Genotypic and phenotypic analysis of the allelic diversity in candidate genes for oil content in exotic plant materials of rapeseed (*Brassica napus L.*)

Dissertation

zur Erlangung des naturwissenschaftlichen Doktorgrades

im Rahmen von GAUSS an der Fakultät für Agrarwissenschaften

der Georg-August-Universität Göttingen

vorgelegt von

Daniela Katja Weis

geboren in Wertheim am Main

Göttingen, im Juli 2014

D 7

1. Referent: Herr PD Dr. Wolfgang Ecke

2. Korreferent: Herr Prof. Dr. Reiner Finkeldey

Tag der mündlichen Prüfung: 24. Juli 2014

Für meine Familie

Contents

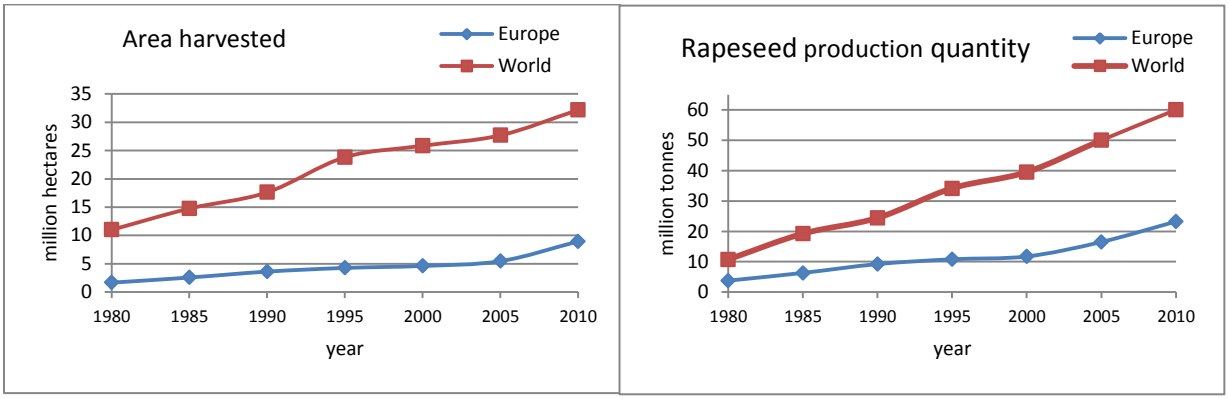
List of tables

I. General Introduction

I.1 Current status of rapeseed production

Today, rapeseed (*Brassica napus* L.) is the major oil crop in Europe (19.2 million tons) followed by sunflower (6.9 million tons), cotton (1.0 million tons) and soybean (0.9 million tons). Worldwide, rapeseed is the second most important oilseed crop after soybean (United States Department of Agriculture “Oil crops year book” August 2013). The demand for rapeseed oil grew significantly in the last decades, as shown in figure 1 (Metzger and Bornscheurer 2006; FAO statistics from November 2013).

Figure 1: Development of oilseed rape harvested area and production quantity in Europe and worldwide for the period 1980 - 2010 (FAO statistics from November 2013)



The worldwide rapeseed production reached an all-time record in 2012 with 64.8 million tons (FAO statistics from November 2013). In Europe, 19.2 million tons of rapeseed were harvested in that year. This is 29.6% of the world production which makes Europe the biggest rapeseed producer, followed by Canada (23.7%) and China (21.6%). In the biggest rapeseed producing European countries Germany and France, together 10.3 million tons were harvested (FAO statistics from November 2013). The production increase is mainly attributable to the increased area of rapeseed cultivation, as shown in figure 1. But also the increasing frequency of using higher yielding hybrids (Osborn *et al.*, 2007; Gehringer *et al.*, 2007) plays an important role. The higher yield in hybrids is based on the phenomenon of heterosis,

which means the superior performance of F_1 hybrids produced by a cross between genetically distinct homozygous parents to their midparent value or to the value of the better parent.

The main usage of rapeseed oil is for industrial purposes, especially for the production of biodiesel. According to the United States Department of Agriculture, in Europe almost 74% of the domestic consumption of rapeseed oil of the crop year 2012/13 was used for industrial purposes (USDA report: “European Union: Oilseeds and Products Supply and Distribution”, November 2013). Rapeseed oil is also used in human nutrition and the byproducts of the oil extraction, meal and press cake, are used as animal feed. This is possible due to the successful development of zero erucic acid and low glucosinolate rapeseed, also described as “canola” quality rapeseed. Because oilseed rape with high erucic acid and glucosinolate content should not be used in human and animal nutrition due to toxicity of these substances. In animal testing myocardial insufficiency and growth disturbance was observed after feeding erucic acid (Thomasson, 1955; Beare *et al.*, 1959; Roine and Uksila, 1959). The toxicity of glucosinolates is attributable to the formation of thiocyanates, oxazolidinethions and nitriles, leading for example to thyroid disorder (for review see Andersson *et al*., 2008).

I.2 The rapeseed genome

The cultivated *Brassica* species represent the group of crops most closely related to *Arabidopsis* (Rana *et al.*, 2004). The lineages of the *Arabidopsis* and *Brassica* species diverged about 20 - 43 million years ago (Yang *et al*., 1999; Town *et al.*, 2006; Beilstein *et al.,* 2010). The genome of *Arabidopsis*, which is one of the smallest known in higher plants (130 Mbp), has only 5 chromosomes in the haploid

complement (Schmidt *et al.*, 2001; *Arabidopsis* Genome Initiative, 2000) and is well characterized. The whole nucleotide sequence of *A. thaliana* has been determined and annotated and is available for the public at the website http://www.arabidopsis.org. The sequences are well cross-referenced with genetic and physical chromosome maps and EST (Expressed Sequence Tags). Also cDNAs are available, which are important for gene annotation (*Arabidopsis* Genome Initiative, 2000). Comparison of coding genes of *Brassica napus* and *Arabidopsis thaliana* revealed an average identity of coding sequences of 87% (Cavell *et al*., 1998). Brunel *et al*. (1999) and Grant *et al.* (1998) discovered similar values when comparing *Brassica napus* and *A. thaliana* at the protein sequence level. Similar values of sequence identity on nucleotide and amino acid level were evaluated by Bach (2007) for six candidate genes on oil content: wrinkled *(WRI),* pyruvate dehydrogenase kinase *(PDHK),* plastidial pyruvate dehydrogenase *(PDH),* plastidial pyruvate kinase 2 *(PKP2),* biotin carboxyl carrier protein 2 *(BCCP2),* and acyl-CoA diacylglycerol acyltransferase 1 *(DGAT).*

