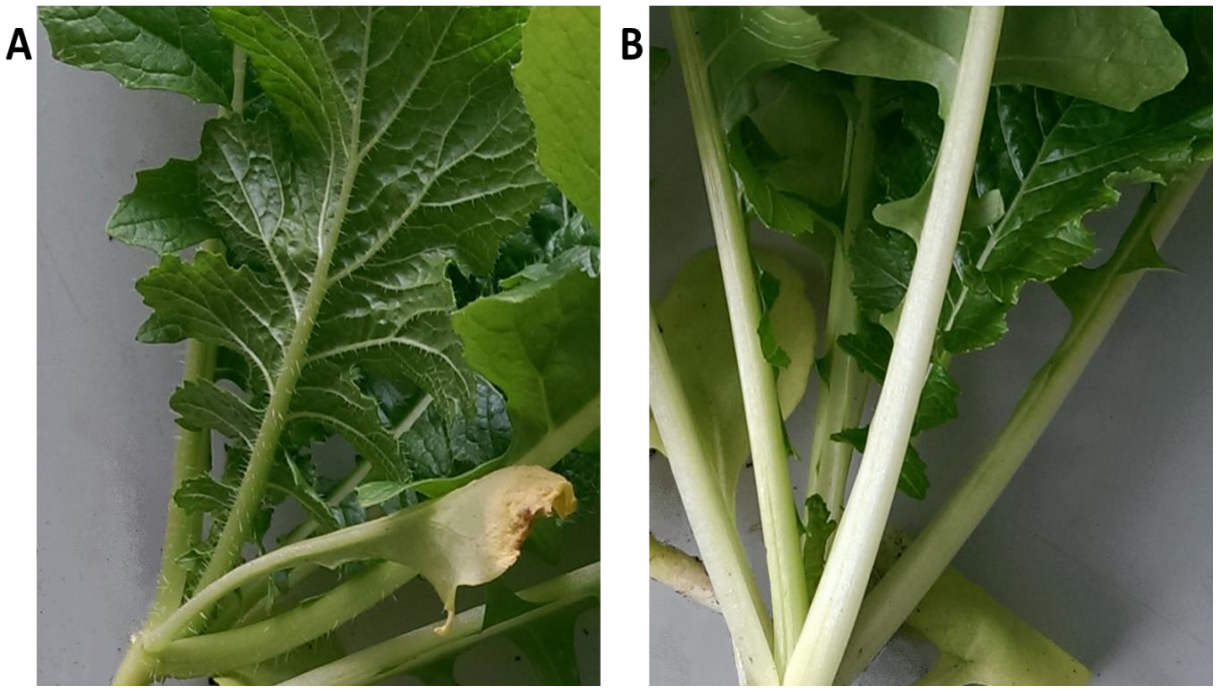

Figure 3: Lack of uniformity in plants of DH VT-012
Both plants are from the same block and the same DH-line at time point 4. On the plant in (A) many trichomes are found both at the upper and lower side of leaves and petioles, whereas the plant in (B) didn't have any visible trichomes.

3.2 Candidate Gene Selection

Initial investigation of the provided datasets for *B. rapa* on Selective Sweeps and QTLs resulted in a total of 31 Selective Sweeps (detected via piHS), 6 QTLs overlapping with some of these Selective Sweeps and 13 QTLs which are not overlapping with any Selective Sweeps (for details see Appendix Table 2). Those regions cover all ten chromosomes (see Figure 4) and in total include approximately 6,000 genes (15 % of all *B. rapa* genes). If compared to the chromosomal distribution of the main B. rapa genomic features (Wang et al. 2011), most of those Selective Sweeps and QTLs are located in rather gene-rich genomic regions.

Due to the sheer number of genes and time limitation, Selective sweeps detected by other signals than piHS and other QTLs were not investigated.

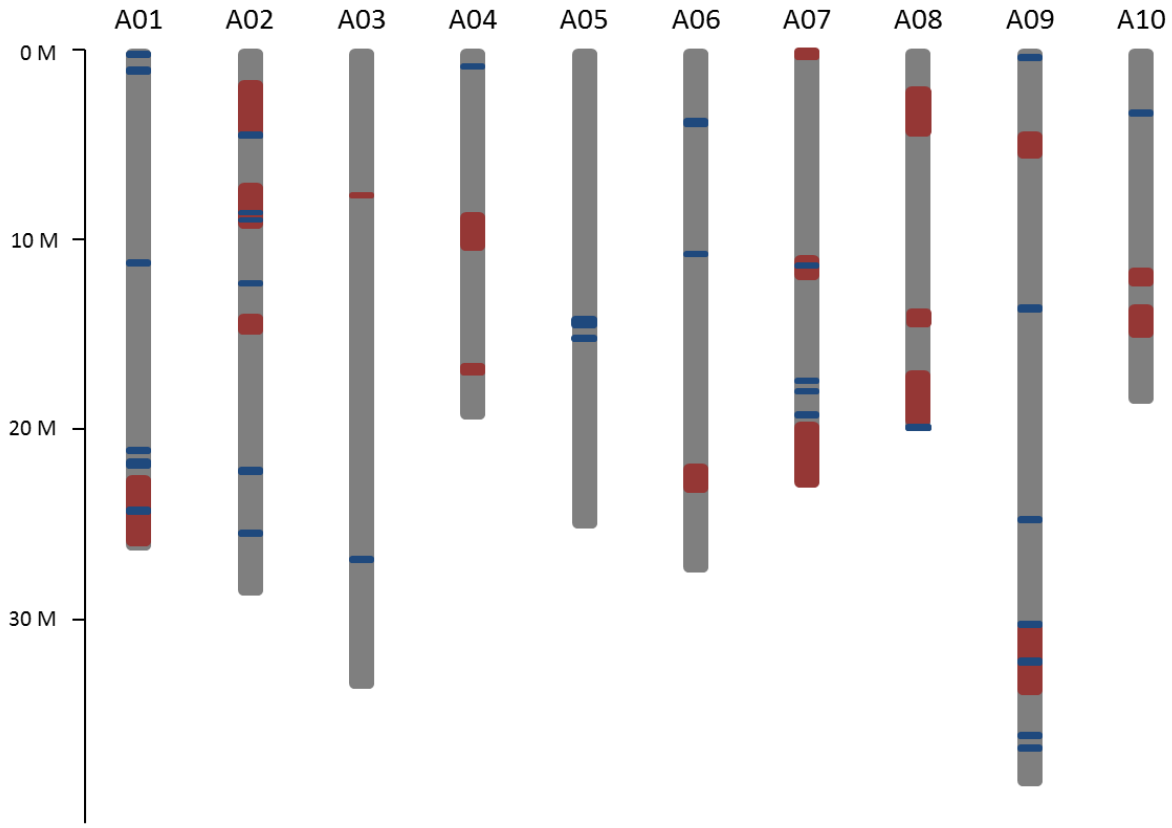

Figure 4: Position of Selective Sweeps and QTLs on *B. rapa* chromosomes
Chromosomes are named as described in Wang et al. (2011). The blue regions indicate the Selective Sweeps which were investigated, whereas red regions indicate QTLs. Only those QTLs were investigated which did not overlap with any Selective Sweep. Chromosomes as well as highlighted regions are on scale in units of million (M) bases.

The 37 Selective Sweeps in *B. oleracea* covered all chromosomes except from chromosome C06 (for details see Appendix Table 3). Based on the fact that five of those Selective Sweeps are at synthenic positions compared to some *B. rapa* Selective Sweeps or QTLs, there is indication of overlap between selected genes in *B. oleracea* and *B. rapa*, although this overlap could also be random.

Based on the mentioned selection criteria 83 genes were selected as potential candidate genes for the 31 *B. rapa* Selective Sweeps, whereas 92 genes were selected as potential candidate genes for the 13 *B. rapa* QTLs. This list of in total 175 genes contains at least one gene per Selective Sweep or QTL. The predicted functions of those genes are quite diverse. The list contains for instance five different expansins and many transcription factors (e.g. five myb-domain transcription factors). Several genes in this list are hormone related, like the three auxin responsive proteins or the two cytokinin trans-hydroxylases. A fraction of those genes are also carbohydrate related, such as sugar transporters. Furthermore, many of the selected genes are reported to be involved in plant developmental processes. However, due to the lack of *Arabidopsis* orthologues there are also several genes in the list with unknown function.

Further reduction of this list via selection of colleagues and predicted function of genes resulted in a total of 60 genes: in this reduced list some Selective Sweeps and QTLs were not represented anymore. The predicted functions of those 60 genes remained similarly diverse, except from the fact that several genes which role is likely downstream in tuberization were eliminated. This for instance lead to the complete lack of expansins in the reduced list. This reduced list is however stronger enriched in transcription factors and hormone related genes as their function is more likely upstream in the process of tuberization.
Out of the 60 *B. rapa* genes, ten had paralogues in different *B. oleracea* Selective Sweeps.

3.3 RT-qPCR

3.3.1 RNA Quality

The isolated RNA had different concentrations but always a very uniform and good quality as no degradation was visible in any sample. RNA was therefore suitable for cDNA conversion and following this for RT-qPCR. Pattern of mRNA main bands on 1 % agarose gel differed between 'root' and leaf samples, as can be seen from the examples in Figure 5.

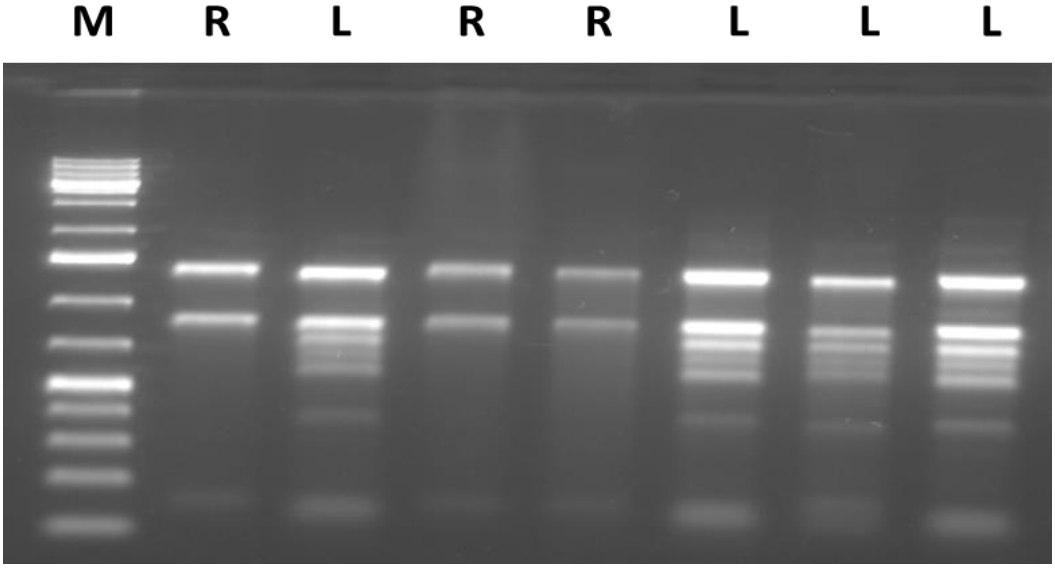

Figure 5: Quality and Concentration of isolated RNA
RNA samples isolated from 'root' are indicated by (R), whereas leave samples are indicated by (L). The first lane shows the 1 kb + marker. Samples were loaded to 1 % agarose gel.

3.3.2 Validation of reference gene

Based on the described calculation procedure for reference gene selection, Cyp (Bra037296) was selected as most optimal for use as a reference gene as it was on rank 215 and therefore was indicated to be very stably expressed. The selected reference gene candidate Cyp was then tested on all samples (*B. rapa* as well as *B. oleracea*). The Cq-values were sufficiently low with a total average of 19.39 (with an average of 19.17 for plate 1 and 19.70 for plate 2), indicating high expression of the gene. The gene had Cq-values that did not differ from the average by more than 1 in almost 90 % of all samples (Figure 6). Therefore, expression of Cyp is stable enough for use as a reference gene. Comparison of the stability of Cyp expression to β-Actin expression on a subset of samples indicated higher stability of Cyp, which also was ranked better. An analysis of the melting curves furthermore indicated that Cyp-primers bind specifically and amplify the same product in all samples (data not shown). Therefore, Cyp was used as a reference gene in the following experiments.

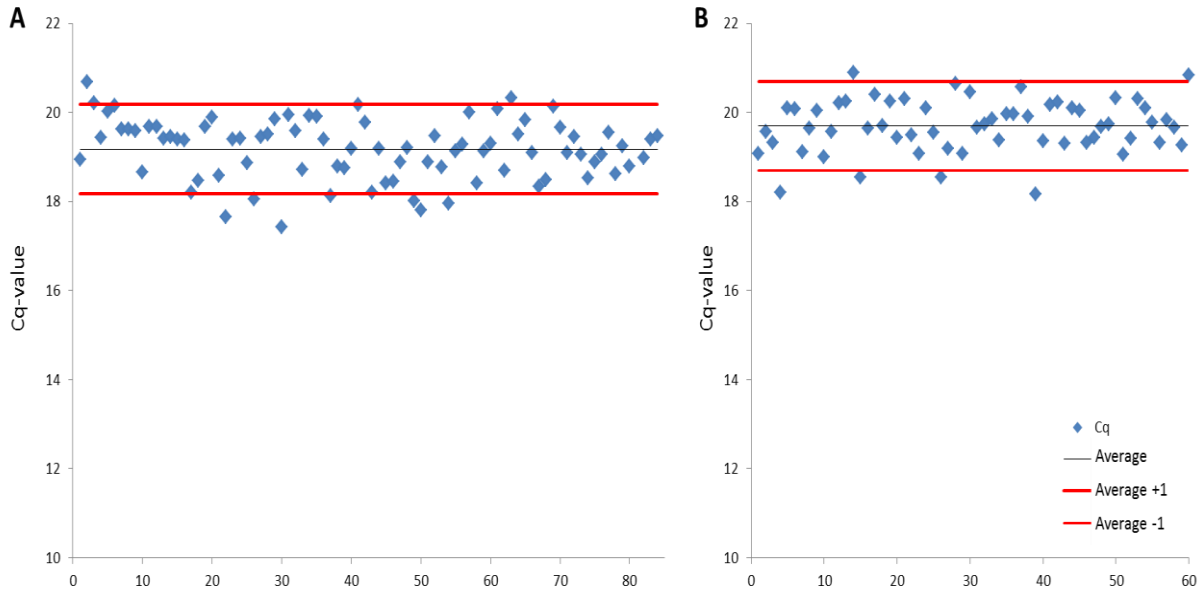

Figure 6: Cq-values of the reference gene Cyp for all samples
The Cq values of the reference gene Cyp for each sample are indicated by a blue square. Graph (A) shows the Cq values of 84 samples, whereas (B) shows the Cq values of additional 60 samples from another PCR plate. The average of the Cq values was calculated separately for each plate and is demonstrated by the black horizontal line. Red lines indicate the area of ±1 around the average.

3.3.3 Initial Testing of 58 genes

To reduce the number of genes, which will be tested over the different accessions, an initial testing was performed for all 58 genes (for gene IDs & information see Appendix Table 4). The number of genes was reduced to 58 since primer design was not optimally possible for two genes out of the 60 selected genes. For this purpose, RT-qPCR was ran for only one time point for the samples VT-117, VT-053 and Pak Choi (both leaves and 'root') without biological replicates. The specific time points differed for genes and were chosen based on

expression patterns of genes during week 1 to 6 in the tubers of VT-117 in the genome wide gene expression analysis by Ningwen Zhang and Niccolo Bassetti (Bassetti, 2015). Those time points were chosen for initial testing which indicated a change in expression. If for instance a gene started to be lower expressed from time point 2 on, time point 3 was chosen. If genes were however stable expressed, usually time point 4 was chosen.

The goal was to find genes showing expression levels which were shared by the two turnips but clearly differed in Pak Choi turnips/roots and leaves (examples in Figure 7). Many genes however did not meet this criterion. Gene 24 for instance has very similar expression levels in 'roots' of all three accessions (Figure 7 A), whereas gene 33 is similar expressed in the 'roots' of VT-053 and Pak Choi but differs in the 'roots' of VT-117 (Figure 7 B). In contrast, genes 34 and 38 are expressed differently in all three accessions (Figure 7 C & D). Similarly, the expression levels of those genes in leaves are also not shared for the two turnips and have different levels in Pak Choi. Therefore, based on the results of initial testing, genes with similar expression characteristics as the previously mentioned examples were not considered as promising. However, some genes showed expression which indicated shared patterns for the two turnips but different expression in Pak Choi (examples in Figure 8). Often those genes are higher (e.g. Figure 8 A, C & D) or lower expressed (e.g. Figure 8 B) in 'roots' of the two turnips than in 'roots' of Pak Choi. Expression levels in leaves, however did not indicate any promising patterns (see Figure 8). Seven of those genes (7, 11, 21, 35, 36, 46 and 49) were chosen for testing with the whole set of samples including biological repeats to verify the expression patterns but also to test whether the results of the initial testing indeed can predict the expression pattern of the genes in the complete set of samples. These seven genes with their predicted function are listed in Table 2.

Table 2: Details for 7 genes selected based on initial testing

The first column (No.) indicates the number with which the gene is identified in the text of this work. The second column shows the official *B. rapa* gene ID, whereas the third column (*A. thal.* orth.) shows the *A. thaliana* orthologue which (predicted) function is displayed in the fourth column. The fifth and the last column indicate, whether the gene is placed in a Selective Sweep (Sel. Sw.) and/or a QTL – A number indicates the number of the Selective Sweep or QTL, (Y) indicates an overlapping QTL and (N) indicates non-presence.