With 1.2 Gb, the *Brassica napus* (oilseed rape) genome is estimated to be about nine times larger (Arumuganathan and Earle, 1991) than the genome of *Arabidopsis*. Most probably several spontaneous hybridizations between turnip rape (*Brassica rapa L*.), genome AA (2n=20), and cabbage (*Brassica oleracea L.),* genome CC (2n= 18), led to the origin of the amphidiploid oilseed rape (*Brassica napus L*.), genome AACC (2n=38) (U, 1935; Kimber and McGregor, 1995; Allender and King, 2010; Bancroft *et al.,* 2011). In genetic mapping experiments of Slocum *et al*. (1990) and Song *et al.* (1991) it has been noted that a high proportion of the genomes of the two progenitor species *B. oleracea* and *B. rapa* are duplicated. Lagercrantz and Lydiate (1996) showed in comparative mapping experiments high genome collinearity for *B. rapa*,

B. oleracea and *B. nigra*. They developed an RFLP map in a particularly polymorphic *B. nigra* (genome BB, 2n=16) cross which revealed eight sub-chromosomal segments, each present in three copies. By comparative mapping they also showed that each part of the B genome has a corresponding part in the A and C genome. This led to the conclusion that these three *Brassica* genomes descended from a common hexaploid ancestor and have a triplicated genic structure. This theory was confirmed by Lysak *et al.* (2005) using a cytogenetic approach. They concluded that a distinctive feature of the tribe *Brassicaceae* is that they contain triplicated genomes and are descended from a common hexaploid ancestor with basic genomes similar to that of *Arabidopsis*. Also further studies of the genomes of *B. oleraceae*, *B. rapa*, and *B. napus* confirmed the fundamentally triplicated nature of the diploid *Brassica* genomes (O´Neill and Bancroft, 2000; Park *et al*., 2005; Rana *et al.,* 2004; Ziolkowski *et al.,* 2006; The *Brassica rapa* Genome Sequencing Project Consortium, 2011; Beilstein *et al.,* 2010; Bancroft *et al*., 2011). The presumed hexaploidization event is thought to have occurred 5 - 22 million years ago (Lysak *et al.,* 2005; The *Brassica rapa* Genome Sequencing Project Consortium, 2011; Beilstein *et al.,* 2010). As shown in many studies, not all genes exist in three copies in the *B. napus* progenitor species or six copies in the *B. napus* genome, respectively (Lan *et al.,* 2000; Babula *et al.,* 2003; Lukens *et al.,* 2003; Rana *et al.,* 2004). For example, Parkin *et al.* (2005) found evidence of segmental gene loss as well as additional segmental duplications in *B. napus* compared to the homologous genes in *Arabidopsis*. Also in the published genome of *B. rapa* (The *Brassica rapa* Genome Sequencing Project Consortium, 2011) it is clearly shown that some genes are less than triplicated. Some genes even have been limited to one copy. This substantial gene loss is thought to be typical after polyploid formation in eukaryotes (The *Brassica rapa* Genome Sequencing Project Consortium, 2011). These results together with the fact that the progenitor A

and C genomes are essentially intact in *B. napus* and have not been rearranged, as shown in genetic mapping experiments (Parkin *et al.,* 1995) and analysis of genome microstructure (Rana *et al.*, 2004), mean that in *B. napus* each gene exists at a minimum of two with many genes having even higher copy numbers.

I.3 Genetic diversity in rapeseed and resources for increasing the genetic diversity in winter rapeseed

The present breeding materials of oilseed rape have a small diversity (Diers and Osborn, 1994; Becker *et al.,* 1995; Lombard and Delourme, 2001; Jesske, 2011). This is attributable to the narrow genetic base as well as the recent origin, and that extensive rapeseed breeding started not more than 70 years ago (Becker *et al.,* 1995; Allender and King, 2010). Although the hybridization of the parental species most probably has occurred several times in the evolutionary formation of rapeseed, the present breeding materials of oilseed rape is derived from very few interspecific hybrid plants that occurred some centuries ago in a limited geographical region (Becker *et al.,* 1995). Also, recent breeding practices have led to a relatively small genetic diversity in modern winter rapeseed breeding materials (Seyis *et al.,* 2001; Hasan *et al.,* 2006). For example, the trait “canola” quality could only be established by using strongly restricted genetic materials. To get an erucic acid free variety, crosses with mutants that lack the ability to synthesize erucic acid were done. To establish “low glucosinolate” content, a Polish variety with very low glucosinolate content was used.

Today, an increased genetic diversity is of special interest to increase oil content in rapeseed cultivars. An increased genetic diversity among parents of hybrid cultivars is potentially interesting with respect to increased heterosis for the triait yield as shown in many studies (Diers *et al.,* 1996; Butruille *et al.,* 1999; Riaz *et al.,* 2001;

Girke *et al*., 2012b; Zou *et al*., 2010). It is also possible that additive effects or dominance effects on oil content could occur due to the effect of novel alleles in genes for oil content (Würschum *et al*., 2013). To broaden the genetic base, different ecotypes and wild relatives are valuable resources (Tanksley and McCouch, 1997; Becker, 2001; Hasan *et al*., 2006; Osborn *et al.,* 2007; Girke *et al*., 2012a; Jesske *et al.,* 2013a). So, introgressing alleles from spring type rapeseed and rapeseed of Asian origin, which represent distinct gene pools in comparison to European winter rapeseed cultivars (Diers and Osborn, 1994; Becker *et al.,* 1995; Plieske and Struss, 2001; Shengwu *et al.,* 2003; Hasan *et al.*, 2006; Girke *et al.,* 2012a), could be used to increase allelic diversity of winter rapeseed cultivated in Europe. But the most important source to increase the genetic diversity of oilseed rape lays in its artificial resynthesis from the two progenitor species. Large genetic and phenotypic variability is observed within *B. rapa* and *B. oleracea* (Becker *et al.,* 1995 and 1999; Girke *et al.*, 2012a; Seyis *et al.,* 2001 and 2003; Jesske *et al*., 2013a). Even larger genetic variability could be attained by using wild relatives from the parental species for the artificial resynthesis of rapeseed (Allender *et al*., 2007; Jesske *et al*., 2013a).

I.4 Allelic diversity in candidate genes for oil content

During an earlier GABI project, GABI BRIDGE, allelic diversity in nine candidate genes for oil content was evaluated in rapeseed. The nine candidate genes were wrinkled *(WRI),* pyruvate dehydrogenase kinase *(PDHK),* plastidial pyruvate dehydrogenase *(PDH),* plastidial pyruvate kinase 2 *(PKP2),* 3-ketoacyl-acyl carrier protein synthase III *(KAS III),* biotin carboxyl carrier protein 2 *(BCCP2),* acetyl Co-enzyme A carboxylase carboxyl transferase alpha subunit *(CAC3),* acetyl Co-enzyme A carboxylase biotin carboxylase subunit *(CAC2)* and acyl-CoA diacylglycerol acyltransferase 1 *(DGAT).* Twenty-eight loci of these candidate genes for oil content

were analyzed by comparative sequencing in a core set of genetically diverse rapeseed genotypes (Ecke and Lange, Schmidt, personal communication). The core set consists of 18 relatively new winter rapeseed genotypes, four old winter rapeseed varieties, four forage rapeseed genotypes, five resynthesized genotypes and one Asian genotype (Ecke and Lange, personal communication).

A large number of polymorphisms were observed at the candidate gene loci. Based on this sequence information, SNP and InDel markers were established and the common alleles were checked for phenotypic effects by association analysis. A number of loci were identified where the allelic diversity significantly affected oil content (Ecke and Lange, personal communication). During the comparative sequencing a number of novel alleles were detected in the few exotic genotypes included in the core set, which were not present in adapted germplasm. Due to the low frequency of these alleles, it was not possible to determine their phenotypic effects by association analysis. This led to the idea to analyze a larger number of exotic genotypes to detect additional novel alleles in candidate genes for oil content and to determine possible phenotypic effects of these alleles in segregating F_2 populations.

I.5 Candidate genes for oil content

The oil content in rapeseed is under polygenic control. The investigated genes in this project were genes coding for proteins involved in fatty acid biosynthesis, in supplying substrates for fatty acid biosynthesis, in the transcriptional control of this biochemical pathway or genes involved in seed maturation and development.

I.5.1 Genes coding for plastidial pyruvate kinase subunits: PKP1 (coding for the α-subunit), PKP2 (coding for the β1-subunit) and PKP3 (coding for the β2-subunit)

Pyruvate kinase catalyzes the reaction from phosphoenolpyruvate to pyruvate (Valentini *et al.,* 2000) which is the main source for the production of precursors for different anabolic pathways. Pyruvate kinases exist in vascular plants in the cytosol and plastids. Those isoenzymes have clearly different physical, immunological and kinetic characteristics (for review see Ambasht and Kayastha, 2002). The plastidial pyruvate kinase influences pathways in the plastids that use phosphoenolpyruvate or pyruvate as substrate like fatty acid biosynthesis (Andre and Benning, 2007). Plastidial pyruvate kinase is a heteromeric complex consisting of equal proportions of alpha and beta subunits (β1 and β2). The α subunit as well as one of the β subunits are required for enzyme activity in *A. thaliana* (Andre *et al.,* 2007). The plastidial enzyme prevalent in developing seeds in *B. napus* and *Ricinus communis* likely has a subunit composition of 3α3β (Plaxton *et al.,* 2002; Negm *et al.,* 1995), in *Arabidopsis thaliana*, 4α4β1 (Andre *et al.,* 2007). Studies with mutants in plastidial pyruvate kinase of *Arabidopsis thaliana* (Baud *et al.,* 2007b; Andre *et al.,* 2007; Lonien and Schwender, 2009; Hajduch *et al.,* 2010) showed that especially the α and β1 subunits are very important for fatty acid synthesis and embryo development. An inactivation of the α-β1 complex leads to wrinkled seeds with a 60 – 70% reduction in seed fatty acid content (Andre *et al.,* 2007; Baud *et al.,* 2007b). Transgenic expression of the β1 subunit results in a full restoration of oil content and transgenic expression of the β2 subunit restores oil content partially (Andre *et al.,* 2007). Also studies on the expression of genes involved in fatty acid metabolism in *Arabidopsis thaliana* imply that the pyruvate kinase β1 subunit is predominantly utilized for fatty acid biosynthesis (Mentzen *et al.,* 2008). Studies in developing seeds and seedlings

of *Arabidopsis thaliana* (Andre *et al.,* 2007; Andre and Benning, 2007) showed that cytosolic glycolysis is unable to compensate for the loss of plastidial pyruvate kinase activity. Flux analysis in *Arabidopsis* embryos revealed that the flux through plastidial pyruvate kinase, which provides the most of the pyruvate for fatty acid biosynthesis, is reduced in *PKβ1PKα*-mutants by 43% with respect to wild type *Arabidopsis thaliana* (Lonien and Schwender, 2009). Flux analysis in developing *Brassica napus* embryos revealed that about 75% of the pyruvate utilized for fatty acid biosynthesis is produced by plastidial pyruvate kinase (Schwender *et al.,* 2006). This leads to the conclusion that plastidial pyruvate kinase catalyzes a key role in the conversion of photosynthesis products into oil through acetyl-CoA in rapeseed.

I.5.2 Genes coding for pyruvate dehydrogenase of the plastidial pyruvate dehydrogenase complex (PDH)

The pyruvate dehydrogenase complex (PDC) is composed of three primary components: pyruvate dehydrogenase, dihydrolipoamide transacetylase and dihydrolipoamide dehydrogenase. The PDC, like the pyruvate kinase, occurs in the cell in two isoforms, one located in the mitochondrial matrix and the other in the plastids (Lernmark and Gardeström, 1994). The PDC catalyzes the oxidative decarboxylation of pyruvate by which acetyl-CoA is formed. Studies with *Arabidopsis thaliana* (Johnston *et al.,* 1997; Mentzen *et al.,* 2008; Ke *et al.,* 2000; Ruuska *et al.,* 2002, Schwender and Ohlrogge, 2002) and *Brassica napus* (Sangwan *et al.,* 1992) indicate that pyruvate is the main precursor of acetyl-CoA in plastids for fatty acid synthesis through PDC.

I.5.3 Gene coding for pyruvate dehydrogenase kinase (PDHK)

Pyruvate dehydrogenase kinase is the negative regulator of the mitochondrial pyruvate dehydrogenase complex (Budde *et al.,* 1988; Tovar-Mendez *et al.*, 2003).