No.	Gene ID	*A. thal.* orth.	predicted function	Sel. Sw.	QTL
7	Bra034022	AT1G67110	CYP735A2 – cytokinin trans-hydroxylase	9	Y
11	Bra029281	AT5G62165	AGL42 – Agamous-like 42	12	Y
21	Bra003743	unknown	unknown	21	N
35	Bra030232	AT2G22490	CYCD2;1 – Cyclin D2;1	N	3
36	Bra032175	AT2G23300	Leucine rich repeat protein kinase family protein	N	3
46	Bra004226	AT1G67440	emb1688 – embryo defective 1688	N	7
49	Bra034889	unknown	unknown (DUF1510)	N	8

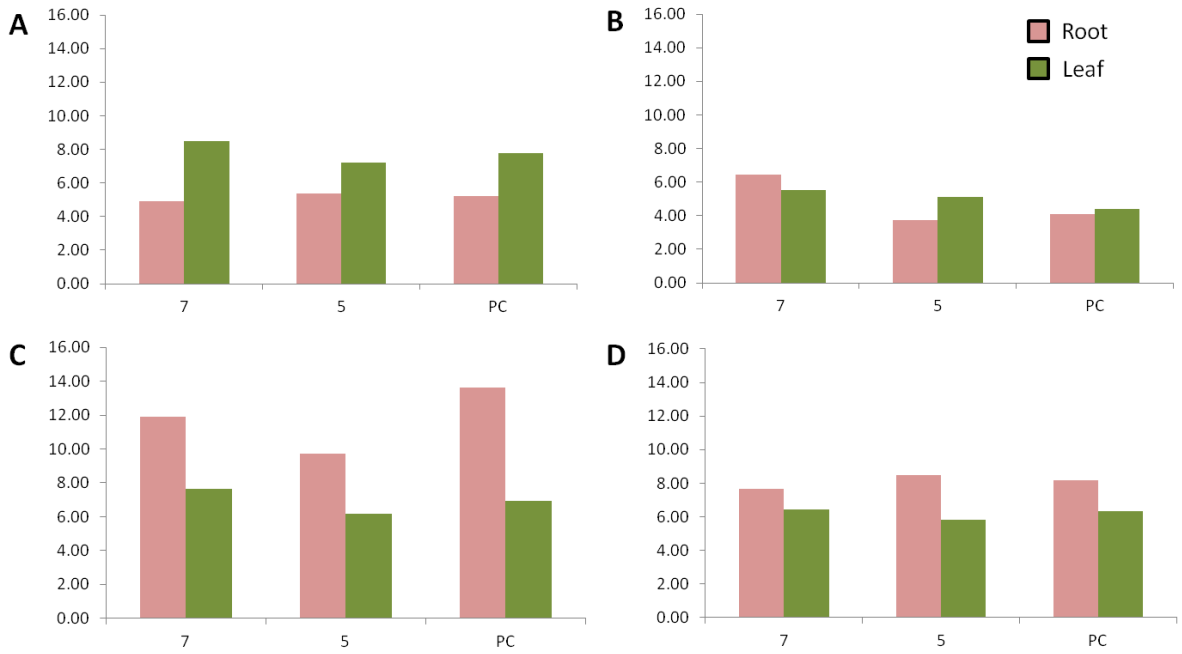

Figure 7: Example of discarded candidates after initial testing
On the X-axis, 7 indicates VT-117, 5 indicates VT-053 and PC indicates Pak Choi. The Y-axis shows ΔCq (relative to reference gene Cyp Bra037296) values of the shown gene: (A) gene 24 (Bra030639 – SKIP16), (B) gene 33 (Bra008495 – tetratricopeptide repeat domain-containing protein), (C) gene 34 (Bra022954 – SPL3) and (D) gene 38 (Bra032056 – DnaJ/Hsp40). Low ΔCq-values indicate high expression. ΔCq-values of 12, or higher are too close to the background and therefore are considered indicate no expression of the target gene.

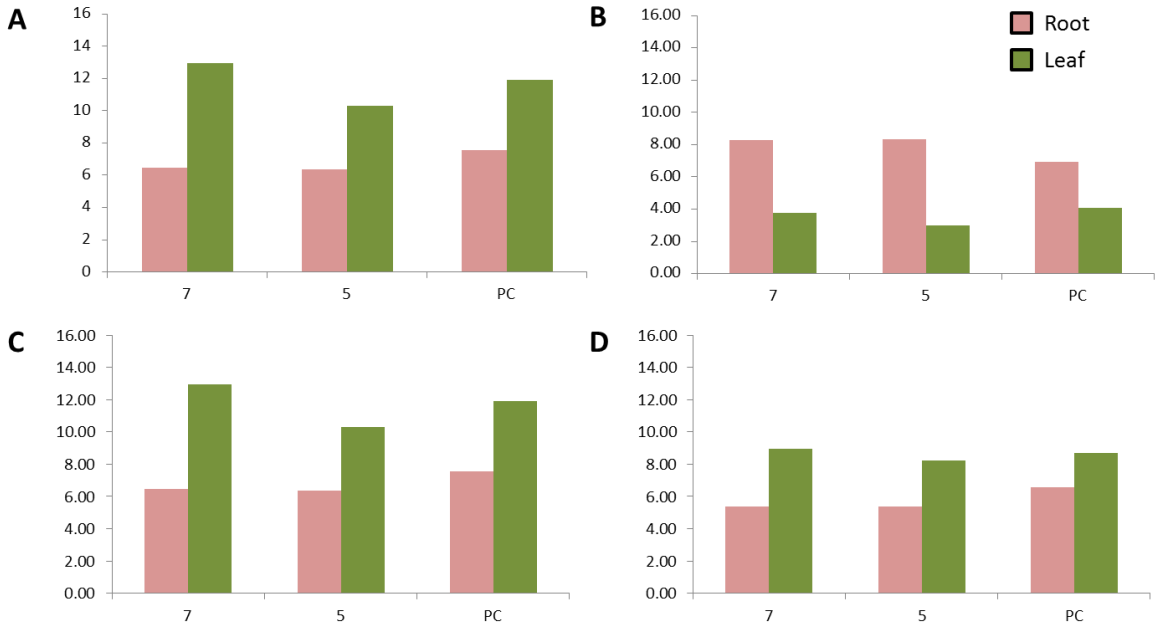

Figure 8: Example of promising candidates after initial testing
On the X-axis, 7 indicates VT-117, 5 indicates VT-053 and PC indicates Pak Choi. The Y-axis shows ΔCq values (relative to reference gene Cyp Bra037296) of the shown gene: (A) gene 16 (Bra019878 – SUB), (B) gene 21 (Bra003743 - unknown), (C) gene 26 (Bra017432 – SGR5) and (D) gene 35 (Bra030232 – CYCD2;1). Low ΔCq-values indicate high expression. ΔCq-values of 12, or higher are too close to the background and therefore are considered indicate no expression of the target gene.

The results of the initial testing were also used to evaluate the quality of primers. Most primers were indicated to amplify specifically only one gene. The melting curve was smooth with one clear peak (example gene 47 in Figure 9 B). A few primer pairs however were indicated to not amplify specifically a single gene as melting curves were not smooth with more than one clear peak (example of gene 8 in Figure 9 A). The RT-qPCR mix of a subset of samples was loaded on a 1 % agarose gel to check the predictive value of the melting curves. In case of single peaked melting curves one band was observed and in case of non-smooth and double peaked melting curves more than a single band were observed (picture not shown). The melting curves therefore reliably predict the specificity and amplification quality of primers. Therefore, those genes with suspicious melting curves were not used for further investigation due to time limitation. An example of such a gene, which initially was promising but had to be discarded is gene 16 (see Figure 8 A for initial testing results and Figure 9 A for the melting curve).

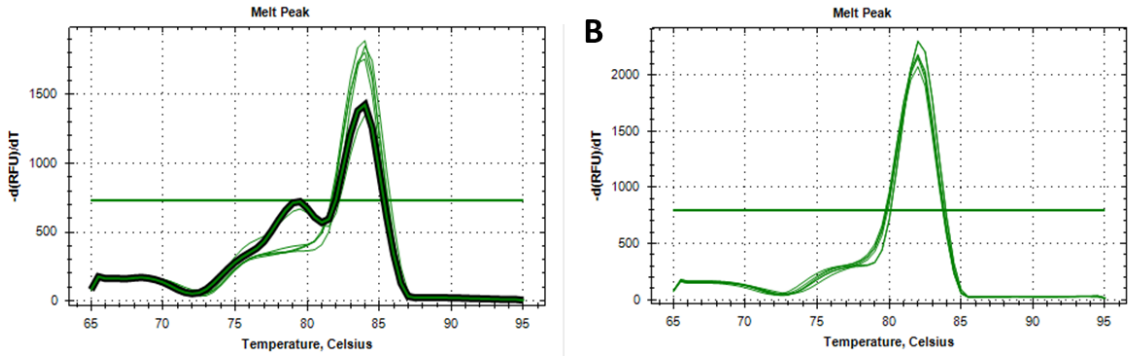

Figure 9: Melting Curves of two exemplary genes
The green graphs indicate the melting curves of a gene in six different samples (VT-117, Vt-053 and Pak Choi, both in leaf and 'root' samples). The graph (A) shows the melting curves of the amplicons of gene 16 (Bra019878 – SUB), which primers did not amplify specifically, whereas graph (B) shows the melting curves of the amplicons of gene 47 (Bra014208 – leucine rich repeat), which primers did amplify specifically.

3.3.4 Detailed screening of selected genes

The expression levels of seven genes selected based on the results of the initial testing were detected in 144 samples. Those samples are the time points week 1 to 3 for both *B. oleracea* accessions kohlrabi and kale, time points week 2 to week 5 for VT-117, VT-053 & Pak Choi and time points 2 to 5 (excluding 3.3 & 3.6) for VT-012 and Yellow Sarson.

It became clear that the expression of genes in Yellow Sarson often was not similar to their expression in Pak Choi. In many cases, such as in the case of gene 21 (others not shown), the expression level of genes in turnip accessions often were similar to those in Yellow Sarson or in between the expression levels of Pak Choi and Yellow Sarson (see e.g. Figure 10 A, time point 2 or 3). As the initial testing was done with two turnip accessions and Pak Choi, this therefore did not optimally predict shared expression in turnips compared to non-tuberizing morphotypes (compare e.g. the results of gene 21 in the initial testing in Figure 8 B and the detailed testing in Figure 10 A). The assumption was that genes with roles in tuber formation have different, yet similar expression in non-turnips. The results from the screening over the larger set however show that this assumption often is wrong.

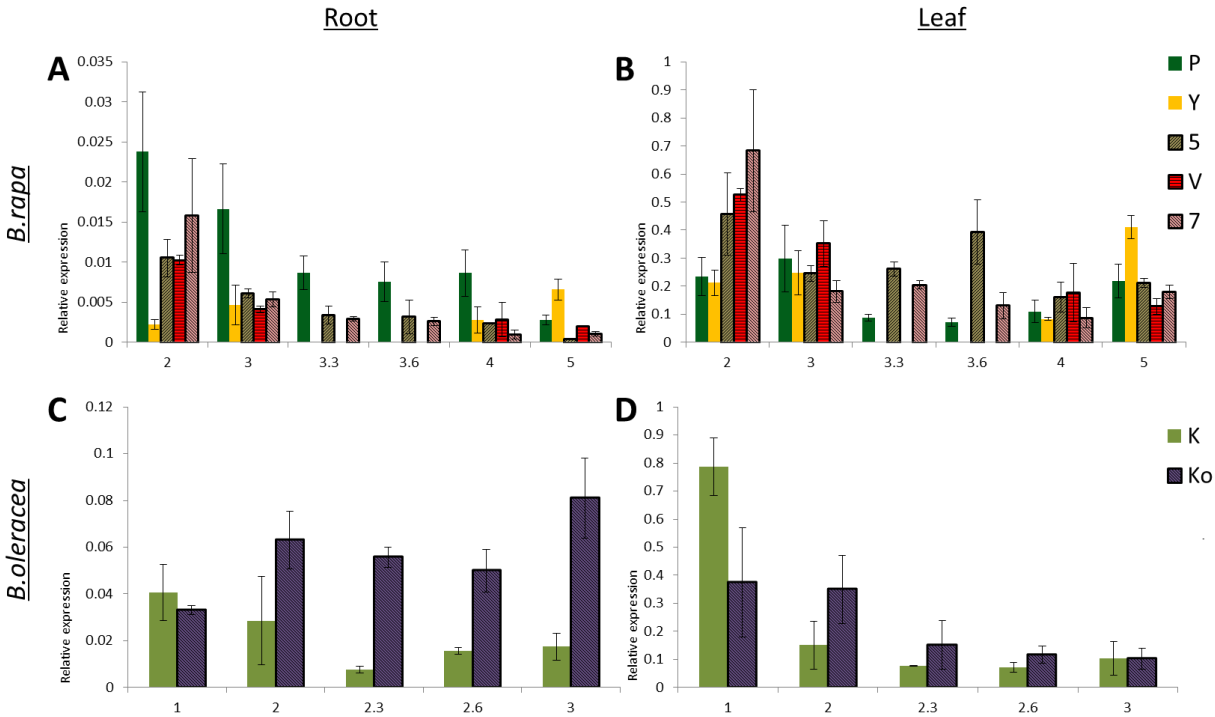

Figure 10: Expression pattern of a potentially not tuberization related gene (gene 21 Bra003743 - unknown)

On the y-axis the relative expression (= $2^{-\Delta Cq}$) of the target gene compared to the reference gene (Cyp Bra037296) is displayed whereas on the x-axis, time points are displayed in weeks. (A) and (B) show expression in *B. rapa*, whereas (C) and (D) show expression in *B. oleracea*. Expression in 'roots' is shown in (A) and (C) and expression in leaves is shown in (B) and (D). Abbreviations: P – Pak Choi, Y – Yellow Sarson, 5 – VT-053, V – VT-012, 7 – VT-117, K – Kale and Ko – Kohlrabi. Error bars represent standard deviation. Note differences in scale on the y-axis.

Nevertheless, for two out of the five genes (7 and 35), patterns specific for tuberizing turnips and differing from plants that do not form tubers were found.

The transcript level of gene 35 increased almost by one fold change for both VT-053 and VT-117 in time point 3.3 compared to time point 3 and remained high until the latest time point, whereas expression levels in other accessions remained relatively stable and only slowly increased until the latest time point (Figure 11 A). Regarding the fact that VT-012 did not form tubers and expression of gene 35 behaves similar to Yellow Sarson and Pak Choi over all time points tested, the expression of this gene is indicated to be correlated to initiation of tuber growth in *B. rapa*. Gene 35 was also one of the examples where the indicated patterns from the initial testing were validated in the detailed screening. This can be seen by comparing Figure 8 D with Figure 11 A.

A very similar expression pattern for *B. rapa* was observed for gene 7; in VT-053 and VT-117 this gene also increased its transcript abundance in 'roots' clearly at time point 3.3, whereas other accesions did not follow this behaviour (Figure 12 A). Similar to gene 35, a difference between transcript levels of VT-053 and VT-117 compared to VT-012 is already indicated in time point 3, although the genes are still expressed rather similar to each other in Pak Choi and Yellow Sarson. However, gene 7 in contrast to gene 35 does not increase in transcript abundance in the other morphotypes in time, leading to a much higher transcript level in the

two tuber forming turnips compared to the non-turnip forming accessions until the latest time point.

Expression patterns in leaves of *B. rapa* do not resemble the expression patterns in 'roots' for any of the genes nor can any obvious correlation to tuberization be observed (e.g. Figure 10 B and Figure 11 B). For gene 7, expression in leaves was even not detected at all.

Furthermore, in *B. oleracea* expression of the seven tested genes was also not indicated to be correlated to tuber formation nor did it resemble the differences between tuber forming and non-tuber forming accessions of *B. rapa* (see examples in Figure 10 C & D, Figure 11 C & D as well as Figure 12 B).

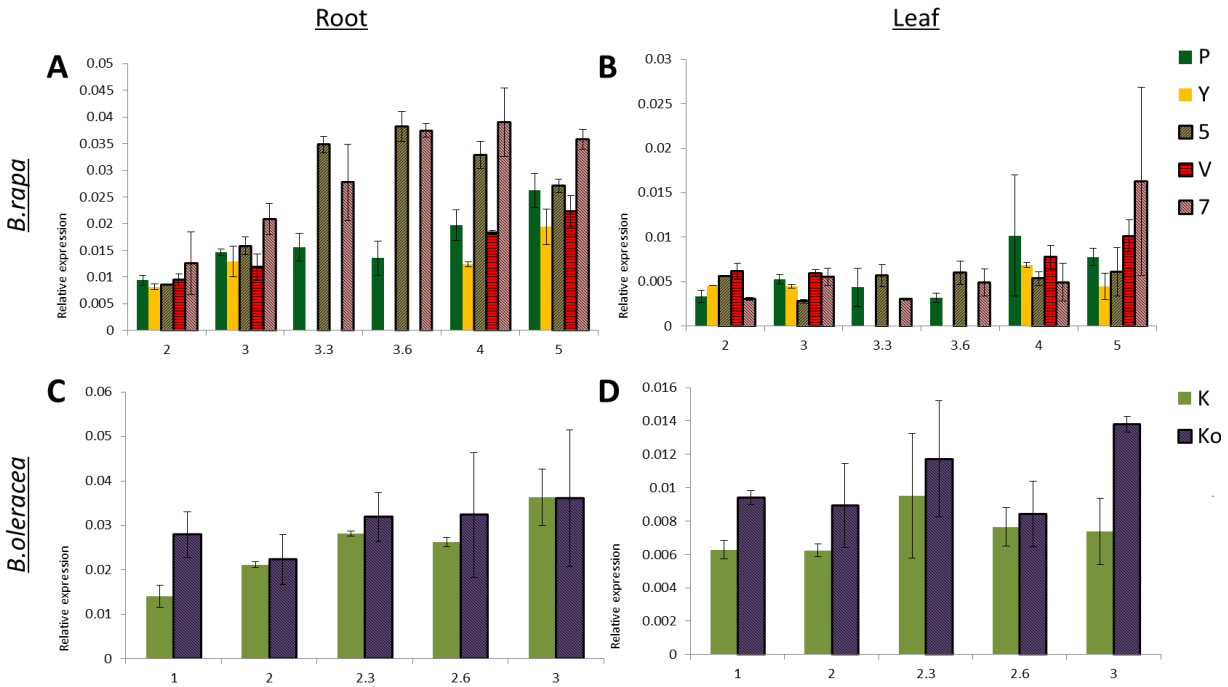

Figure 11: Expression pattern of potentially tuberization related gene 35 (Bra030232 – CYCD2;1)
On the y-axis the relative expression (=$2^{-\Delta Cq}$) of the target gene compared to the reference gene (Cyp Bra037296) is displayed whereas on the x-axis, the time points are displayed. (A) and (B) show expression in *B. rapa*, (C) and (D) show expression in *B. oleracea*. Expression in 'roots' is shown in (A) and (C), expression in leaves is shown in (B) and (D). Abbreviations: P – Pak Choi, Y – Yellow Sarson, 5 – VT-053, V – VT-012, 7 – VT-117, K – Kale and Ko – Kohlrabi. Error bars represent standard deviation. Note the difference in scales of the y-axis.

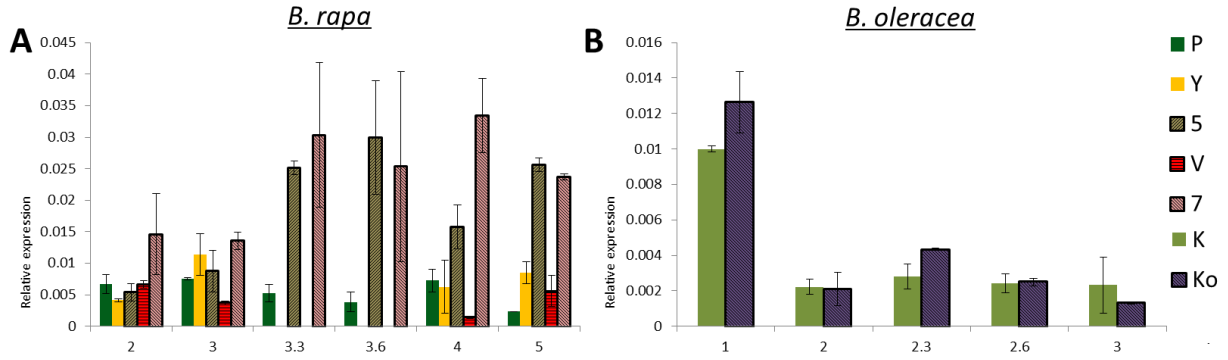

Figure 12: Expression pattern of potentially tuberization related gene 7 (Bra034022 – CYP735A2)
On the y-axis the relative expression (=$2^{-\Delta Cq}$) of the target gene compared to the reference gene (Cyp Bra037296) is displayed whereas on the x-axis, the time points are displayed. (A) shows expression in *B. rapa* 'roots' whereas (B) shows expression in *B. oleracea* 'roots'. Abbreviations: P – Pak Choi, Y – Yellow Sarson, 5 – VT-053, V – VT-012, 7 – VT-117, K – Kale and Ko – Kohlrabi. Error bars represent standard deviation. Note the difference in scales of the y-axis.

Since genes with promising expression patterns in 'root' tissues did not show any indications in their expression in leaves and in *B. oleracea* accessions, these latter samples were no longer considered in the analyses. Furthermore due to the limited prediction quality of the initial testing, selection of genes for further screening was done without considering results of the initial screening but only considering putative function of genes in tuberization. This led to the selection of 22 additional genes out of the list of 58, which were screened in all previously mentioned *B. rapa* 'root' samples. Eight of those genes had expression patterns indicative for a role in tuberization (for details see Table 3).

Table 3: Details for 8 promising genes after detailed screening
The first column (No.) indicates the number with which the gene is identified in the text of this work. The second column shows the official *B. rapa* gene ID, whereas the third column (*A. thal.* orth.) shows the *A. thaliana* orthologue which (predicted) function is displayed in the fourth column. The fifth and the last column indicate, whether the gene is placed in a Selective Sweep (Sel. Sw.) and/or a QTL – A number indicates the number of the Selective Sweep or QTL, (Y) indicates an overlapping QTL and (N) indicates non-presence.

No.	Gene ID	*A. thal.* orth.	predicted function	Sel. Sw.	QTL
3	Bra038700	AT3G12145	FLOR 1 – leucine rich repeat	6	Y
8	Bra034021	AT1G67110	CYP735A2 – cytokinin trans-hydroxylase	9	Y
17	Bra019808	AT1G13110	CYP71B7 – cytochrome P450 71B7	17	N
20	Bra003665	AT1G77850	ARF17 – Auxin Response Factor 17	20	N
34	Bra022954	AT2G33810	SPL3 – Squamosa Promoter binding protein-like 3	N	2
47	Bra014208	AT1G49750	leucine-rich repeat like protein	N	8
51	Bra039855	AT4G14550	IAA14 – auxin-responsive protein IAA14	N	8
57	Bra009091	AT5G05160	RUL1 – Reduced in Lateral Growth 1	N	13

The expression patterns of the eight promising genes are shown in Figure 13. For proper interpretation of the data it is however necessary to mention that VT-117 generally shows a very large standard deviation in time point two which makes reliable comparisons for this time point difficult.
The generally very highly expressed gene 3 is on average much higher expressed in turnips, although expression in VT-053 during time point 2 and 5 and in VT-117 during time point 3 is close to expression levels of Pak Choi and Yellow Sarson (Figure 13 A). Nevertheless this gene is often at least one fold higher expressed in turnips and sometimes even reaches transcript levels comparable to the reference gene.
Gene 8, which shares the same *Arabidopsis* orthologue as gene 7 and also is located in tandem with this gene in the *B. rapa* genome, has almost the same expression pattern as gene 7, although it is generally approximately three fold lower expressed across all samples (Figure 12 A and Figure 13 B).
An expression pattern which is similar in the two turnips VT-117 and VT-012 but different in the turnip VT-053 and the two non-turnips was observed for gene 17 (Figure 13C). For all time points except time point two, VT-117 has the highest transcript levels of this gene, while VT-012 also has higher transcript levels than the other three accessions (VT-053, Pak Choi and Yellow Sarson). This indicates a shared expression pattern for Japanese turnips, where abundance of the transcript increases slowly with the time.
Gene 20 is an example where differences of expression levels between turnips and non-turnips become clear with the onset of tuberization. Already at timepoint 3.3 transcript

abundance of this gene is slightly elevated in turnips and continues to be higher until the latest time point (Figure 13 D). Differences are especially clear in time point 4, where expression in turnips is up to two folds higher than in non turnips. However, in contrast to gene 7, 8 and 35, this gene also shows increased expression in VT-012 which did not form a tuber.

Another gene which expression is shared by all turnips in the time point that preceded tuberization is gene 34. Although not or almost not expressed in Pak Choi and Yellow Sarson, at all time points and for all *B. rapa* turnip accessions in time point 2, gene 34 is clearly expressed in all three turnip accessions at time point three (Figure 13 E). However, with the time, transcript levels of this gene go down for VT-117 until it is not expressed anymore in time point 5, whereas the gene remains expressed in the other two turnips. As this gene is also in a selective sweep of *B. oleracea*, expression of this gene was also tested in 'root' as well as leaf samples of this species. In both, 'root' and leaves, the gene was expressed from time point 2 on (Figure 14). In 'roots' the expression levels increase in Kale, whereas in Kohlrabi the transcript levels remain more or less stable (Figure 14 A). For leaves, expression levels of the gene increase in both morphotypes, but in Kohlrabi expression is always higher (Figure 14 B).

Coming back to the results of the initial testing, gene 34 is an example of a gene, which was clearly not indicated as promising after the initial testing (see Figure 7 C) but became a promising candidate for a tuberization related gene during the detailed testing.

Similar to gene 8, gene 47 indicates a shared expression pattern for Japanese turnips as it is slightly to clearly higher expressed in VT-117 and VT-012 in all time points except from time point 5 (Figure 13 F). The differences are however not that clear in time points 3.3 as well as time point 3.6, where also samples of VT-012 are missing.

Gene 51 shows an expression pattern which distinguishes turnips from non-turnips at a subset of time points. At time point 2, this gene is approximately one fold lower expressed in all turnips compared to the two non- turnips (Figure 13 G). Those differences are not that clear anymore in the next three time points. However, in time points 4 and 5 transcript levels are clearly higher in VT-117 and VT-053 again. The turnip VT-012 which did not form a tuber has similar low expression levels of this gene like the two non-turnip morphotypes. Therefore, gene 51 distinguishes turnips from non-turnips in early time points and changes it's expression levels during development. Whereas it's expression decreases in non-turnips, it increases in turnips during tuber formation but stays rather stable in turnip morphotypes that failed to form tubers.

A potential role in tuberization based on transcript levels can also be assumed for gene 57 which shows slightly elevated transcript levels in time point 3.6 (Figure 13 H). Due to the large standard deviation of the expression values in Pak Choi at time point 3.3 it is hard to evaluate, whether transcript levels are also higher in turnips at time point 3.3 and 4. Based on expression values and the assumption of similar expression levels of this gene in Pak Choi and Yellow Sarson, the results for this gene indicate an elevated expression during the onset of tuberization. Differences however clearly diminish at time point 5.

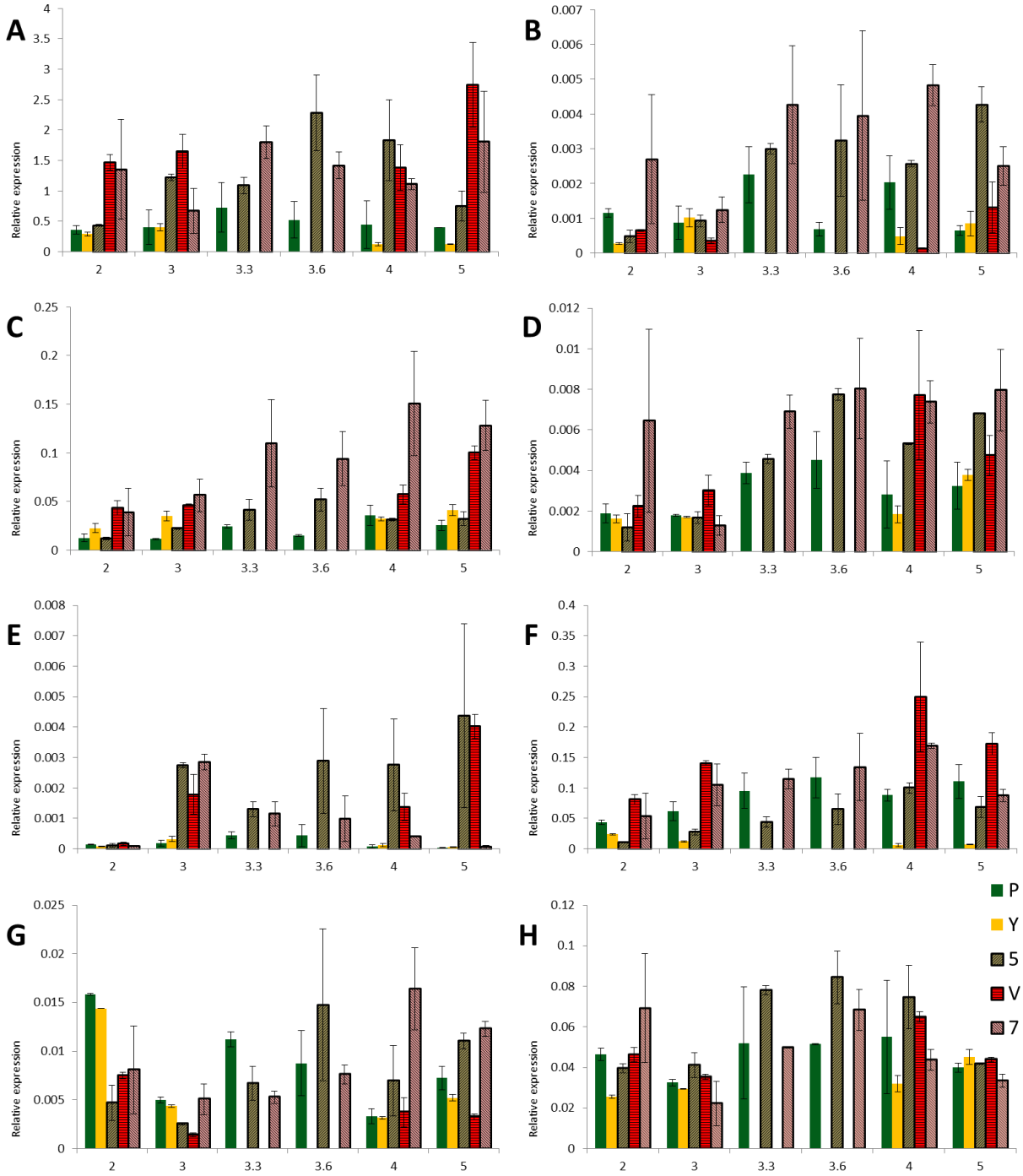

Figure 13: Expression pattern of eight potentially tuberization related genes
The y-axis represents the expression of the target gene relative to the reference gene Cyp - Bra037296 ($=2^{-\Delta Cq}$), whereas the x-axis represents the time points in weeks after germination. The following genes are displayed: (A) gene 3 (Bra038700 – FLOR1), (B) gene 8 (Bra034021 – CYP735A2), (C) gene 17 (Bra019808 – CYP71B7), (D) gene 20 (Bra003665 – ARF17), (E) gene 34 (Bra022954 – SPL3), (F) gene 47 (Bra014208 – leucine-rich repeat-like protein), (G) gene 51 (Bra039855 – IAA14), (H) gene 57 (Bra009091 – RUL1). Abbreviations: P – Pak Choi, Y – Yellow Sarson, 5 – VT-053, V – VT-012 and 7 – VT-117. Error bars represent standard deviation. Note the difference in scales of the y-axis.

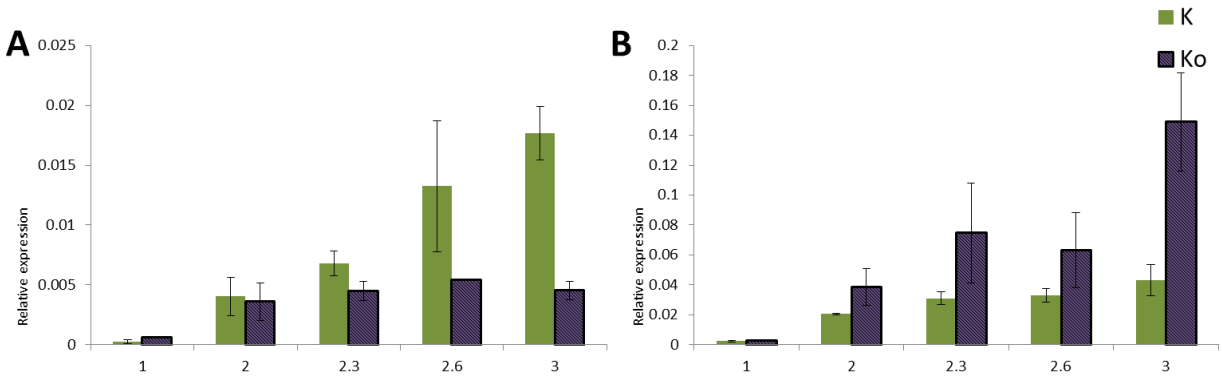

Figure 14: Expression levels of Gene 34 (SPL3) in *B. oleracea*
The y-axis represents the expression of the target gene relative to the reference gene Cyp - Bra037296 ($=2^{-\Delta Cq}$), whereas the x-axis represents the time points. (A) displays expression levels in 'roots', whereas (B) displays expression levels in leaves. Abbreviations: K – Kale and Ko – Kohlrabi. Error bars represent standard deviation. Note the difference in scales of the y-axis.

3.4 Anatomy

Fixed material lost colour during the ethanol series and appeared whitish to yellowish after some time. After long storage in ethanol RNA as well as DNA was still present in the samples (compare Figure 15 B (for DNA) & D (for RNA) with Figure 15 A & C). This was tested with an Acridine Orange staining. RNA as well as DNA was equally intact over the whole area of the sample (see Figure 15 B and D), so fixative penetrated the samples sufficiently. Higher signal intensity was observed around the vessels, both for RNA and DNA.

As the technique was not tested with these kind of samples yet only one biological repeat per time point was used for embedding to avoid loss of many fixed material due to potential problems. For embedding, Pak Choi and VT-117 were chosen as accessions with the time points 2, 3, 3.3, 3.6, 4 and 5.
During paraplast infiltration, two samples (VT-117, time point 3.6 as well a 4 – see Figure 17 D) heavily shrinked after the first paraplast exchange, whereas other samples remained their shape and stayed intact until embedding.
Sectioning of samples in cold paraplast blocks resulted in good and intact slices until 20 µm thickness. Thinner slices tended to rip heavily and the structure of the embedded tissue was disturbed by sectioning (compare Figure 16 A and B). Therefore all slices were cut to 20 µm thickness, although the quality of light microscopy pictures is suboptimal due to the many cell layers.
Even though the infiltrated wax prevented proper and equal penetration of acridine orange, the cut surface was still possible to stain to some extent. The staining indicated presence of both RNA and DNA (compare Figure 15 F (for DNA) & G (for RNA) with Figure 15 E & F), allowing the assumption that both molecules are still intact after embedding. The fluorescent images of the embedded non-stained samples, however are different as the ones of the unembedded tissue, making clear comparisons of RNA & DNA quality of embedded and unembedded tissue impossible (e.g. Figure 15 A compared to Figure 15 E).