Mature seeds from transgenic *Arabidopsis thaliana* lines, repressing PDHK, showed a significant increase in the accumulation of seed storage lipids per 100 seeds and an increase in the average 1000-kernel seed weight (Marillia *et al.*, 2003). Li *et al.* (2011) showed that overexpression of a PDHK, detected in *Brassica napus* (*BnPDK1),* resulted in the decrease of seed oil content in transgenic *Arabidopsis* lines. This leads to the conclusion that PDHK, controlling mitochondrial localized PDC, plays a role in the fatty acid biosynthesis in developing seeds. An observed decrease in leaf photosynthesis, while overexpressing *BnPDK,* leads to the conclusion that a possible way to influence seed oil content by PDHK could be through photosynthetic activity (Li *et al.,* 2011). Firming this theory, several studies reported on the importance of photosynthesis for oil accumulation in green seeds. The light reactions of photosynthesis in green seeds can provide ATP and reductant, and the Rubisco bypass allows more efficient lipid synthesis (Ruuska *et al.,* 2004; Schwender *et al*., 2003 and 2004; Goffman *et al.*, 2005; Li *et al.,* 2006).

I.5.4 Genes coding for biotin carboxyl carrier protein (BCCP2), alpha-carboxyltranserfase (CAC3) and biotin carboxylase (CAC2) of the heteromeric isoform of the acetyl-coenzyme A carboxylase

Dicotyledonous plants have two isoforms of acetyl-coenzyme A carboxylase (ACCase): a homomeric isoform in the cytosol and a heteromeric isoform in the plastids. They have several important differences in their biochemical properties and functions. It is thought that the homomeric isoform in the cytosol is needed for a number of different classes of specialized metabolites like the biosynthesis of flavonoids and the elongation of oleic acid to erucic acid, which is a major malonyl-CoA dependent pathway in *Brassica* species (for review see Ohlrogge and Browse, 1995). Plastids from *Brassicaceae* plants contain a homomeric and a heteromeric isoform of acetyl-coenzyme A carboxylase. The heteromeric isoform (ACCase II)

catalyzes the formation of malonyl-coenzyme A which is a key regulatory step for de novo fatty acid biosynthesis (Post-Breittenmiller *et al.,* 1992; Jaworski *et al.,* 1993; for review see Thelen and Ohlrogge, 2002a and 2002b). The heteromeric acetyl-coenzyme A carboxylase consists of four subunits: alpha - carboxyltransferase (CAC3), biotin carboxylase (CAC2), biotin carboxyl carrier protein (BCCP) and beta - carboxyltransferase. The latter one is encoded by the plastidial genome, while the others are nuclear encoded (Sasaki *et al.,* 1993 and 1995). In *Arabidopsis* two paralogous genes exist for BCCP, *CAC1A* (*At5g16390*, codes for BCCP1) and *CAC2A* (*At5g15530*, codes for BCCP2) (Thelen *et al*., 2001), while oilseed rape contains at least 6 *BCCP* copies of which four encode class 2 BCCPs (BCCP2) (Elborough *et al.,* 1996). This gene complexity, and the observations that all BCCP isoforms from *Arabidopsi*s and oilseed rape are up-regulated during the period of maximal oil accumulation in seeds, point to a potential importance of BCCPs in ACCase function and fatty acid biosynthesis (Turnham and Northcote, 1983; Focks and Benning, 1998). Thelen *et al.* (2001) showed in semiquantative immunoblot analyses that class one BCCPs are strongly expressed in most plant organs of *Arabidopsis* and oilseed rape, while class two BCCPs accumulate predominantly in flowers and developing seeds. This led to the hypothesis that BCCP1 might be important for housekeeping functions and BCCP2 might have a role in fatty acid biosynthesis for lipid deposition (Thelen *et al.,* 2001).

I.5.5 Gene coding for the short chain 3-ketoacyl carrier protein synthase (KAS III)

KAS III catalyzes the condensation reaction of acetyl-CoA and malonyl-ACP which initiates the fatty acid synthesis in plants (Jaworski *et al.,* 1989 and 1993). Further condensation reactions, which elongate the nascent fatty acid molecule, are

catalyzed by KAS I, II and IV by utilizing acyl-ACPs and malonyl-ACP. KAS I catalyzes the elongations from butyryl-ACP to palmitic acid (C4:0-ACP-C16:0-ACP), while KAS II is specific for the further elongation to stearic acid (C16:0-ACP to C18:00-ACP). KAS IV is most active on medium-chain acyl-substrates (C4-C14) (for review see Dehesh *et al.,* 2001). KAS III is thought to have a rate-limiting role in fatty acid biosynthesis because its enzyme activity correlates well with the rate of fatty acid synthesis in vitro (Jaworski *et al.,* 1989). Abbadi *et al.* (2000) developed an unregulated KAS III enzyme (KAS III is normally inhibited by its medium-chain-acyl ACP products) which produced in vitro a 1.5 fold increase in medium-chain fatty acids and an increase of 20% in total fatty acid products in comparison to the wild type enzyme. Dehesh *et al.* (2001) and Stoll *et al.* (2006) detected an increase in palmitic acid in transgenic plants of *Arabidopsis*, rapeseed and tobacco, but unfortunately the transgenic seeds contained lower levels of oil as well as reduced rates of the lipid biosynthesis compared to the wild type seeds (Dehesh *et al.,* 2001).

I.5.6 Transcription factors: Leafy cotelydon 2 (LEC2), FUSCA3 (FUS3) and wrinkled (WRI)

The transcription factors leafy cotelydon 1 (LEC1), leafy cotelydon 2 (LEC2), abscisic acid insensitive 3 (ABI3) and FUSCA3 (FUS3) have been identified as master regulators of seed maturation processes in *Arabidopsis thaliana* (Kroj *et al.,* 2003; Crowe *et al.,* 2000; Santos-Mendoza *et al.,* 2005 and 2008; To *et al.,* 2006; Baud *et al.,* 2007a; Kagaya *et al.,* 2005; Mu *et al.,* 2008; Stone *et al.,* 2008; Roschzttardtz *et al.,* 2009). These master regulators exhibit a broad control on seed maturation by acting together in a complex regulatory network, where LEC1 and LEC2 act as positive regulators upstream of ABI3 and FUS3 (Kroj *et al.,* 2003; Crowe *et al.,* 2000; Santos-Mendoza *et al.,* 2005 and 2008; To *et al.,* 2006; Baud *et al.,* 2007a; Kagaya