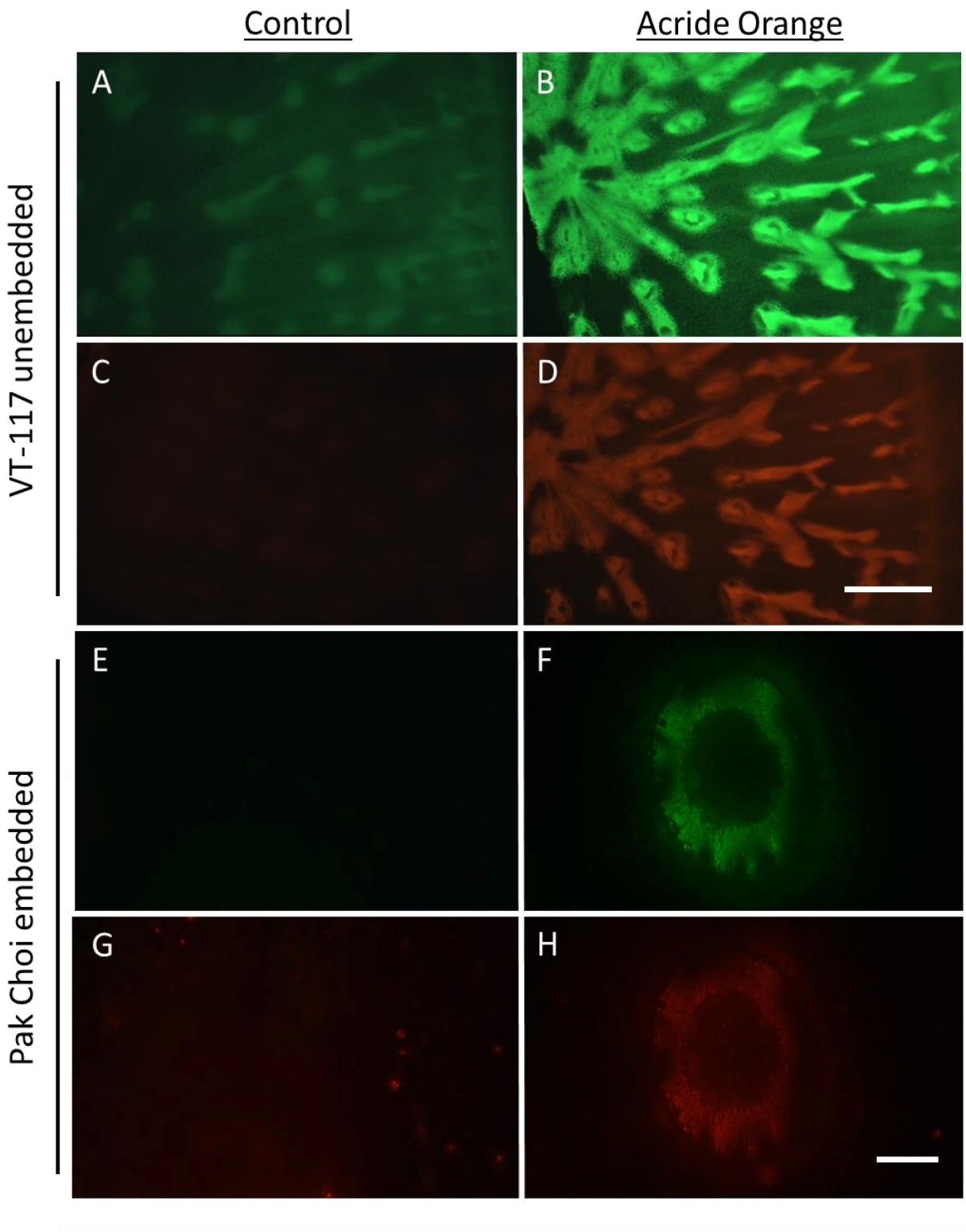

Figure 15: Stability of RNA and DNA in fixed or embedded 'root' tissue (time point 5)
The tissue samples are tubers of VT-117 taken out of 70 % Ethanol after long storage of several weeks (A - D) or are paraplast embedded 'roots' of Pak Choi (E - H). All samples were stained with acridine orange. (A), (C), (E) and (G) show unstained tissue, whereas (B), (D), (F) and (H) show the stained material. Pictures were taken under UV-light with a GFP filter for (A), (B), (E) and (F) and a ds-red filter for (C), (D), (G) and (H). The bar in (D) corresponds to 0.5 cm for (A – D) whereas the bar in (H) correspond to 2mm for (E – H).

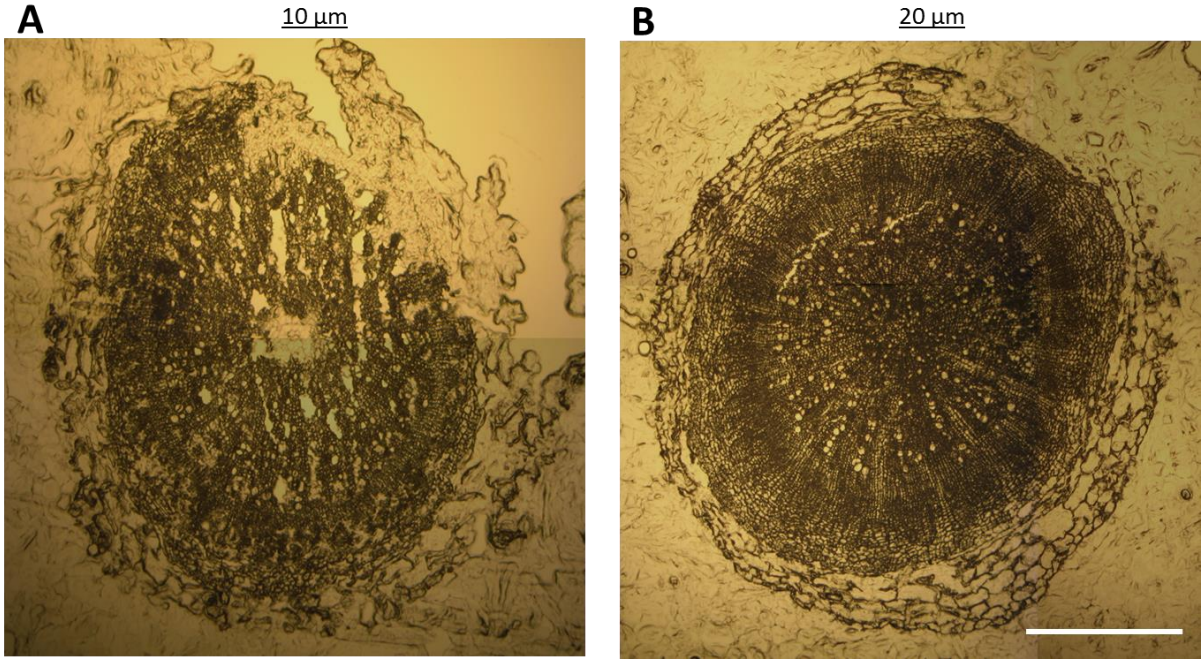

Figure 16: Effect of slice thickness on quality of sections
Both (A) and (B) are Pak Choi from the time point 3. (A) is a 10 µm thick section, whereas (B) is 20 µm thick. Both slices were cut from cold blocks. Picture (A) is aligned from two separate pictures, whereas (B) is aligned from four separate pictures. The Scale bar corresponds to 1 mm.

Generally embedded samples are well maintained overall with well preserved cellular morphology (see Figure 17). The shrunken tissue however showed strongly malformed sections with shrunken cells, which are often not clearly distinguishable anymore (Figure 17 D). Furthermore, the overall morphology does not resemble the morphology of non-shrunken tissue (compare e.g. to Figure 17 A to C).
Comparison of turnips (VT-117) to corresponding tissues in Pak Choi revealed clear differences in morphology. Sections from the hypocotyl of Pak Choi show clear distinction between pith and cortex with dense vascular bundles in between. (Figure 17 A – right). In turnip 'roots', this contrast is much less clear with more dense cells in cortex and pith but less dense cells in vascular bundles (Figure 17 A – left). The vascular bundles are furthermore much broader, which indicates initiated secondary growth. In the transition zone from hypocotyl to root in Pak Choi, cortex and pith tend to contribute a smaller proportion of the total section, whereas vascular bundles are still dense and clearly visible (Figure 17 B – right). In the corresponding section of turnip, overall morphology is very similar to the hypocotyl section, although different tissues become less clearly distinguishable. (Figure 17 B – left). The root part of Pak Choi shows adventitious root formation and a thick vascular cylinder with a rather thin cortex (Figure 17 C – right). Corresponding turnip sections are still rather comparable to hypocotyl, although the pith contributes a much smaller proportion (Figure 17 C – left).
Longitudinal sections of a turnip (time point 5) show petioles as well as leaf primordia close to the apical meristem characterized by small, dense and round cells (Figure 17 E). Longitudinal vessels connecting leaves and turnip are also visible in the periphery of the turnip.

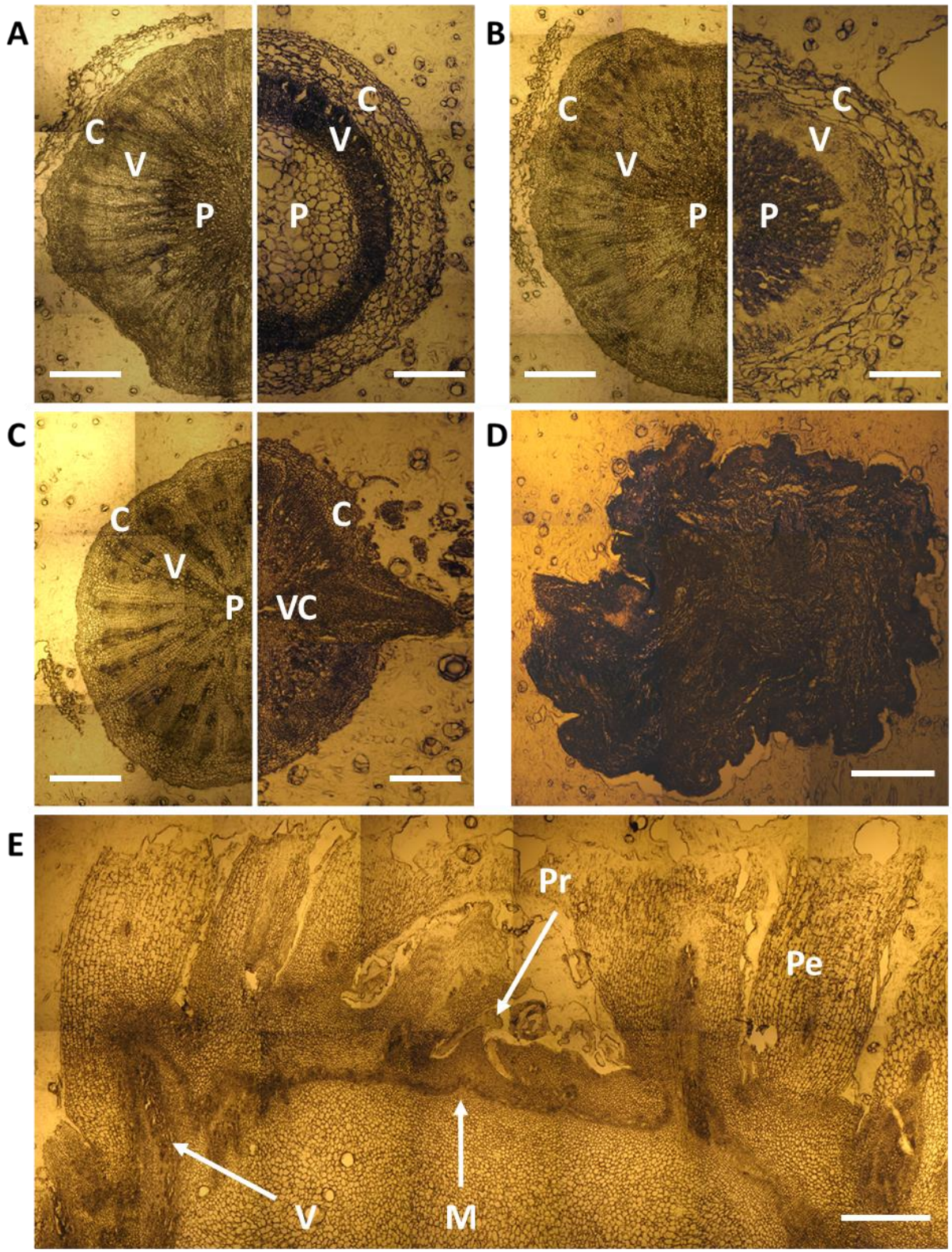

Figure 17: Morphology of embedded and sectioned Turnip and Pak Choi Tissues
Comparison of Turnip and Pak Choi time point 3 is shown in (A), (B) and (C), whereas VT-117 is displayed left and Pak Choi right. (A) shows upper part of the hypocotyl, (B) shows the part below hypocotyl and above root, (C) shows the root parts. Picture (D) is an example of a section of the turnips which shrunk (VT-117, time point 4). (E) shows the morphology of the apical part of a turnip (VT-117, time point 5) which was cut longitudinal. All sections are 20 µm thick. Pictures are aligned from several separate microscopy pictures. Scale bars: 0.5 mm for all Pak Choi samples and VT-117 in (C), 0.75 mm for VT-117 in (A) and (B), 1.5 mm for (D) and 1 mm for (E). Initials: P – Pith, V – Vascular bundles, C – Cortex, VC – Vascular Cylinder, M – Meristem, Pr – Leaf Primordium, Pe – Leaf Petiole.

4 Discussion

4.1 Datasets & Initial Selection of candidate genes

The investigation of datasets revealed a number of genes with possible involvement in tuberization but also revealed the small overlap of QTLs and selective Sweeps. Furthermore, the rather low percentage of genes which remained promising candidates after detailed testing revealed a non-optimal selection method.

4.1.1 Limited overlap between QTLs and Selective Sweeps

After comparing positions of Selective Sweeps and Positions of QTLs for turnip formation, it became clear that overlap was rather limited. However, this small overlap does not necessarily indicate that both datasets are not correct or not giving valuable indications.

Selective Sweeps for instance can be all genetic loci linked to genes with different functions than tuber initiation, for example some Selective Sweeps may surround a gene, which causes strong tuber enlargement rather than tuber induction. This gene could for instance be responsible for letting tubers grow larger, which is a desired characteristic for farmers and may also lead to fixation in the population and therefore to a selective sweep. Thus, an expansin could for instance be the candidate gene for such a selective sweep (Jung et al., 2010), but may not be detected in a QTL study, where effects on tuber formation rather than effects on tuber size are scored. In a QTL study however it would be rather hard to score exclusively for turnip absence or presence without taking into account turnip width, which the population also segregates for due to the many genes involved.

Due to the more diverse set of accessions used for detection of Selective Sweep, it is also rather difficult to detect Selective Sweeps for traits which are very quantitative and can have different underlying genes with same or similar effects. If turnips have been domesticated independently several times, which is at least the case for Japanese and European turnips (Zhao et al., 2005), different genes responsible for tuberization may have become fixed in different accessions. In this case, genes that are not shared between these different groups can be missed if a large and very diverse set of turnips is used for studies, whereas other genes which are shared may result in selective sweeps. Those shared genes may include genes not responsible for initiation of tuberization but for other desirable characteristics like strong tuber expansion, taste or sugar composition.

Furthermore detection of selective sweeps strongly depends on used parameters as well as window and stepping size. Too small windows for instance are leading to more false positives, whereas too large windows will lead to more false negatives.

Therefore, some Selective Sweeps can be false positives.

In contrast, QTL studies are limited to the alleles present in the two parents and therefore turnip-QTLs are only absolutely representative for those parents and may differ for other parental combinations. In both of the two QTL studies used for this work, the Japanese turnip VT-115 was used as one parent and therefore some QTLs may be characteristic for Japanese turnips. Hence for detecting Selective Sweeps a larger set of turnips and non-turnips is

compared, detection of relevant genes is not limited only to the alleles present in the parents. Considering other QTLs retrieved from different mapping populations than only the ones used in this work may therefore already lead to a larger overlap between QTLs and Selective Sweeps.
Furthermore for an optimal QTL study it is necessary that offspring segregates in absence of any skewness. This requires that alleles segregate 1 to 1. If however gametes, zygotes or offsprings differ in fitness, there will be skewed segregation. This has been the case for the two QTL studies used in this work. Thus, relevant genomic fragments might be absent in the final mapping population or detection of important QTLs may fail due to severe skewness as this can reduce significance of QTLs.

4.1.2 Quality of selection methods

Although several promising genes were detected, the possibility that the investigation missed important tuberization related genes cannot be excluded due to various reasons.

First of all due to time and other limitations, not all available data on genomic regions potentially responsible for tuber formation have been used in this work. For instance only turnip Selective Sweeps detected by Cheng Feng via piHS were used. Selective Sweeps detected by other signals or in the dataset of Lin Ke were not considered in the investigations while they did not completely overlap with the Selective Sweeps used for this work. Furthermore due to difficulties of detailed localization, turnip QTLs reported in publications (Kubo et al., 2010; Lou et al., 2007; Lu et al., 2008)were also not used in this work, while they were indicated to not always overlap with QTLs detected by Ningwen Zhang. Additionally often not the complete QTLs detected by Ningwen Zhang were investigated, as in case of overlap with Selective Sweeps, only the Selective Sweeps were investigated.
Therefore, some potential genomic regions harbouring tuberization related genes were not investigated in this work.

Second, although several regions were not considered, the genomic area covered by the investigated QTLs and Selective Sweeps still harboured about 15 % of all *B. rapa* genes. As only few genes were selected as candidates for each Selective Sweep and QTL, some causal genes can have been missed. Selection of genes was not only based on predicted function but initially exclusively based on differences in expression. This selection via gene expression can miss important genes, as some genes have a very large effect on the tuber trait with only very small expression changes. Similarly some genes situated in a QTL or a Selective Sweep may have a large difference in expression level but do not have an effect on the investigated trait. Furthermore important expression patterns can be missed due to very narrow peaks with very defined expression patterns between the investigated time points.

Third, genes which are causing tuberization and therefore result in a Selective Sweep or QTL, may not be expressed differently but may contain functional differences caused by mutation. Therefore, the genes which are involved in initial stages of tuber formation may not be selected by the methods which were used in this work as their effect may not be caused by

differences in expression. In these cases, the initial genes will definitely not be detected and only downstream genes can be detected with the used methodology (but could be detected via sequence comparison). The initial genes are the genes which are actually regulating the tuberization by acting upstream in the tuberization mechanisms. As turnips are the only *B. rapa* morphotypes that form tubers, those initial genes should differ between turnips and non-tuberizing morphotypes. Therefore, the initial/upstream genes will be placed in turnip Selective Sweeps or QTLs. Downstream genes which are regulated by the initial genes may also differ in their expression between turnips and non-tuberizing morphotypes. However this differences only result from differences in the initial genes or their expression. Thus, downstream genes are not causal for Selective Sweeps or QTLs. Nevertheless, as the investigated QTLs and Selective Sweeps are covering 15 % of the genome, it is likely that some downstream genes are placed in those regions by chance. The expression of upstream genes may not differ between turnips and non-tuberizing morphotypes or may differ less than those of the downstream genes. Therefore, some downstream genes might be detected as candidate genes, even though they are regulated by the upstream genes.

Therefore, with this methodology, some key genes involved in tuberization should have been detected. Those genes may not always be initial for tuberization but likely still indicate which genetics or molecular pathways are involved in tuberization and therefore provide hints for further investigation.

4.2 Experiments and Methodology

Several problems or challenges like no turnip formation for VT-012 or the predictive quality of the initial testing were faced during the experiments. This often resulted in the necessity of adapting the used methods. The reasons for those problems were various. Furthermore the fixation and embedding techniques require evaluation of their results for further experiments and follow up research.

4.2.1 Growth of VT-012

The Japanese turnip DH-line VT-012 did not form any tubers. Furthermore it was less uniform than the other accessions as it was varying for plant vigor and trichome number.
Morphological non-uniformity can be caused by many non-genetic factors such as environmental stress etc. Trichomes however are under genetic control (Agren and Schemske, 1994) and therefore indicate that the seeds for VT-012 are actually no pure DH-line seeds. As VT-012 still behaved rather uniform regarding tuberization (no plant formed tubers), it has been reasonable to still include in the experiments.
Failed tuberization for VT-012 might be explained by a too small pot, as this accession formed turnips in another experiment which took place at the same time with larger pots (Villegas Paliz, personal communication). Pot size is correlated to availability of nutrients and for turnips as well as other tuberizing crops, it is reported that nutrients, especially nitrogen, are important regulators for tuberization (Aksenova et al., 2012; Simonne et al., 1993). VT-012 may require more nutrients or may be less tolerant to abiotic stress than the other two

turnip accessions and therefore did not form tuber under identical conditions. It is however not clear at what point in the mechanism of tuber initiation VT-012 started to differ from the other two turnip accessions that did form tubers at the expected time. The expression patterns of genes 34 and 35 for instance indicate that some genes are expressed similar in VT-012 compared to like in other turnips, whereas some genes are expressed more similar to non-tuberizing morphotypes. Therefore, it is likely that some upstream tuberization related genes or pathways are activated but some other essential tuberization related pathways or downstream genes were not activated in this experiment for VT-012. VT-012 can therefore still be included in the set of accessions, as it can serve either as third turnip accession or as a negative control.

4.2.2 Prediction Quality of Initial Testing

As previously mentioned, only two out of seven genes, which were selected based on their expression pattern over three genotypes at a single time point, indeed showed promising expression patterns when screened over more time points and genotypes. Therefore, the prediction quality of the initial testing is rather low, which may have different reasons. One explanation is that without biological repeats, expression values are not very reliable, which is also due to the large standard deviation of the biological repeats that often exceeds differences in expression between morphotypes. Most likely these large values for the standard deviation resulted from sampling of plants in different developmental stages. Another explanation is that Yellow Sarson is not included in the sample set used for initial testing. In the detailed screening Yellow Sarson however often clearly had different gene expression levels than Pak Choi, which resulted in many false positives in the initial testing. Those false positives were revealed when expression of genes in all turnip accessions was compared to all non-turnip accessions. This should however not lead to the conclusion that gene expression in Yellow Sarson is more similar to that of turnips than gene expression in Pak Choi.

Furthermore, the time points selected for initial testing were often not optimal. This is due to differences of expression in plants used for this work which were grown in August 2015 and the plants used for the microarray experiment which were grown in September 2014. Gene 34 for instance displayed increased expression until week 6 in VT-117 based on the microarray data. In the plants of the same accession which were used for this work, this gene was however almost not expressed at week 6 anymore.

The initial testing nevertheless was a good test of primer quality with a rather diverse set of samples (turnip and non-turnips as well as leaves and ‘roots’). Primers that annealed specifically in this set of samples, also annealed specifically in the whole set of samples. Reasons for non-specific annealing of primers in some samples may be differences of genomic sequences in different accessions compared to the reference genome sequence or alternative splicing, which is very common in Arabidopsis but also *B. rapa* (Filichkin et al., 2010; Tong et al., 2013; Yuan et al., 2009).

4.2.3 Fixation and Embedding of Samples

Although samples for fixation and embedding are usually recommended to be very small (Jackson, 1992), fixation of large tissue samples in this work was sufficient, indicating that vacuum infiltration sufficiently enabled penetration of fixative. For the samples used in this work it is therefore not necessary to dissect tissue in smaller pieces and fixation of even larger samples may also be possible. This could in some cases enable a better illustration of overall morphology.

However, some samples shrunk during the paraplast infiltration. Several possible explanations are described for this behaviour, such as inappropriate fixation, too long exposure to xylene, incomplete removal of ethanol or xylene as well as overheating (Drobysheva et al., 2008; Ross, 1953). As only two samples shrunk and no biological repeats were treated in the same way, it is however hard to definitely determine the reasons for the observed shrinkage. Nevertheless, due to the tested quality of RNA in other samples, it is unlikely that inappropriate fixation is the cause of the observed shrinkage.

Infiltration of paraplast also has been sufficient, as insufficiently infiltrated samples usually show compressed tissue due to sectioning. Damage of sections and samples during cutting of slices thinner than 20 µm therefore likely result from other reasons. The observed ripping of sections is likely to be caused by dull blades or small particles adhered to the blade or placed in the paraplast. The main reason however for the observed problems with thinner sections is likely the differences of hardness of samples and the surrounding paraplast. For optimal sections, samples should be similar hard as embedding media. Due to strong lignification in stem and upper root parts as well as hardening of samples during long storage in 70 % ethanol and long keeping in xylene or liquid paraplast (Chung et al., 2008), most samples therefore are likely to be too hard for optimal paraplast embedding and sectioning. Embedding in Technovit®, which is much harder may therefore be a better option to obtain thin slices, which allows better and more detailed investigation.

Even though sections of 20 µm thickness are not optimal for microscopy due to the presence of too many cell layers, those sections should still allow conclusions about the general location of gene expression and transcripts.

The purpose of fixation and embedding was the preparation of suitable samples for in-situ hybridization. As overall as well as cellular morphology is retained after sectioning and RNA quality if good, those sections are likely to give good results when used for in situ hybridization. Furthermore, Technovit embedding may even result in more suitable sections as they likely can be cut thinner.

4.3 Comparison of candidate genes of *B. rapa* to *B. oleracea*

Comparing the data on tuberization of the two *Brassica* species, only little overlap was found as gene expression patterns for the tested genes in *B. rapa* and *B. oleracea* were not similar.

Only five out of 37 Selective Sweeps in *B. oleracea* are potentially in syntenic regions of some Selective Sweeps in *B. rapa*. This conclusion was drawn from the high proportion of similar orthologues in a similar or same order in those Selective Sweeps. Considering this small proportion, this overlap can also be interpreted as random overlap. However, selection

of one orthologue can lead to similar effects like selection of another orthologue in another subgenome as similar genes can have the same or similar functions. Therefore, underlying genes for tuberization can still be the same even though selective sweeps may not overlap. Regarding the fact that almost 20 % of the 60 candidate genes in *B. rapa* also had orthologues in *B. oleracea* Selective Sweeps, there is indeed to some extent indication for selection of similar genes during domestication.

In cases when expression patterns were determined for genes in *B. rapa* as well as their predicted orthologues in *B. oleracea*, patterns were not comparable. In some cases like in gene 34, expression levels distinguishing tuberizing and non-tuberizing morphotypes even had opposite patterns in *B. oleracea* compared to *B. rapa*. This observation can lead to the conclusion that in *B. oleracea* different genes and mechanisms are involved in tuberization than in *B. rapa*. There are however several points which prevent this clear conclusion.

Some of those points relate to the differences in sampling and morphology of *B. oleracea*. The first point is that *B. oleracea* forms tubers on the stem, above hypocotyls. This tissue is not present in the time points before approximately week five. Therefore, expression in *B. rapa* tissue which forms tubers later on is compared to *B. oleracea* tissue which will never form tubers.

Due to this difference in growth and morphology, harvesting of the whole stems including regions above cotyledons as well as a small proportion of root was necessary to maintain comparability of *B. oleracea* samples. This sampling however lead to the second point which prevented a clear conclusion. The tubers in *B. oleracea* present a much smaller proportion of the complete sample than in *B. rapa* and hence putative expression differences in tuberous organs are 'diluted' in *B. oleracea*.

Another reason why it cannot be concluded definitely that different genes are involved in tuberization of *B. oleracea* and *B. rapa* is the fact that the expression of only the one *B. oleracea* orthologue which was most similar to the *B. rapa* candidate gene was determined. As genome triplication occurred before the divergence of *B. rapa* and *B. oleracea* (Liu et al., 2014; Wang et al., 2011), the most similar orthologues are also most likely in the same subgenome. However, even though mutation rate in the subgenomes are not equal (Cheng et al., 2012), still orthologues in different subgenomes can be selected for. In this case, comparing expression patterns of the most similar orthologues will lead to wrong conclusions. Nevertheless this approach was chosen due to time limitation. Furthermore, even if all three orthologues would have been investigated in *B. oleracea*, the set of just one kohlrabi and one non-tuberizing kale would not have allowed representative conclusions.

4.4 Potential roles of selected *B. rapa* genes in tuberization

Ten genes were identified which show expression patterns that are indicative of involvement in tuberization. The important candidate genes from previous studies *Br-FLC2* (Lou et al., 2007; Zhao et al., 2010) was however not tested, as the overlap of the QTL including this gene with a Selective Sweep indicated that this gene may not be involved in tuberization. However, *Br-FLC2* can only be removed from the list of candidates, if Selective Sweeps indeed are very reliable. Further experiments with FLC2 would therefore provide additional information. Additionally, no genes from the class III peroxidases, which are overrepresented

in the turnip genome (Lin et al., 2014), were found as promising candidate genes. This may be due to downstream activity of those genes in the tuberization process.

4.4.1 Paralogues with indications of different function

Gene 3 which is Bra038700 and an orthologue of the *A. thaliana* FLOR1 (At3g12145) is usually higher expressed in turnips. In *A. thaliana*, the leucine rich repeat FLOR1 is described to promote flowering by being involved in transition from vegetative meristem to floral meristem (Torti et al., 2012). In this function it very likely interacts with the MADS-domain transcription factor AGAMOUS which expression is also flower specific (Acevedo et al., 2004). The fact that it is generally very high expressed in *B. rapa* 'root' issue is interesting as in *A. thaliana* it is described to either be exclusively expressed in floral meristem and tissues (Gamboa et al., 2001) or to be only very weakly expressed in vegetative meristems (Torti et al., 2012). This is clearly not the case for *B. rapa* where it is high expressed in all the 'root' samples, which exclude vegetative shoot meristems. In turnips, however, expression of this gene is almost always even one to two fold higher and sometimes reaches expression levels of the reference gene CYP (Bra037296). Considering the fact that Bra038700 is the only one out of the three *B. rapa* paralogues that increases in expression in VT-117 from time point three on (microarray data), there might be a change in function for this paralogue in tuberization. This assumption is supported by the fact that this gene is located both in a Selective Sweep and a QTL (from the mapping population VT-115 x wutacai105). However, to increase evidence for a role of this paralogue in tuber formation, an investigation of the expression of the other two paralogues in the set of the five accessions. Furthermore, a comparison of sequences of gene as well as its regulatory regions may reveal differences between this paralogue and the other paralogues. This would give hints, whether indeed the coding region of this paralogue changes or whether differences in upstream regulatory elements are a more likely explanation.

Another gene which is indicated to have a novel function in turnip is gene 34 (Bra022954) which is an orthologue of *Arabidopsis* SPL3 (AT2G33810). In the tested 'root' samples, this gene remains not expressed or maximally very low expressed in the non-turnip morphotypes at all time points. In turnip 'roots' however it is clearly expressed from time point 3 on, although its expression later on behaves differently in the thee turnips. SPL3 is a member of the Squamosa binding-like (SPL) transcription factors that contain the SBP domain (Fornara and Coupland, 2009). In Arabidopsis this gene is very low expressed in juvenile plants and expression levels increase with age, especially short before and during flowering (Cardon et al., 1997). It was reported to be involved in floral transition but more likely in vegetative phase change (Cardon et al., 1997; Wu and Poethig, 2006). The fact that this gene is clearly expressed in young turnip accessions which are far from flowering is interesting and suggests a potential role in tuber induction – especially as expression levels tend to decrease again at least in VT-117 as soon as radial tuber growth starts. This behaviour is contrary to what is reported for *Arabidopsis*, where expression of this gene increases continuously with age of plants (Cardon et al., 1997).

SPL3 is regulated by miR156, a micro RNA which is only expressed in juvenile plants and negatively regulates SPL3 transcript abundance post-transcriptionally (Fornara and Coupland, 2009; Wu and Poethig, 2006). This micro RNA is an important regulator of the juvenile phase in plants and has been shown to play a role in the tuberization process in potato and even in induced tuberization in tomato where it is high expressed during tuberization (Eviatar-Ribak et al., 2013; Lakhotia et al., 2014). In turnip however, the elevated expression levels of SPL3 indicate decreased levels of miR156. Furthermore, a recent study in tuber development of radish, reported decreasing miR156 levels during tuber formation (Sun et al., 2015). Therefore, a decreasing level of miR156 during turnip development is indeed likely, even though SPL3 expression is indicated to not exclusively being regulated by miR156 but also by other elements (Fornara and Coupland, 2009; Wu and Poethig, 2006).
Although the role of SPL3 and miR156 in turnip formation cannot be determined based on the available data, the transcript abundance of SPL3 indicates that they are factors contributing to turnip formation. Analysis of miR156 levels may shed more light to this question. Furthermore, as SPL4 and SPL5 are suggested to have redundant roles to SPL3 and are also regulated by miR156 (Fornara and Coupland, 2009), analysis of mRNA levels of those genes might also be an interesting approach.

4.4.2 Gene with putative direct effect on meristem/cambium formation

The expression of gene 35 (Bra030232) which is an orthologue of Cyclin D2;1 in *Arabidopsis* (CYCD2;1 - AT2G22490) indicates involvement in tuberization in some *B. rapa* morphotypes. This gene is clearly higher expressed in the two turnips which did form tubers from time point 3.3 until time point 4. In VT-012, the expression level of CYCD2;1 remains however similar to that of non-tuberizing morphotypes.
CYCD2;1 is a member of the D-type cyclin family in plants which are G_1-specific cyclins (Boucheron et al., 2004; Soni et al., 1995). Studies in Arabidopsis revealed expression in all tissue types but higher expression in tissue with meristematic activity (Sanz et al., 2011). If higher expressed, CYCD2;1 results in increased cell division rates in meristems, resulting in enlarged root apical meristems and more but smaller cells in produced tissue (Qi and John, 2007; Sanz et al., 2011). Overexpression in tobacco furthermore resulted in enlarged shoot apical meristems with increased periclinal divisions in the L3-layer (Boucheron et al., 2004). Those results indicate a role of this gene in tuber growth of turnips when higher expressed.
CYCD2;1 expression is induced by high sucrose but not hormones (Riou-Khamlichi et al., 2000). Nevertheless, it plays a role in the formation of lateral roots together with auxin (Sanz et al., 2011). The induction of this gene by sucrose corresponds to the previously formulated idea that sugar metabolisms is a key regulator in turnip formation (e.g. Nishijima et al., 2005; Peterson, 1973). These results as well as the fact that expression in VT-012 does not follow the trend of the two tuber forming turnips, indicate that this gene may not act as an initial upstream regulator in tuber-induction.