et al., 2005; Mu *et al.,* 2008). Whether the regulation by LEC2 is direct or indirect is not known, but experiments of Braybrook *et al.* (2006) indicate that ABI3 and FUS3 may not be direct targets. *LEC1* encodes a CBF transcription factor (Lotan *et al.,* 1998). The proteins of *LEC2*, *FUS3* and *ABI3* are related and all of them contain B3 DNA-binding domains. They could directly bind the promoter of maturation specific genes (Giraudat *et al.,* 1992; Luerssen *et al.,* 1998; Stone *et al.,* 2001). LEC1, LEC2 and FUS3 (Yamamoto *et al.,* 2010) also regulate, possibly in different parallel or partially redundant pathways, the transcription factor wrinkled (*WRI*) which in turn controls the expression of genes that are involved in late glycolysis and fatty acid biosynthesis, as well as biotin and lipoid acid synthesis (Ruuska *et al.,* 2002; To *et al.,* 2006; Andre *et al.,* 2007; Baud *et al.,* 2007a and 2007b; Mu *et al.,* 2008, Baud *et al.,* 2009; Lonien and Schwender, 2009; Yamamoto *et al.,* 2010). *WRI* was characterized as coding for an APETALA2/EREB (ethylene responsive element binding) domain protein that is involved in the control of storage compound biosynthesis in *Arabidopsis thaliana* (Cernac and Benning, 2004). The *wri1*-mutants showed a decreased conversion of sucrose and glucose into triacylglycerols. The seed oil content was reduced up to 80%. This was most probably caused by reduced activities of key glycolytic enzymes (Focks and Benning, 1998; Baud and Graham, 2006; Lonien and Schwender, 2009). An overexpression of this transcription factor gene in transgenic lines resulted in an increased level of triacylglycerols in seeds and leaves (Cernac and Benning, 2004). Masaki *et al.* (2005) and Cernac *et al.* (2006) even demonstrated in their studies that a regulatory cross-talk between *WRI* and sucrose or abscisic acid exists. The target genes of *WRI* involved in late glycolysis and fatty acid biosynthesis are (among others) *PKP1*, *PKP2*, *PDH E1 α*, *PDH E1 β*, *BCCP2* and *KAS I* (Santos-Mendoza *et al.,* 2008; Baud *et al.,* 2007a and 2009; Maeo *et al.,* 2009; Lonien and Schwender, 2009).

I.5.7 Gene involved in seed development and regulation of developmental gene expression: polycomb-group protein fertilization-independent endosperm (FIE)

All putative polycomb group protein complexes (PcG-complexes) in *Arabidopsis thaliana* are predicted to comprise at least the WD40 motif containing proteins fertilization-independent endosperm (FIE) and multicopysupressor of IRA1 (MSI1) as basis, and a zinc-finger protein and a SET-domain protein as further subunits (Köhler *et al.,* 2003a; Jullien *et al.,* 2006). The SET-domain protein is thought to be the catalytic subunit (Mosquna *et al.,* 2009). The zinc-finger protein contains a VEFS domain which may interact with a SET-domain protein (Chanvivattana *et al.,* 2004; Jullien *et al.,* 2006). These subunits are variable depending on the cell type and function (Kinoshita *et al.,* 2001; Katz *et al.,* 2004; Wood *et al.,* 2006; Jullien *et al.,* 2006; Jiang *et al.,* 2008). Possible PcG-subunits of the zinc-finger protein group are embryonic flower 2 (EMF2), vernalization 2 (VRN2) or fertilization-independent seed 2 (FIS2). The SET-domain protein subunits are curley leaf (CLF), swinger (SWN) or MEDEA (MEA) (Chaudhury *et al.,* 1997; Goodrich *et al.,* 1997; Luo *et al.,* 1999; Spillane *et al.,* 2000; Yadegari *et al.,* 2000; Köhler *et al.,* 2003b; Katz *et al.,* 2004; Chanvivattana *et al.,* 2004; Wood *et al.,* 2006; Jullien *et al.,* 2006; Jiang *et al.,* 2008). Polycomb group proteins play critical roles in the regulation of developmental gene expression (Jiang *et al.,* 2008). They regulate vegetative and reproductive programs in flowering plants (Ohad *et al.,* 1996 and 1999, Chaudhury *et al.,* 1997; Yadegari *et al.,* 2000; Kinoshita *et al.,* 2001; Katz *et al.,* 2004; Wood *et al.,* 2006; Jullien *et al.,* 2006; Bouyer *et al.,* 2011). The PcG-complex, built of the “Fis-class” genes, regulates endosperm and embryo development in *Arabidopsis* (Ohad *et al.,* 1999; Luo *et al.,* 1999 and 2000; Spillane *et al.,* 2000; Yadegari *et al.,* 2000; Sørensen *et al.,* 2001; Köhler *et al.,* 2003a and 2003b; Bouyer *et al.,* 2011). The FIS

class genes consist of *FIE*, *MSI1*, *FIS2* and *MEA* (Ohad *et al.,* 1999; Chaudhury *et al.,* 1997; Grossniklaus *et al.,* 1998; Kiyosue *et al.,* 1999; Spillane *et al.,* 2000; Köhler *et al.,* 2003a and 2003b). These genes prevent endosperm development before double fertilization (Ohad *et al.,* 1999; Luo *et al.,* 2000). Mutants in the Polycomb Group genes of the "FIS class" show phenotypically similarities of an increased parental dosage in the endosperm which leads to an increase of seed size (Scott *et al.,* 1998). Those mutants showed autonomous initiation of endosperm development and ovule growth (Ohad *et al.,* 1996 and 1999; Chaudhury *et al.,* 1997; Grossniklaus *et al.,* 1998; Luo *et al.,* 1999 and 2000). This is of interest because bigger seeds provide more space for storage components like proteins and oil, as demonstrated by Zhou *et al.* (2009). The PcG-complexes also repress the activity of seed developmental genes (Luo *et al.,* 1999; Köhler *et al.,* 2003b; Bouyer *et al.,* 2011). Bouyer *et al.* (2011) analyzed the genomic and developmental consequences of the complete loss of the PcG-activity during embryo and subsequent sporophyte development with homozygous *fie*-mutant plants. Genes acting at different hierarchical levels were up-regulated in the plants with loss of FIE-function. For example, the transcriptional master regulators *LEC2* and *FUS3* as well as the more specific transcription factor *WRI* and downstream targets like oleosine genes were up-regulated. The oleosine genes were strongly overrepresented with 73% of the PcG-targeted oleosine genes being up-regulated. This was also confirmed by the observation of strong lipid accumulation in cotyledons and roots in *fie*-mutant seedlings.

I.5.8 Gene involved in seed development: *IKU2*

IKU2 encodes a leucine-rich repeat kinase which is expressed uniquely in the endosperm (Luo *et al.,* 2005). Garcia *et al.* (2003) isolated the two single loci mutants *IKU1* (*haiku1*) and *IKU2* (*haiku2*) in *Arabidopsis thaliana* plants. These mutations specifically affect seed size. Mutations in *IKU* lead to a premature stop of endosperm growth which triggers precocious cellularization, reduction of cell proliferation in the embryo and limitation of cell elongation of the maternally derived seed integument (Garcia *et al.,* 2003). Plants homozygous for *iku* produce smaller seeds (they phenocopy maternal excess) but do not show any vegetative or reproductive phenotypic variation (Garcia *et al.,* 2003). The *IKU2* locus colocalizes with a quantitative trait locus for seed size that had been detected in a study of Alonso-Blanco *et al.* (1999). As already mentioned for *FIE*, the size of the seed is of interest because of the space for storage components.

I.6 The main objectives of this study were:

(i) to analyze the allelic diversity of candidate genes for oil content.

For this approach the allelic diversity was evaluated in exotic genotypes represented by Chinese, spring and resynthesized rapeseed in comparison to reference sequences from the winter rapeseed varieties “Express” or “Tapidor” on the DNA sequence level for the candidate genes *PKP1*, *PKP2*, *PKP3*, *LEC2*, *FUS3*, *FIE*, *IKU2, WRI*, *BCCP2*, *PDH* and *PDHK*.

(ii) to detect novel alleles of the candidate genes in the exotic genotypes.

For this approach, the polymorphisms and haplotypes detected in the allelic diversity study were compared to the sequences of a set of current winter rapeseed breeding materials.

(iii) to evaluate effects on oil content of novel alleles in candidate genes for oil content.

In this approach, novel alleles detected in the candidate genes *PKP1*, *PKP2*, *PKP3*, *LEC2*, *IKU2*, *WRI*, *BCCP2*, *CAC2*, *CAC3*, *KAS III*, *PDH* and *PDHK* in genotypes analyzed in the earlier GABI BRIDGE project and the exotic plant materials analyzed in Chapter II, were tested for effects on oil content in segregating F_2 populations.

I.7 Utilization of data from the earlier GABI BRIDGE project and partners in GABI OIL

This thesis was a subproject of the project GABI FUTURE “OIL - OMICS based strategies for increasing the seed oil content in rapeseed" (GABI OIL). It was conducted in close collaboration with other partners within GABI OIL. It also was a follow-up on the project GABI BRIDGE “*Brassica napus*: Allelic diversity in Candidate Genes” (GABI BRIDGE), from which data was used.

- Sequences of the conventional winter rapeseed materials were provided by Wolfgang Ecke, Martin Lange (Georg-August-Universität in Göttingen, Germany) and Renate Schmidt´s groups at the Max-Planck-Institut für molekulare Pflanzenphysiologie (MPI) in Golm, Germany, and the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) in Gatersleben, Germany.
- The reference sequences as well as locus specific primer pairs were provided by Renate Schmidt’s groups.
- DNA sequencing in the GABI OIL project was done by the sequencing facility of Syngenta Seeds GmbH: Syngenta Biotechnology Inc. (NC, USA).

- Allelic diversity in candidate genes for oil content in the “core set” of GABI BRIDGE had been evaluated by Martin Lange (Georg-August-Universität Göttingen, Germany) and Renate Schmidt´s group at the MPI in Golm, Germany. On this basis, genotypes as donor of novel alleles have been selected for the F_2 populations investigated in the field trail 2007/08.
- Segregating F_2 populations were grown at the nurseries of the Georg-August-Universität Göttingen (Göttingen Reinshof, Germany), Deutsche Saatveredelung AG (Thüle, Germany), Syngenta Seeds GmbH (Biemsen, Germany), KWS SAAT AG (Einbeck, Germany) and the Justus-Liebig-Universität Gießen (Rauischholzhausen, Germany).
- Saaten-Union Biotec GmbH (Gatersleben, Germany) and KWS SAAT AG (Einbeck, Germany) developed markers for the novel alleles. Also, marker analyses in the F_2 populations were carried out by these companies.

II. Analyzing the diversity of candidate gene loci for oil content in exotic rapeseed materials

II.1 Materials and Methods

II.1.1 Plant materials

In this study a total of 32 genotypes were included in the exotic genotype set (Table II.1 & Table II.2). Six Chinese rapeseeds, six spring rapeseed varieties and 19 resynthesized rapeseeds were selected from the materials of Girke (2002). One additional resynthesized rapeseed was chosen from the materials of Jesske (2011), which was resynthesized from "Yellow Sarson" (*Brassica rapa var. trilocularis, A-genome)* and the wild species *Brassica montana (C-genome)*. Seeds were manually sown in multi pot trays and grown in the greenhouse under standard conditions at the Department of Crop Sciences, Georg-August-Universität Göttingen, Germany.

Table II.1: Designation, origin and ancestry of the resynthesized genotypes in the set of 32 exotic rapeseed genotypes (Girke 2002; Jesske 2011)

ID	genotype	source[§]	mother plant	form/ variety	father plant	form/variety
7	G39*	GAU	*B. oleracea conv. capitata var. capitata*	Sun up H	*B. rapa ssp. oleifera*	Yellow seeds
4	G43	GAU	*B. oleracea conv. capitata var. capitata*	S-S Cross	*B. rapa ssp. chinensis*	forma
3	G50	GAU	*B. oleracea conv. acephala var. gongyloides*	Erfordia	*B. rapa ssp. oleifera*	Yellow seeds
15	H128	GAU	*B. oleracea conv. acephala var. sabellica*	Halbhoher Grüner Mooskrauser	*B. rapa ssp. chinensis*	forma
8	H149	GAU	*B. oleracea conv. capitata var. sabauda*	Cavalier Rouge	*B. rapa ssp. chinensis*	forma
18	H365	GAU	*B. oleracea conv. capitata var. capitata*	Ho41F	*B. rapa ssp. rapa*	Just Right
13	H40*	GAU	*B. oleracea conv. botrytis var. italica*	Spartan Early	*B. rapa ssp. pekinensis var. laxa*	Sandun
2	H44	GAU	*B. oleracea conv. capitata var. sabauda*	Eisenkopf	*B. rapa ssp. pekinensis*	Ho46
9	K29*	GAU	*B. oleracea conv. acephala var. sabellica*	Halbhoher Grüner Mooskrauser	*B. rapa ssp. oleifera*	Yellow seeds