Another gene which may directly contribute to secondary growth is gene 57 (Bra009091) which is an orthologue of the *Arabidopsis* RUL1 (AT5G05160). This gene is higher

expressed in time point 3.6 in the two turnips, which were tested at this time point. In other time points, expression is not clearly higher.
RUL1 is a recently described receptor like kinase belonging to the LRR-III subfamily (Agusti et al., 2011; Shiu et al., 2004). As a positive regulator, RUL1 acts as promoter of cambium growth in the shoot, whereas MOL1 acts as its negative regulator (Agusti et al., 2011). An increased expression therefore could explain increased radial growth of turnips. The differences of expression are however not clear enough to assume a significant function of RUL1 in tuberization of *B. rapa*. Therefore it is more likely to assume, that increased cambium activity is due to different stimuli.

4.4.3 Genes indicating hormonal differences between turnips and non-turnips

A function in tuberization via alternation of hormonal levels can be assumed for gene 7 (Bra034022) and 8 (Bra034021). Both genes are orthologues of *A. thaliana* CYP735A2 (At1g67110) and are located in tandem. From the week where tuber swelling starts, those two genes are both higher expressed in the tuber forming turnip accessions than in non-turnip morphotypes and the turnip accession which did not form tubers. In *Arabidopsis* CYP735A2 is mainly expressed in root vascular bundles (Kiba et al., 2013). Considering the fact that the proportion of root tissue compared to stem tissue at least in samples of VT-117 is smaller than in Pak Choi, those difference are likely to be even larger if only root tissue would have been compared. However, this only would be true if expression in *B. rapa* would also be located in roots as it was described for *Arabidopsis*.
In *Arabidopsis*, this gene is described to catalyse the biosynthesis of trans-zeatin, one of the most active cytokinins (Schmitz et al., 1972; Takei et al., 2004). A higher expression therefore would result in a higher level of trans-zeatin. As trans zeatin is an inhibitor of root growth but a positive regulator for shoot growth (Kiba et al., 2013; Ramireddy et al., 2014) and turnips are composed at least partially by stem tissue, increased trans zeatin levels can indeed explain turnip growth. Inductive effects of cytokinins on root cambium growth in radish were reported (Jang et al., 2015) suggesting further putative effects of cytokinins on growth of turnips. In contrast, small in vitro experiments with turnips had some indications of negative effects of cytokinin on turnip formation (Zhang, personal communication, Guan, 2009). In tomato however, it was shown that ectopic expression of an cytokin biosynthesis gene induced tuberization (Eviatar-Ribak et al., 2013).
Both genes are located in a very small Selective Sweep which is overlapping with a QTL. However, as VT-012 did share this expression patterns and hormonal regulation is rather complex, it is not clear whether this gene is indeed an upstream regulator of turnip formation.

Another gene which indicates a role of hormones in tuber formation is gene 20 (Bra003665) which is an orthologue of *Arabidopsis* Auxin Response Factor 17 (ARF17 - AT1G77850). This gene is higher expressed in turnip ‘roots’ from time point 3.3 on. ARF17 is a transcription regulator of several auxin response genes and is itself a target of the micro RNA miR160 (Gutierrez et al., 2009). This micro RNA is the key regulator in transcript abundance of ARF17, which is involved in a wide range of plant developmental processes such as adventitious root formation and light signalling via its involvement in early auxin response

(Gutierrez et al., 2009; Sorin, 2005). In radish, miR160 is also reported to be upregulated during tuber development (Sun et al. 2015).
There is furthermore evidence for involvement of ARF17 in many more developmental processes, as accumulation of ARF17 transcript led to severe changes of almost all organs and also developmental processes (Mallory et al., 2005). ARF17 furthermore regulates the expression of several other ARFs. Furthermore, miR160 also regulates ARF10 and ARF16 post-transcriptionally (Mallory et al., 2005).
As transcript abundance of ARF17 is increased in all turnips including VT-012, it is indicated that in turnips there are general upstream differences in regulation of auxin response in some time points. Those differences are however not absolutely clear, as differences in transcript abundance of ARF17 between turnips and non-turnips are not big. An investigation of the abundance of miR160 or the other two ARFs which are regulated by miR160 would therefore shed more light to the question, whether there are really differences in auxin response regulation.

The expression of Gene 51 (Bra039588) which is an orthologue of *Arabidopsis* IAA14 (AT4G14560) also indicates some changes in hormone level or hormone response in turnips. This gene is lower expressed in turnip morphotypes during early time point but increases in expression after tuber initiation has passed and tuber growth is in progress. In contrast to CYCD2;1, IAA14 is inhibiting periclinal cell divisions and blocks lateral root formation (Fukaki et al., 2005, 2002). IAA14 represses auxin response genes and therefore alters auxin signalling in the development of plants (Vanneste et al., 2005). Its role in tuberization may therefore enable periclinal cell divisions in early stages of tuberization but later on reduction of the division rate to enable priority to expansion of cells, which seems restrained with high activity of CYCD2;1 (Qi and John, 2007). However, based on the available information, this assumption is too speculative and further experiments would be needed to investigate the effect of IAA14 on tuberization.

4.4.4 Genes indicating differences between Japanese and European turnips

Expression of some candidate genes indicated different expression in Japanese turnips, compared to similar expression in European turnip and the non-tuberizing morphotypes.

One of these genes in gene 17, which is Bra019808. This gene is an orthologue of the *Arabidopsis* CYP71B7 (At1g13110), which is a cytochrome P450 from the CYP71 family. CYP71B7 has a reported function in stress response or terpenoid metabolism (Maughan et al., 1997; Narusaka et al., 2004). This gene is higher expressed in the Japanese turnips VT-117 and VT-012 compared to the European turnip VT-053, although differences are rather unclear due to the expression in VT-012 which is lower than in VT-117 but still to some extent higher than in the non-tuberizing morphotypes. However, due to its expression profile and the reported function, this gene is not further considered as a candidate gene for turnip formation.

A similar expression pattern was identified for gene 47 (Bra014208) which is an orthologue of an uncharacterized leucine rich repeat in *Arabidopsis* (At1g49750). Prediction of function

in tuberization for this gene is hardly possible due to the lack of available published information as well as the lack of characterization of genes with similar sequences.. Also for this gene, expression seems to be higher in the two Japanese turnips than in the other accessions. However, similar to the previously mentioned gene, the differences are not clear.

Although due to their independent domestication, differences between Japanese turnips and European turnips can be expected, in this work, no genes with likely functions in turnip formation were found which clearly distinguish Japanese turnips from European turnips based on their expression. Only two genes indicated expression differences between those two turnip groups which however were not clear.

5 Conclusion

The investigation of several genes in *B. rapa* turnip Selective Sweeps and QTLs lead to the identification of eight genes which are likely to be involved in tuberization at different developmental stages and with different effects.

Two of these genes - FLOR1 and SPL3 - are plant development related genes involved in phase transition. Specific paralogues of these genes are indicated to have a changed role in turnips being now involved in tuberization. They have a likely function upstream in tuber induction. However, further research would be necessary to investigate and validate these indications. Other genes like CYP735A2 and ARF17 point to roles of auxins and cytokinins in turnip tuberization. Even though it is not clear whether those genes act upstream in the induction of tuberization, it would be interesting to investigate internal hormonal levels in turnips to shed light on the hormonal effect in tuberization of *B. rapa.* The investigation also lead to the identification of potential downstream genes of which CyclinD2;1 is the most likely candidate. These genes may act in the process of tuberization enabling tuber growth but are rather unlikely to induce tuberization.

Even though two genes indicated differences between the mechanism of tuberization in Japanese and European turnips, these genes were not indicative enough to conclude that these differences indeed exist. Further investigation of this question with more turnip genotypes or different candidate genes might be more promising if understanding of tuberization in *B. rapa* is better.

Two of the identified genes (SPL3 and ARF17) are reported to be regulated by micro RNA. On the one hand this indicates an upstream role of those genes but on the other hand it opens new questions on the micro RNA levels in turnips. Therefore investigation of microRNA, especially miR156 and miR160 are promising approaches for future research to investigate whether changed micro RNA levels also have an effect on tuberization in *B. rapa*.

This work therefore adds some concepts to the available knowledge on the tuberization of turnips but cannot provide many details for the understanding of this process. The deeper investigation of those concepts however may provide some more details on tuberization in *B. rapa*.

The RNA in situ hybridization, for which this work provides basics, could be a first step of these investigations as it can provide further hints for the function of some identified genes and may also enable initial investigation of micro RNA levels.

Literature

Abelenda, J.A., Navarro, C., Prat, S., 2013. Flowering and tuberization: a tale of two nightshades. Trends Plant Sci.

Abelenda, J.A., Navarro, C., Prat, S., 2011. From the model to the crop: genes controlling tuber formation in potato. Curr. Opin. Biotechnol. 22, 287–292.

Acevedo, F.G., Gamboa, A., Paéz-Valencia, J., Jiménez-García, L.F., Izaguirre-Sierra, M., Alvarez-Buylla, E.R., 2004. FLOR1, a putative interaction partner of the floral homeotic protein AGAMOUS, is a plant-specific intracellular LRR. Plant Sci. 167, 225–231.

Agren, J., Schemske, D.W., 1994. Evolution of Trichome Number in a Naturalized Population of Brassica rapa. Am. Nat. 143, 1–13.

Agusti, J., Lichtenberger, R., Schwarz, M., Nehlin, L., Greb, T., 2011. Characterization of Transcriptome Remodeling during Cambium Formation Identifies MOL1 and RUL1 As Opposing Regulators of Secondary Growth. PLoS Genet. 7, 1 – 14.

Aksenova, N.P., Konstantinova, T.N., Golyanovskaya, S.A., Sergeeva, L.I., Romanov, G.A., 2012. Hormonal regulation of tuber formation in potato plants. Russ. J. Plant Physiol. 59, 451–466.

Basnet, R.K., Moreno-Pachon, N., Lin, K., Bucher, J., Visser, R.G., Maliepaard, C., Bonnema, G., 2013. Genome-wide analysis of coordinated transcript abundance during seed development in different Brassica rapa morphotypes. BMC Genomics 14, 840.

Bassetti, N., 2015. Transcriptomic profiling of turnip (Brassica rapa spp. rapa) during tuber development (Master thesis).

Bonnema, G., Del Carpio, D.P., Zhao, J., 2011. Diversity Analysis and Molecular Taxonomy of Brassica Vegetable Crops, in: Genetics, Genomics and Breeding of Vegetable Brassicas, Genetics, Genomics and Breeding of Crop Plants. CRC Press, Boca Raton, pp. 81–124.

Borlido, J., Pereira, S., Ferreira, R., Coelho, N., Duarte, P., Pissarra, J., 2002. Simple and fast in situ hybridization. Plant Mol. Biol. Report. 20, 219–229.

Boucheron, E., Healy, J.H.S., Bajon, C., Sauvanet, A., Rembur, J., Noin, M., Sekine, M., Riou Khamlichi, C., Murray, J.A.H., Van Onckelen, H., Chriqui, D., 2004. Ectopic expression of Arabidopsis CYCD2 and CYCD3 in tobacco has distinct effects on the structural organization of the shoot apical meristem. J. Exp. Bot. 56, 123–134.

Braissant, O., Wahli, W., 1998. A simplified in situ hybridization protocol using non-radioactively labeled probes to detect abundant and rare mRNAs on tissue sections. Biochemica 1, 16.

Burns, W., 1946. Corm and Bulb Formation in Plants, with Special Reference to the Gramineae. Trans. Bot. Soc. Edinb. 34, 316–347.

Cardon, G.H., Höhmann, S., Nettesheim, K., Saedler, H., Huijser, P., 1997. Functional analysis of the Arabidopsis thaliana SPB-box gene SPL3: a novel gene involved in the floral transition. Plant J. 12, 367–377.

Casson, S.A., Lindsey, K., 2006. The turnip Mutant of Arabidopsis Reveals That LEAFY COTYLEDON1 Expression Mediates the Effects of Auxin and Sugars to Promote Embryonic Cell Identity. PLANT Physiol. 142, 526–541.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A.P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., et al., 2014. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science 345, 950–953.

Chapin, F.S., Schulze, E.-D., Mooney, H.A., 1990. The Ecology and Economics of Storage in Plants. Annu. Rev. Ecol. Syst. 21, 423–447.

Cheng, F., Wu, J., Fang, L., Sun, S., Liu, B., Lin, K., Bonnema, G., Wang, X., 2012. Biased Gene Fractionation and Dominant Gene Expression among the Subgenomes of Brassica rapa. PLoS ONE 7, e36442.

Cheng, F., Wu, J., Wang, X., 2014. Genome triplication drove the diversification of Brassica plants. Hortic. Res. 1, 14024.

Chung, J.-Y., Braunschweig, T., Williams, R., Guerrero, N., Hoffmann, K.M., Kwon, M., Song, Y.K., Libutti, S.K., Hewitt, S.M., 2008. Factors in tissue handling and processing that impact RNA obtained from formalin-fixed, paraffin-embedded tissue. J. Histochem. Cytochem. 56, 1033–1042.

Citovsky, V., Zambryski, P., 2000. Systemic transport of RNA in plants. Trends Plant Sci. 5, 52–54.

Cunningham, F., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., et al., 2015. Ensembl 2015. Nucleic Acids Res. 43, D662–D669.

Dobranszki, J., Magyar-Tabori, K., Hudak, I., 2008. In Vitro Tuberization in Hormone-Free Systems on Solidified Medium and Dormancy of Potato Microtubers. Fruit Veg. Cerearl Sci. Biotechnol. 1, 82–94.

Dohm, J.C., Minoche, A.E., Holtgräwe, D., Capella-Gutiérrez, S., Zakrzewski, F., Tafer, H., Rupp, O., Sörensen, T.R., Stracke, R., Reinhardt, R., Goesmann, A., Kraft, T., Schulz, B., Stadler, P.F., Schmidt, T., Gabaldón, T., Lehrach, H., Weisshaar, B., Himmelbauer, H., 2013. The genome of the recently domesticated crop plant sugar beet (Beta vulgaris). Nature 505, 546–549.

Drobysheva, D., Ameel, K., Welch, B., Ellison, E., Chaichana, K., Hoang, B., Sharma, S., Neckameyer, W., Srinakevitch, I., Murphy, K.J., others, 2008. An optimized method for histological detection of dopaminergic neurons in Drosophila melanogaster. J. Histochem. Cytochem. 56, 1049–1063.

Elias, M., Panaud, O., Robert, T., 2000. Assessment of genetic variability in a traditional cassava (Manihot esculenta Crantz) farming system, using AFLP markers. Heredity 85, 219–230.

Eviatar-Ribak, T., Shalit-Kaneh, A., Chappell-Maor, L., Amsellem, Z., Eshed, Y., Lifschitz, E., 2013. A Cytokinin-Activating Enzyme Promotes Tuber Formation in Tomato. Curr. Biol. 23, 1057–1064.

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.-K., Mockler, T.C., 2010. Genome-wide mapping of alternative splicing in Arabidopsis thaliana. Genome Res. 20, 45–58.

Fornara, F., Coupland, G., 2009. Plant Phase Transitions Make a SPLash. Cell 138, 625–627.

Francoz, E., Ranocha, P., Nguyen-Kim, H., Jamet, E., Burlat, V., Dunand, C., 2015. Roles of cell wall peroxidases in plant development. Phytochemistry 112, 15–21.

Fukaki, H., Nakao, Y., Okushima, Y., Theologis, A., Tasaka, M., 2005. Tissue-specific expression of stabilized SOLITARY-ROOT/IAA14 alters lateral root development in Arabidopsis: Tissue-specific Aux/IAA signaling in lateral root development. Plant J. 44, 382–395.

Fukaki, H., Tameda, S., Masuda, H., Tasaka, M., 2002. Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of Arabidopsis. Plant J. 29, 153–168.

Gamboa, A., Paéz-Valencia, J., Acevedo, G.F., Vázquez-Moreno, L., Alvarez-Buylla, R.E., 2001. Floral Transcription Factor AGAMOUS Interacts in Vitro with a Leucine-Rich Repeat and an Acid Phosphatase Protein Complex. Biochem. Biophys. Res. Commun. 288, 1018–1026.

Guan, A., 2009. Determination of Sucrose and Gibberellin effects on Brassica rapa Turnip Growth (Report).

Gutierrez, L., Bussell, J.D., Pacurar, D.I., Schwambach, J., Pacurar, M., Bellini, C., 2009. Phenotypic Plasticity of Adventitious Rooting in Arabidopsis Is Controlled by Complex Regulation of AUXIN RESPONSE FACTOR Transcripts and MicroRNA Abundance. Plant Cell 21, 3119–3132.

Havis, A.L., 1940. A DEVELOPMENTAL ANALYSIS OF KOHLRABI AND CABBAGE STEMS. J. Agric. Res. 61, 459–470.

Hendrikse, Y., 2015. Uncover selective sweeps for variation in diverse morphotypes of Brassica oleracea (Master thesis).

Jackson, D., 1992. In situ hybridization in plants, in: Molecular Plant Pathology. Oxford University Press, Oxford, pp. 163–174.

Jang, G., Lee, J.-H., Rastogi, K., Park, S., Oh, S.-H., Lee, J.-Y., 2015. Cytokinin-dependent secondary growth determines root biomass in radish (*Raphanus sativus* L.). J. Exp. Bot. 66, 4607–4619.

Jung, J., O'Donoghue, E.M., Dijkwel, P.P., Brummell, D.A., 2010. Expression of multiple expansin genes is associated with cell expansion in potato organs. Plant Sci. 179, 77–85.

Kiba, T., Takei, K., Kojima, M., Sakakibara, H., 2013. Side-Chain Modification of Cytokinins Controls Shoot Growth in Arabidopsis. Dev. Cell 27, 452–461.