* genotypes of the screening set, [§]GAU: Georg-August-Universität Göttingen, FUB: Freie Universität Berlin

Table II.1: Designation, origin and ancestry of the resynthesized genotypes in the set of 32 exotic rapeseed genotypes (Girke 2002; Jesske 2011) (continued)

ID	genotype	source[§]	mother plant	form/ variety	father plant	form/variety
11	K332	GAU	*B. rapa ssp. oleifera*	forma annua	*B. napus ssp. napobrassica*	G50
16	L122	GAU	*B. oleracea conv. capitata var. sabauda x B. oleracea conv. capitata var. sabauda*	Praeco x Savoy King	*B. rapa ssp. pekinensis*	Early Hybrid G
32	MOY4	GAU	*B. rapa ssp. trilocularis*	Yellow Sarson (accession number RO18)	*B. montana*	wild species (accession number 6835)
12	R140	GAU	*B. oleracea conv. capitata var. capitata*	Hybrid 0	*B. rapa ssp. oleifera*	Eskisehir II
19	R54	GAU	*B. oleracea conv. capitata var. capitata*	Stone Head	*B. rapa ssp. pekinensis*	forma
5	R76	GAU	*B. oleracea conv. botrytis var. albograbra*	966-1-4 SI	*B. rapa ssp. oleifera*	Eskisehir II
14	R8	GAU	*B. oleracea conv. capitata var. sabauda*	Eisenkopf	*B. rapa ssp. pekinensis*	Ho46
1	R99	GAU	*B. oleracea conv. capitata var. capitata*	Stone Head	*B. rapa ssp. pekinensis*	forma
10	RS1/2*	FUB	*B. rapa*	I 85-326	*B. oleracea*	I 85-241
6	S17	FUB	*B. napus ssp. napus x B. oleracea conv. gemifera*	rapeseed x Brussel sprout	*B. rapa ssp. oleifera*	Lemkes Rübsen
17	S29	GAU	*B. oleracea conv. acephala var. sabellica*	Japanese ornamental cabbage	*B. rapa ssp. pekinensis*	Sandun

* genotypes of the screening set, [§]GAU: Georg-August-Universität Göttingen, FUB: Freie Universität Berlin

Table II.2: Designation, seed quality and origin of the Chinese and spring rapeseed genotypes in the set of 32 exotic genotypes (Girke 2002)

ID	genotype	gene pool	quality[$]	breeding company	country	first registration
23	87-50182*	Chinese rapeseed	++	Wanxian Institute of Agricultural Science, Sichuan Provine	China	
29	Altex*	Spring rapeseed	00	University of Alberta, Dept of Plant Science, Edmonton, Alberta T6G 2P5	Canada	1980
27	Barossa	Spring rapeseed	00	New South Wales Dept of Agriculture, Agricultural Research Institute, Wagga Wagga, New South Wales 2650	Australia	1990
22	Ganyu 3	Chinese rapeseed	++	Huazong Agricultural University, Wuhan, Hubei Province	China	1977
20	Italy	Chinese rapeseed	++	Huazong Agricultural University, Wuhan, Hubei Province	China	
25	Linyou 5*	Chinese rapeseed	++	Jiangshu Academy of Agricultural Sciences, Nanjing, Jiangshu Province	China	
28	Oro*	Spring rapeseed	0+	Agriculture Canada, 107 Science Place, Saskatoon, Saskatchewan S7N 0X2	Canada	1968
31	Roy 12	Spring rapeseed	0+	Western Australian Dept of Agriculture, Baron-Hay Court, South Perth, Western Australia 6151	Australia	
30	Shiralee	Spring rapeseed	00	New South Wales Dept of Agriculture, Agricultural Research Institute, Wagga Wagga, New South Wales 2650	Australia	1988
26	Tanto	Spring rapeseed	00	INRA, Station d'Amelioration des Plantes, BP No 29, 3560 Le Rheu/ Serasem – Reserches et Selections, Ferme de la Cueillerie, 10/12, rue Roger-Lecerf, Premesques; 59840 Perchenies	France	1990
21	Xiangyou 11	Chinese rapeseed	00	Huan Academy of Agricultural Sciences, Changsha, Huan Province	China	
24	Zhenyou 11	Chinese rapeseed	++	Zhejiang Agriculture University	China	

* genotypes of the screening set, [$]seed quality with zero (0) or high (+) erucic and low (0) or high (+) glucosinolate content

II.1.2 Selected candidate genes for oil content and locus specific primer pairs

The candidate gene loci investigated in this study as well as the functions of the encoded proteins are shown in Table II.3. Renate Schmidt´s groups provided locus specific primer pairs on the basis of sequences from a BAC-library of the rapeseed cultivar “Express” for 1 to 5 fragments per candidate gene locus. Only for the candidate gene PDHK, a BAC-library of the rapeseed cultivar “Tapidor” was used (Bach, 2007; Schmidt, personal communication). Those locus specific primer pairs were necessary for analyzing the allelic diversity at individual loci since in the polyploid rapeseed genome each gene exists in two or more copies.

Table II.3: Selected candidate genes for oil content

Gene abbreviation	Gene name	Loci	Function of the protein
PKP2	Plastidial pyruvate kinase 2	*K14, K48, K141*	Plastidial pyruvate kinase subunit β1: Catalyze the reaction from phosphoenolpyruvate to pyruvate
PKP3	Plastidial pyruvate kinase 3	*123P16, 123J24*	Plastidial pyruvate kinase subunit β2: Catalyze the reaction from phosphoenolpyruvate to pyruvate
PKP1	Plastidial pyruvate kinase 1	*PPK17, PPK196, PPK214*	Plastidial pyruvate kinase subunit α: Catalyze the reaction from phosphoenolpyruvate to pyruvate
LEC2	Leafy cotelydon 2	*L65, L83*	Transcription factor: Control of seed maturation and regulation of developmental gene expression
FUS3	FUSCA 3	*F161, F183*	Transcription factor: Control of seed maturation and regulation of developmental gene expression
FIE	Fertilization-independent endosperm	*FIE67*	Transcription factor: Control of seed development and regulation of developmental gene expression
IKU2	IKU2	*I103, I124*	Leucine rich repeat kinase: Control of seed size
WRI	Wrinkled	*W2, W9, W80, W102*	Transcription factor: Control of storage compound biosynthesis
BCCP2	Biotin carboxyl carrier protein 2	*B51, B62, B114, B173*	Biotin carboxyl carrier protein: Catalyze the reaction from Acetyl-CoA to Malonyl-CoA