Kubo, N., Saito, M., Tsukazaki, H., Kondo, T., Matsumoto, S., Hirai, M., 2010. Detection of quantitative trait loci controlling morphological traits in Brassica rapa L. Breed. Sci. 60, 164–171.

Lakhotia, N., Joshi, G., Bhardwaj, A.R., Katiyar-Agarwal, S., Agarwal, M., Jagannath, A., Goel, S., Kumar, A., 2014. Identification and characterization of miRNAome in root, stem, leaf and tuber developmental stages of potato (Solanum tuberosum L.) by high-throughput sequencing. BMC Plant Biol. 14, 6.

Langercrantz, U., 1998. Comparative Mapping Between Arabidopsis thaliana and Brassica nigra Indicates That Brassica Genomes Have Evolved Through Extensive Genome Replication Accompanied by Chromosome Fusions and Frequent Rearrangements. Genetics 150, 1217–1228.

Lin, K., Zhang, N., Severing, E.I., Nijveen, H., Cheng, F., Visser, R.G., Wang, X., de Ridder, D., Bonnema, G., 2014. Beyond genomic variation-comparison and functional annotation of three Brassica rapa genomes: a turnip, a rapid cycling and a Chinese cabbage. BMC Genomics 15, 250.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I.A.P., Zhao, M., Ma, J., Yu, J., Huang, S., et al., 2014. The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. Nat. Commun. 5.

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

Lu, G., Cao, J., Yu, X., Xiang, X., Chen, H., 2008. Mapping QTLs for root morphological traits inBrassica rapa L. based on AFLP and RAPD markers. J. Appl. Genet. 49, 23–31.

Mallory, A.C., Bartel, D.P., Bartel, B., 2005. MicroRNA-Directed Regulation of Arabidopsis AUXIN RESPONSE FACTOR17 Is Essential for Proper Development and Modulates Expression of Early Auxin Response Genes. PLANT CELL ONLINE 17, 1360–1375.

Maughan, J.A., Nugent, J.H., Hallahan, D.L., 1997. Expression of CYP71B7, a Cytochrome P450 Expressed Sequence Tag fromArabidopsis thaliana. Arch. Biochem. Biophys. 341, 104–111.

McFadden, G.I., Bonig, I., Cornish, E.C., Clarke, A.E., 1988. A simple fixation and embedding method for use in hybridization histochemistry on plant tissues. Histochem. J. 20, 575–586.

McVean, G., 2006. The Structure of Linkage Disequilibrium Around a Selective Sweep. Genetics 175, 1395–1406.

Meyerowitz, E.M., 1987. In situ hybridization to RNA in plant tissue. Plant Mol. Biol. Report. 5, 242–250.

Mitsui, Y., Shimomura, M., Komatsu, K., Namiki, N., Shibata-Hatta, M., Imai, M., Katayose, Y., Mukai, Y., Kanamori, H., Kurita, K., Kagami, T., Wakatsuki, A., Ohyanagi, H., Ikawa, H., Minaka, N., Nakagawa, K., Shiwa, Y., Sasaki, T., 2015. The radish genome and comprehensive gene expression profile of tuberous root formation and development. Sci. Rep. 5, 10835.

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

Narusaka, M., Seki, M., Umezawa, T., Ishida, J., Nakajima, M., Enju, A., Shinozaki, K., others, 2004. Crosstalk in the responses to abiotic and biotic stresses in Arabidopsis: analysis of gene expression in cytochrome P450 gene superfamily by cDNA microarray. Plant Mol. Biol. 55, 327–342.

Nielsen, R., 2005. Molecular signatures of natural selection. Annu Rev Genet 39, 197–218.

Nishijima, T., Sugii, H., Fukino, N., Mochizuki, T., 2005. Aerial tubers induced in turnip (Brassica rapa L. var. rapa (L.) Hartm.) by gibberellin treatment. Sci. Hortic. 105, 423–433.

O'Brien, P.J., 1972. The Sweet Potato: Its Origin and Dispersal. Am. Anthropol. 74, 342–365.

Passardi, F., Penel, C., Dunand, C., 2004. Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci. 9, 534–540.

Perl, A., Aviv, D., Willmitzer, L., Galun, E., 1991. In Vitro tuberization in transgenic potatoes harboring β-glucuronidase linked to a patatin promoter: effects of sucrose levels and photoperiods. Plant Sci. 73, 87–95.

Peterson, R.L., 1973. Control of cambial activity in roots of turnip (Brassica rapa). Can. J. Bot. 51, 475–480.

Qi, R., John, P.C.L., 2007. Expression of Genomic AtCYCD2;1 in Arabidopsis Induces Cell Division at Smaller Cell Sizes: Implications for the Control of Plant Growth. PLANT Physiol. 144, 1587–1597.

Ramireddy, E., Chang, L., Schmülling, T., 2014. Cytokinin as a mediator for regulating root system architecture in response to environmental cues. Plant Signal. Behav. 9, e27771.

Riou-Khamlichi, C., Menges, M., Healy, J.S., Murray, J.A., 2000. Sugar control of the plant cell cycle: differential regulation of Arabidopsis D-type cyclin gene expression. Mol. Cell. Biol. 20, 4513–4521.

Ross, K.F.A., 1953. Cell shrinkage caused by fixatives and paraffin-wax embedding in ordinary cytological preparations. Q. J. Microsc. Sci. 3, 125–139.

Sanz, L., Dewitte, W., Forzani, C., Patell, F., Nieuwland, J., Wen, B., Quelhas, P., De Jager, S., Titmus, C., Campilho, A., Ren, H., Estelle, M., Wang, H., Murray, J.A.H., 2011. The Arabidopsis D-Type Cyclin CYCD2;1 and the Inhibitor ICK2/KRP2 Modulate Auxin-Induced Lateral Root Formation. Plant Cell 23, 641–660.

Schmitz, R.Y., Skoog, F., Playtis, A.J., Leonard, N.J., 1972. Cytokinins: Synthesis and Biological Activity of Geometric and Position Isomers of Zeatin. PLANT Physiol. 50, 702–705.

Selman, I.W., Kulasegaram, S., 1967. Development of the stem tuber in kohlrabi. J. Exp. Bot. 18, 471–490.

Shiu, S.-H., Karlowski, W.M., Pan, R., Tzeng, Y.-H., Mayer, K.F.X., Li, W.-H., 2004. Comparative Analysis of the Receptor-Like Kinase Family in Arabidopsis and Rice. Plant Cell 16, 1220–1234.

Simonne, E.H., Smittle, D.A., Mills, H.A., 1993. Turnip growth, leaf yield, and leaf nutrient composition responses to nitrogen forms. J. Plant Nutr. 16, 2341–2351.

Soni, R., Carmichael, J.P., Shah, Z.H., Murray, J.A., 1995. A family of cyclin D homologs from plants differentially controlled by growth regulators and containing the conserved retinoblastoma protein interaction motif. Plant Cell 7, 85–103.

Sorin, C., 2005. Auxin and Light Control of Adventitious Rooting in Arabidopsis Require ARGONAUTE1. PLANT CELL ONLINE 17, 1343–1359.

Stewart, C.N., Halfhill, M.D., Warwick, S.I., 2003. Genetic modification: Transgene introgression from genetically modified crops to their wild relatives. Nat. Rev. Genet. 4, 806–817.

Sun, Y., Qiu, Y., Zhang, X., Chen, X., Shen, D., Wang, H., Li, X., 2015. Genome-wide identification of microRNAs associated with taproot development in radish (Raphanus sativus L.). Gene 569, 118–126.

Takei, K., Yamaya, T., Sakakibara, H., 2004. Arabidopsis CYP735A1 and CYP735A2 Encode Cytokinin Hydroxylases That Catalyze the Biosynthesis of trans-Zeatin. J. Biol. Chem. 279, 41866–41872.

Tong, C., Wang, X., Yu, J., Wu, J., Li, W., Huang, J., Dong, C., Hua, W., Liu, S., 2013. Comprehensive analysis of RNA-seq data reveals the complexity of the transcriptome in Brassica rapa. BMC Genomics 14, 689.

Torti, S., Fornara, F., Vincent, C., Andres, F., Nordstrom, K., Gobel, U., Knoll, D., Schoof, H., Coupland, G., 2012. Analysis of the Arabidopsis Shoot Meristem Transcriptome during Floral Transition Identifies Distinct Regulatory Patterns and a Leucine-Rich Repeat Protein That Promotes Flowering. PLANT CELL ONLINE 24, 444–462.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., Leunissen, J.A.M., 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 35, W71–W74.

Urieli-Shoval, S., Meek, R.L., Hanson, R.H., Ferguson, M., Gordon, D., Benditt, E.P., 1992. Preservation of RNA for in situ hybridization: Carnoy's versus formaldehyde fixation. J. Histochem. Cytochem. 40, 1879–1885.

Vanneste, S., De Rybel, B., Beemster, G.T.S., Ljung, K., De Smet, I., Van Isterdael, G., Naudts, M., Iida, R., Tasaka, M., Inzé, D., Fukaki, H., Beeckman, T., 2005. Cell Cycle Progression in the Pericycle Is Not Sufficient for SOLITARY ROOT/IAA14-Mediated Lateral Root Initiation in Arabidopsis thaliana. PLANT CELL ONLINE 17, 3035–3050.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.-H., Bancroft, I., Cheng, F., et al., 2011. The genome of the mesopolyploid crop species Brassica rapa. Nat. Genet. 43, 1035–1039.

Willemsen, J., 2014. The identification of candidate genes responsible for variation in potato tuber shape (Master thesis).

Wu, G., Poethig, R.S., 2006. Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. Development 133, 3539–3547.

Xiao, D., Zhang, N.-W., Zhao, J.-J., Bonnema, G., Hou, X.-L., 2012. Validation of reference genes for real-time quantitative PCR normalisation in non-heading Chinese cabbage. Funct. Plant Biol. 39, 342.

Xiao, D., Zhao, J.J., Hou, X.L., Basnet, R.K., Carpio, D.P.D., Zhang, N.W., Bucher, J., Lin, K., Cheng, F., Wang, X.W., Bonnema, G., 2013. The Brassica rapa FLC homologue FLC2 is a key regulator of flowering time, identified through transcriptional co-expression networks. J. Exp. Bot. 64, 4503–4516.

Yuan, Y.-X., Wu, J., Sun, R.-F., Zhang, X.-W., Xu, D.-H., Bonnema, G., Wang, X.-W., 2009. A naturally occurring splicing site mutation in the Brassica rapa FLC1 gene is associated with variation in flowering time. J. Exp. Bot. 60, 1299–1308.

Zhang, N., Zhao, J., Lens, F., de Visser, J., Menamo, T., Fang, W., Xiao, D., Bucher, J., Basnet, R.K., Lin, K., Cheng, F., Wang, X., Bonnema, G., 2014. Morphology, Carbohydrate Composition and Vernalization Response in a Genetically Diverse Collection of Asian and European Turnips (Brassica rapa subsp. rapa). PLoS ONE 9, e114241.

Zhao, J., Kulkarni, V., Liu, N., Pino Del Carpio, D., Bucher, J., Bonnema, G., 2010. BrFLC2 (FLOWERING LOCUS C) as a candidate gene for a vernalization response QTL in Brassica rapa. J. Exp. Bot. 61, 1817–1825.

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

Appendix A

Appendix Table 1: List of primers used for this work

Column one and four indicate the gene number which is used in this work. The corresponding gene ID is displayed in the primer name in column two and five in the primer name. In column three and six, the sequence of the primers if displayed. Abbreviations: G. No. – Gene Number, F – Forwards, R – Reverse.

G. No.	Primer name	Sequence (5' – 3')	G. No.	Primer name	Sequence (5' – 3')
1	1 -- Bra037585 F	GGCTGGTCTCAACATGAACA	**30**	30 -- Bra007380 R	GCATTGGCTTTGAACCTCAT
1	1 -- Bra037585 R	GATGAACTCGAACCATCTTGC	**31**	31 -- Bra007821 F	GACAGGTCCGGTCCATTTTA
2	2 -- Bra037574 F	ACGTTGAGGAAAACGCTGAT	**31**	31 -- Bra007821 R	GACCTAACACGGCAGCAAAC
2	2 -- Bra037574 R	AGCTTCGAAGGTCTGCTCTG	**32**	32 -- Bra031753 F	GGAAACGGATGAAGAAGAGG
3	3 -- Bra038700 F	AAGAATGCCCTCCTCCAAAT	**32**	32 -- Bra031753 R	ATCTTCGTCGGGGTCGAG
3	3 -- Bra038700 R	AAGCGTACGGAGCTCAAGAA	**33**	33 -- Bra008495 F	GCTGGGAGAGAAGGTCAGTG
4	4 -- Bra039755 F	CCGGTCAAAGCCAACAATAG	**33**	33 -- Bra008495 R	GCGTAGTTTCCGGTCAGAAG
4	4 -- Bra039755 R	TGGCACTGTAGATGGAGCTG	**34**	34 -- Bra022954 F	GCCAAACAGTACCACAAACG
5	5 -- Bra039749 F	AGCATACGATCAAATCGGAGA	**34**	34 -- Bra022954 R	CTCCTTCTCTCATTGTGTCCAG
5	5 -- Bra039749 R	CCGATAAAGGAGCGGAGTC	**35**	35 -- Bra030232 F	TGATGAGGAGCCTCACAGG
5	6 -- Bra039749(BOL) F	AGCATACGATCAAATCGGAGA	**35**	35 -- Bra030232 R	TTGTTGTTGTTGGTGTCTGGA
5	6 -- Bra039749(BOL) R	GCCACATCCAACGACTGAAG	**36**	36 -- Bra032175 F	TCATGGAGCTACGATCACGA
7	7 -- Bra034022 F	GCGACCTGATCATCCCTAAA	**36**	36 -- Bra032175 R	GCGAAAAACTCAACCGGTAA
7	7 -- Bra034022 R	CCTCCATCATCGCAAAATTC	**37**	37 -- Bra032097 F	CGACAGCAAGGAAGACAACA
8	8 -- Bra034021 F	ATGGAACGTCAAGGCATCAC	**37**	37 -- Bra032097 R	ATGGAAAGCTTGGTGGCTAA
8	8 -- Bra034021 R	GACATCCAGGAACGTTGGAC	**38**	38 -- Bra032056 F	ATTTGCATGTGTTGCGAGTC
9	9 -- Bra033109 F	TAAGCTCGACCTGGAGGAGA	**38**	38 -- Bra032056 R	AGGCTTGTTCCATACGCTTAG
9	9 -- Bra033109 R	TTGTTGCAGACTCATGTCCTG	**39**	39 -- Bra034257 F	ATCGTACGCGAGTCCGTTTA
10	10 -- Bra029302 F	TATGGGAATATGCCCTTTGG	**39**	39 -- Bra034257 R	TTCGCTGGAGTCACAAAGTC
10	10 -- Bra029302 R	AGGAGCAACTGGTTGACTAGG	**40**	40 -- Bra024902 F	TGCTTCTGCCTCCGTTACTT
11	11 -- Bra029281 F	TGGTGAGAGGAAAGATAGAGATGA	**40**	40 -- Bra024902 R	TGGAACACGTGTCTGCAATC
11	11 -- Bra029281 R	TTGGAAAATTCGTAAAGCCTTC	**41**	41 -- Bra024859 F	CCCGACTTTGAGACAGGAAG
12	12 -- Bra029261 F	AACGGTCCGGTCCTTAGTTT	**41**	41 -- Bra024859 R	TCCCCTGCCTCTAACTCCTT
12	12 -- Bra029261 R	CCAATAAGGGTAAACCGGGTA	**42**	42 -- Bra025121 F	ATGAAGCGCTGTTCCTTACC
13	13 -- Bra024194 F	CAAATTACCACCGGGGTTTC	**42**	42 -- Bra025121 R	TGGTTAGACTGGAGGCGTTT
13	13 -- Bra024194 R	GCCATGGCTCACAAATATGC	**43**	43 -- Bra033738 F	GCTTTTTATAGTGGCGGTGCT
14	14 -- Bra010210 F	AGCTCAAGTTCCGAGCACAT	**43**	43 -- Bra033738 R	TCCGAATTCGAGTTCTTTGTT
14	14 -- Bra010210 R	TCCACTTCCAAATCCTGCAT	**44**	44 -- Bra004097 F	CAATCATCAGTTAACAGCATCACC
15	15 -- Bra019821 F	CGATGGGGAGCTCAGATTTA	**44**	44 -- Bra004097 R	GCCGTGTGAATCTTGCTATG
15	15 -- Bra019821 R	CTCGGCTTTCGAATGAGAAC	**45**	45 -- Bra004057 - 59 F	TGCACATGGTGAGGATCAGT
16	16 -- Bra019878 F	ACCATGCTCCTCCTCTTCCT	**45**	45 -- Bra004057 - 59 R	AGAAGAACGTCCGCTTGGT
16	16 -- Bra019878 R	ACCCTGAACCTGAAGGAGGT	**46**	46 -- Bra004226 F	CGAGCAAAAACATGCTCAGA
17	17 -- Bra019808 F	AAAACGATCTGTTGGTTAAAAAGC	**46**	46 -- Bra004226 R	TTCAGCACTGCTCTCACCAC
17	17 -- Bra019808 R	GAGCTCGCAAGAACCATCTC	**47**	47 -- Bra014208 F	CCGCTACTTGGCTTTGAGAG
18	18 -- Bra019809 F	TGTCAACGTTCATCTACTTCCTCT	**47**	47 -- Bra014208 R	CTCGCGAGTGCCTTATCAGT
18	18 -- Bra019809 R	AAACGGAGAAGCATCACTGG	**48**	48 -- Bra014073 F	GTCGGTTCTCGACGACTTCA
19	19 -- Bra018223 F	CACGTATGTCGTCAGGGAAA	**48**	48 -- Bra014073 R	GGCAGACACGACAAACCAA
19	19 -- Bra018223 R	GCAATAGCTTTGCTCTAAATAGATGT	**49**	49 -- Bra034889 F	TGAGGAGGTAGAGAAACAACAAAG
20	20 -- Bra003665 F	GCGAAGTCGACCCTCAGAT	**49**	49 -- Bra034889 R	CTTGAACTTCAGAGAACTCCCAAT

continued on next page

G. No.	Primer name	Sequence (5' – 3')	G. No.	Primer name	Sequence (5' – 3')
20	20 -- Bra003665 R	GGAGACGACGCAGGGAAC	**50**	50 -- Bra039732 F	ATGAAGAAGTCATCAATGAGCTTAAC
21	21 -- Bra003743 F	CACGGTGCAAGCTCTCAAGT	**50**	50 -- Bra039732 R	ATTCTTCTTTTAGATTGTGCTCGTC
21	21 -- Bra003743 R	ACCGTGGGTGGTGAGGTTAG	**51**	51 -- Bra039855 F	AAACGGAGCTCTGTCTTGG
22	22 -- Bra003969 F	CACGGGGCTGCATAGAGTAG	**51**	51 -- Bra039855 R	GGCTTAGAAGGGTCTTTGTGG
22	22 -- Bra003969 R	TGGCGTCCTCTCTTCTGTTT	**52**	52 -- Bra010581 F	TCCGTAAGCATCCTTTGTCA
23	23 -- Bra007385 F	GGGATGTGTGTTTGCTCCAT	**52**	52 -- Bra010581 R	ACTCACCTCTCCAGCTACGC
23	23 -- Bra007385 R	CTGCTGCTGGAGAGCAAAG	**53**	53 -- Bra016485 F	CTGGAGCTGGAGACAAGAGC
24	24 -- Bra030639 F	CAGTGACGGGCAGCTTTATAC	**53**	53 -- Bra016485 R	GCGGAAGAGATGCTGAGAAC
24	24 -- Bra030639 R	CCTCCGGGAACAAACTGATA	**54**	54 -- Bra016498 F	TCAGCGACGACAGCTATGAC
25	25 -- Bra037401 F	ATGTCAGGAGGGCAAAACTG	**54**	54 -- Bra016498 R	ACCCATTCTTAGGAGTGAAACAA
25	25 -- Bra037401 R	TGATGTTCATCACGATCTCCA	**55**	55 -- Bra016553 F	GCTGCAATAGCTCGTGGATT
26	26 -- Bra017432 F	GGCGGAAAACGAGTTTATGA	**55**	55 -- Bra016553 R	GAATAACGTTTTGCTTCTCTGGA
26	26 -- Bra017432 R	CCGTCAGTATTCGCTGATGA	**56**	56 -- Bra036693 F	GCCCAGAGAGACGGTAGATG
27	27 -- Bra017416 F	ATCGTCCAATGCCTTATGGT	**56**	56 -- Bra036693 R	GTTCTGTTGAGGGTGGTCGT
27	27 -- Bra017416 R	GCGTTAGCGCCACTGTAGTT	**57**	57 -- Bra009091 F	TCTGCTCCTCATGGATAGGC
28	28 -- Bra024672 F	GATCATCACGTTCAAGACAGTCA	**57**	57 -- Bra009091 R	TCAGAAGGCAAAGATCCAAAG
28	28 -- Bra024672 R	GGATAAGGTGTTGGAGCCTTG	**58**	58 -- Bra009202 F	ACGCAACATCCGTCTACAAG
29	29 -- Bra007349 F	GTCATGGCTGTGGAGGAAG	**58**	58 -- Bra009202 R	CATCATCCACACCACCAGAC
29	29 -- Bra007349 R	AGCTGAAGAGGGACACAAGG	**59**	59 -- Bra009242 F	GGAGGAGCTAGACCATAAACG
30	30 -- Bra007380 F	GTCAACGGCAATGTTGGTC	**59**	59 -- Bra009242 R	CAAGAATTGGCTCGTTGACA

Appendix Table 2: Positions of investigated Selective Sweeps (piHS) and QTLs on *B. rapa* chromosomes
Number indicates the number as which the Selective Sweep or the QTL (which not overlap with any Selective Sweep) is named in this work. The next column indicates the Chromosome on which the Selective Sweep or QTL is located with detailed Start and Stop positions in the following two columns. Abbreviations: Chrom. – Chromosome, App. - Approximate

Selective Sweep				QTL			
Number	**Chrom.**	**Start**	**Stop**	**Number**	**Chrom.**	**App. Start**	**App. Stop**
1	A01	90,001	455,000	1	A02	14,244,720	15,751,914
2	A01	1,145,001	1,530,000	2	A03	7,840,039	8,044,007
3	A01	11,260,001	11,515,000	3	A04	8,809,811	11,887,136
4	A01	21,170,001	21,420,000	4	A04	16,890,061	17,684,759
5	A01	21,650,001	22,325,000	5	A06	22,181,738	24,422,022
6	A01	24,200,001	24,570,000	6	A07	197,035	864,340
7	A02	4,960,001	5,175,000	7	A07	19,595,000	end
8	A02	8,705,001	8,905,000	8	A08	2,266,825	6,440,078
9	A02	8,930,001	9,175,000	9	A08	14,203,925	15,551,142
10	A02	12,370,001	12,575,000	10	A08	17,255,004	18,290,476
11	A02	22,420,001	22,875,000	11	A09	4,877,357	6,952,267
12	A02	25,445,001	25,810,000	12	A10	12,028,635	13,384,144
13	A03	26,775,001	27,100,000	13	A10	13,805,747	16,409,300
14	A04	875,001	1,165,000				
15	A05	14,265,001	15,090,000				
16	A05	15,295,001	15,640,000				

continued on next page

Selective Sweep				QTL			
Number	Chrom.	Start	Stop	Number	Chrom.	App.Start	App. Stop
17	A06	3,800,001	4,400,000				
18	A06	10,810,001	11,035,000				
19	A07	11,340,001	11,575,000				
20	A07	17,425,001	17,625,000				
21	A07	18,030,001	18,230,000				
22	A07	19,275,001	19,595,000				
23	A08	20,145,001	20,475,000				
24	A09	330,001	670,000				
25	A09	13,770,001	14,185,000				
26	A09	24,700,001	25,010,000				
27	A09	30,215,001	30,455,000				
28	A09	32,365,001	32,760,000				
29	A09	36,265,001	36,690,000				
30	A09	36,755,001	36,985,000				
31	A10	3,200,001	3,605,000				

Appendix Table 3: Position of Selective Sweeps (piHS) on *B. oleracea* chromosomes

Number indicates the number as which the Selective Sweep is named in this work. The next column indicates the Chromosome on which the Selective Sweep is located with detailed Start and Stop positions in the following two columns. Abbreviation: Chrom. – Chromosome.

Number	Chrom.	Start	Stop	Number	Chrom.	Start	Stop
1	C01	11305001	11680000	20	C07	7360001	7765000
2	C02	2340001	2595000	21	C07	10520001	11960000
3	C03	3000001	3265000	22	C07	16165001	16425000
4	C03	6135001	7115000	23	C07	30560001	30825000
5	C03	7270001	8230000	24	C07	30875001	31245000
6	C03	9415001	9675000	25	C07	40295001	40580000
7	C03	9745001	10065000	26	C07	42245001	42515000
8	C03	16850001	17060000	27	C07	42545001	43040000
9	C03	21110001	21710000	28	C07	44355001	44990000
10	C03	22070001	23430000	29	C07	45160001	45630000
11	C03	24605001	25495000	30	C08	490001	1035000
12	C03	26110001	26655000	31	C08	1440001	1900000
13	C03	38110001	38455000	32	C08	8480001	8935000
14	C03	44670001	45105000	33	C08	35350001	35630000
15	C03	53270001	54745000	34	C08	36170001	36600000
16	C03	55000001	55385000	35	C08	38895001	39280000
17	C04	29140001	29515000	36	C09	20455001	21695000
18	C05	21370001	21785000	37	C09	24680001	25035000
19	C07	3865001	4370000				

Appendix Table 4: Characteristics of the 58 selected genes

The first column (No.) indicates the number with which the gene is identified in the text of this work. The second column shows the official *B. rapa* gene ID, whereas the third column (*A. thal.* orth.) shows the *A. thaliana* orthologue which (predicted) function is displayed in the fourth column. The fifth and the sixth column indicate, whether the gene is placed in a Selective Sweep (Sel. Sw.) and/or a *B. rapa* QTL – A number indicates the number of the Selective Sweep or QTL, (Y) indicates an overlapping QTL and (N) indicates non-presence. The last column (*Bol* Sel. Sw.) indicates whether an orthologue of this gene is also found in a *B. oleracea* Selective Sweep – a number the number of the Selective Sweep in case of presence, whereas (N) indicates non-presence of an orthologue of this gene in a *B. oleracea* Selective Sweep.

No.	Gene ID	*A. thal.* orth.	predicted function	Sel. Sw.	QTL	*Bol* Sel. Sw.
1	Bra037585	AT3G18400	NAC domain containing protein 58	4	N	10
2	Bra037574	AT3G18600	DEAD-box ATP-dependent RNA helicase 51	4	N	10
3	Bra038700	AT3G12145	FLOR 1 - LRR, interacts with agamous	6	Y	N
4	Bra039755	AT1G07440	tropinone reductase-like protein	8	Y	31
5	Bra039749	AT1G66080	MYB75 - transcription factor	8	Y	N
7	Bra034022	AT1G67110	CYP735A2 - cytokinin trans-hydroxylase	9	Y	N
8	Bra034021	AT1G67110	CYP735A2 - cytokinin trans-hydroxylase	9	Y	N
9	Bra033109	AT5G52060	BCL-2-associated athanogene 1	11	Y	5
10	Bra029302	AT5G61890	ERF114	12	Y	N
11	Bra029281	AT5G62165	AGL42 - AGAMOUS-like 42	12	Y	N
12	Bra029261	AT5G62430	K19B1.4 - cycling DOF factor 1	12	Y	N
13	Bra024194	AT4G28530	NAC074 - NAC domain containing protein 74	13	N	(28 & 12)
14	Bra010210	AT1G33760	ERF022	15	N	N
15	Bra019821	AT1G13260	RAV1 - related to ABI3/VP1 1;	17	N	N
16	Bra019878	AT1G11130	SUB - STRUBBELIG	17	N	N
17	Bra019808	AT1G13110	CYP71B7 - cytochrome P450 71B7	17	N	N
18	Bra019809	AT1G13110	CYP71B7 - cytochrome P450 71B7	17	N	N
19	Bra018223	AT3G46130	myb domain protein 48	18	N	N
20	Bra003665	AT1G77850	ARF17 - auxin response factor 17	20	N	N
21	Bra003743	unknown	unknown	21	N	N
22	Bra003969	AT1G70000	myb-like transcription factor-like protein	22	Y	N
23	Bra007385	AT3G58190	LBD29 - LOB domain-containing protein 29	23	Y	N
24	Bra030639	AT1G06110	SKIP16 - F-box protein SKIP16	23	Y	N
25	Bra037401	AT4G00940	Dof zinc finger protein DOF4.1	24	Y	N
26	Bra017432	AT2G01940	SGR5 - SHOOT GRAVITROPISM 5	25	N	N
27	Bra017416	AT2G02450	NAC035 - LONG VEGETATIVE PHASE 1	25	N	N
28	Bra024672	AT1G26945	KDR - KIDARI	26	Y	N
29	Bra007349	unknown	unknown	27	Y	N
30	Bra007380	AT3G58120	leucine zipper transcription factor-like	27	Y	N
31	Bra007821	AT2G25060	ENODL14 - early nodulin-like protein	28	Y	N
32	Bra031753	AT1G10720	BSD domain-containing	29	N	35
33	Bra008495	AT1G80130	tetratricopeptide repeat domain protein	N	1	N
34	Bra022954	AT2G33810	SPL3	N	2	25
35	Bra030232	AT2G22490	CYCD2;1 - Cyclin D2;1	N	3	N
36	Bra032175	AT2G23300	Leucine-rich repeat protein kinase family	N	3	N
37	Bra032097	AT2G24400	SAUR-like auxin-responsive protein	N	3	N
38	Bra032056	AT2G24860	DnaJ/Hsp40	N	3	N
39	Bra034257	AT1G11280	S-locus lectin protein kinase-like protein	N	3	N
40	Bra024902	AT1G15100	RHA2A - RING-H2 finger A2A	N	5	N
41	Bra024859	unknown	unknown	N	5	N
42	Bra025121	AT2G03620	magnesium transporter CorA-like	N	5	24
43	Bra033738	AT5G44350	ethylene-responsive nuclear protein-like	N	5	N
44	Bra004097	AT1G67260	transcription factor TCP1	N	7	N

continued on next page

No.	Gene ID	*A. thal.* orth.	predicted function	Sel. Sw.	QTL	*Bol* Sel. Sw.
45	Bra004057 – Bra004059	AT3G45140	LOX2 - lipoxygenase 2	N	7	N
46	Bra004226	AT1G67440	emb1688 - embryo defective 1688	N	7	N
47	Bra014208	AT1G49750	leucine-rich repeat-like protein	N	8	30
48	Bra014073	unknown	unknown	N	8	N
49	Bra034889	unknown	unknown	N	8	N
50	Bra039732	AT4G14560	IAA1 - auxin-responsive protein IAA1	N	8	N
51	Bra039855	AT4G14550	IAA14 - auxin-responsive protein IAA14	N	8	N
52	Bra010581	AT4G37180	myb family transcription factor	N	9	N
53	Bra016485	AT1G19960	unknown	N	10	N
54	Bra016498	AT1G19640	JMT processes and defense responses	N	10	N
55	Bra016553	AT1G18570	myb domain protein 51	N	10	N
56	Bra036693	AT1G20750	RAD3-like DNA-binding helicase protein	N	11	N
57	Bra009091	AT5G05160	RUL1 - REDUCED IN LATERAL GROWTH1	N	13	N
58	Bra009202	unknown	unknown	N	13	N
59	Bra009242	unknown	unknown	N	13	N

Appendix B

RNA in situ hybridization protocol for Arabidopsis embryos (Weijers Lab)

Based on the Jeff Long *in situ* protocol and adapted by Shunsuke Saiga. Reference for this protocol:

SAIGA, S., FURUMIZU, C., YOKOYAMA, R., KURATA, T., SATO, S., KATO, T., TABATA, S., SUZUKI, M. & KOMEDA, Y. 2008. The Arabidopsis OBERON1 and OBERON2 genes encode plant homeodomain finger proteins and are required for apical meristem maintenance. *Development,* 135**,** 1751-1759.

Jeff Long protocol: http://pbio.salk.edu/pbiol/in_situ_protocol.html

Fixating and sectioning of tissue

Fixation of tissue

Collect siliques at the stage required from the desired plants.

Prepare the fixation (4% PFA) solution:

Fixation solution	50 ml tube	
PFA	2 g	Dissolve at 70°C
Sterilized MQ	Fill to 43 ml	
1M NaOH	50 µl	
10xPBS	5 ml	Add after cooling down
Dimethyl sulfoxide (DMSO)	2 ml	Add after cooling down

Cut the siliques in pieces of about 1 cm. Cut both sides of the siliques open to allow the fixative to easily enter the tissue. Place the cut pieces in a sterilized vial on ice with about 10 ml of ice cold fixative.

Place the glass containers, while still on ice, in a near vacuum for 15-30 minutes.

Repeat the previous step if the siliques have not sunk to the bottom.

Replace the fixation solution with fresh fixative and place overnight at 4 °C.

Replace the fixation solution with subsequently, all at 4 °C

1xPBS (30 minutes)

1xPBS (30 minutes)

30%EtOH (60 minutes)

40% EtOH (60 minutes)

50% EtOH (60 minutes)

60% EtOH (60 minutes)

70% EtOH (60 minutes)

*Note: t*he tissue can be stored in 70% EtOH for several months at 4 °C.

Replace the 70% EtOH with subsequently:

85% EtOH (60 minutes) at 4 °C

96% EtOH (overnight) at 4 °C

100% EtOH (30 minutes) at room temperature

100% EtOH (30 minutes) at room temperature

100% EtOH (60 minutes) at room temperature

100% EtOH (60 minutes) at room temperature

Replace the 100% EtOH with subsequently, all at room temperature in the fume hood:

25% xylene (1 hour)

50% xylene (1 hour)

75% xylene (1 hour)

100% xylene (1 hour)

100% xylene (1 hour)

100% xylene with 10 to 20 paraffin chips (overnight)

Melt paraffin chips at 63 °C overnight

Replace the 100% xylene with melted paraffin. Doing this next to an open flame, in a warm environment will make this easier. Keep at 63 °C.

Replace the melted paraffin with fresh after several hours. Repeat this wax change for two more days, twice a day.

Preheat the plastic base molds and their covers at 63 °C.

Prepare a heat block covered in aluminium foil on which a mold can be placed while it is being handled.

Using a glass Pasteur pipette with the end cut off pipet three siliques into the mold. Fill the mold with more melted paraffin.

Using a heated needle place the siliques aligned at the bottom of the mold.

Place a mold cover on top of the mold, fixate with more paraffin and let slowly set.

Store at 4 °C until use.

Sectioning tissue

Make 8 µm sections of siliques and transfer a ribbon of film (several sections attached to each other) into a 45°C water bath. Put them on a superfrost slide. After each slide (two ribbons per slide) check for the presence of embryos in the sections by DIC optics. At least 3 good embryos should be present per slide. "Bake" the slides overnight at 45°C and store at 4°C until use. It is advisable to use a new section of the microtome blade every one to two blocks.