Table II.3: Selected candidate genes for oil content (continued)

Gene abbreviation	Gene name	Loci	Function of the protein
PDH	Pyruvate dehydrogenase E1 α	*H25, H29, H40, H71, H81*	Pyruvate dehydrogenase: Catalyze the reaction from pyruvate to Acetyl-CoA
PDHK	Pyruvate dehydrogenase kinase	*P12, P33, P57, P133*	Pyruvate dehydrogenase kinase: Inhibition of PDH-complex by phosphorylation

Some of the locus specific primer pairs used in this study already have been published by Cao and Schmidt (2013). Due to the allocation table Table A.1 in the appendix it is possible to match the primers published by Cao and Schmidt (2013) to the analyzed fragments in this study. All primers were synthesized by Eurofins MWG operon (Ebersberg, Germany). The number of fragments per candidate gene locus is shown in Table II.5.

II.1.3 Reference sequence and intron/ exon structure

For identifying polymorphisms, the full length sequences of the cultivar “Express” for the different loci were used as reference for all candidate genes except for *PDHK*. For the candidate gene *PDHK*, the full length sequences of the cultivar “Tapidor” were used as reference. This sequence information was provided by Renate Schmidt´s groups (MPI Glom, IPK Gatersleben). For the candidate genes *BCCP2*, *PDH*, *PDHK*, *PKP2* and *WRI* Renate Schmidt´s groups also provided information about intron and exon structure, also derived from the full length sequences of the loci from the BAC-libraries of “Express” and “Tapidor”, respectively. For these five candidate genes, the sequences of the open reading frame as well as the full length sequences of the loci from the reference sequences are published in the PhD thesis of Bach (2007). In the allocation table Table A.2 in the appendix, the corresponding candidate gene is shown in comparison to the label of the sequences and open

reading frames published by Bach (2007). The full length sequence information of the other investigated candidate genes have been personally communicated by Renate Schmidt.

II.1.4 Amplicons and genotypes of the conventional winter rapeseed materials

To check for novel alleles, the sequences of the exotic rapeseed materials, shown in Table II.1 and Table II.2, were compared with the sequences of conventional breeding materials, shown in Table II.4. The conventional breeding set consists of 20 winter rapeseed and four forage rape varieties and breeding lines. Sequence information of the conventional breeding materials was provided by Renate Schmidt (personal communication) and Martin Lange (personal communication). The corresponding amplicons are listed in Table II.5.

Table II.4: Genotypes of winter rapeseed varieties and breeding lines used as reference for identifying novel alleles (conventional winter rapeseed materials)

ID	genotype	gene pool	quality
52	Amor	Winter rapeseed	00
39*	Askari	Winter rapeseed	++
27	Aviso	Winter rapeseed	00
50	Capitol	Winter rapeseed	00
70	Contact	Winter rapeseed	00
80	DH1 AxM	Winter rapeseed	00
79*	DH7 AxM	Winter rapeseed	00
53	Duell	Winter rapeseed	00
34*	Erox	Winter rapeseed	+0
2	Express	Winter rapeseed	00
67*	Lipid	Winter rapeseed	00
4*	Mansholt	Winter rapeseed	++
29	Milena	Winter rapeseed	00
28	Musette	Winter rapeseed	00
26	Orlando	Winter rapeseed	00
30	Pollen	Winter rapeseed	00
25	Prince	Winter rapeseed	00
5*	Samourai	Winter rapeseed	00

* Screening set

Table II.4: Genotypes of winter rapeseed varieties and breeding lines used as reference for identifying novel alleles (conventional winter rapeseed materials) (continued)

ID	genotype	gene pool	quality
6	Sollux	Winter rapeseed	++
32*	Wotan	Winter rapeseed	00
20	Akela	Forage rape	++
19*	Campari	Forage rape	00
22	Caramba	Forage rape	00
24	Licapo	Forage rape	00

***** Screening set

Table II.5: Candidate gene loci and amplicons for which sequence information was available from the conventional winter rapeseed materials

candidate gene, locus	analyzed fragments	candidate gene, locus	analyzed fragments
BCCP2, B51	1*, 2, 3	*PDHK, P12*	1*, 2, 3
BCCP2, B62	1*, 2, 3	*PDHK, P33*	1, 2, 3*
BCCP2, B114	1, 2*, 3	*PDHK, P57*	1, 2, 3*
BCCP2, B173	1, 2, 3*	*PDHK, P133*	1, 2, 3*
WRI, W2	1, 2, 3, 4, 5	*PDH, H25*	1, 2, 3*
WRI, W9	1, 3, 4, 5	*PDH, H29*	1*, 2, 3
WRI, W80	1, 3*, 4, 5	*PDH, H40*	1*, 2, 3
WRI, W102	1, 2, 3, 5*	*PDH, H71*	1, 2*, 3
PKP2, K14	1, 2, 3*	*PDH, H81*	1, 2, 3*
PKP2, K48	1, 2, 3, 4*	*LEC2, L65*	1, 2*, 3, 4
PKP2, K141	1*	*LEC2, L83*	1*, 2, 2a, 3a
PKP3, 123P16	1*, 2, 3	*FUS3, F161*	1*, 2
PKP3, 123J24	1, 2, 3	*FUS3, F183*	1b*, 2b, 3
PKP1, PPK17	1*, 2*, 3	*FIE, FIE67*	1, 2, 3b, 4b*
PKP1, PPK196	1, 2*, 3, 4	*IKU2, IKU103*	1b, 2, 3, 4*, 5
PKP1, PPK214	1, 2, 3, 4, 5*	*IKU2, IKU124*	1*, 2, 3b, 4, 5b

* the whole set of the conventional winter rapeseed materials was analyzed for this fragment

II.1.5 DNA extraction

Leave samples were harvested from four to six week old seedlings and were immediately frozen in liquid nitrogen. Total genomic DNA was isolated from 100 mg frozen leaf materials, using Nucleon®PhytoPure® genomic DNA Extraction Kits from

GE Healthcare (Freiburg, Germany) following the manufacturer's instructions. DNA was quantified with a Biorad VersaFluor™ fluorometer (Bio-Rad Laboratories CA, USA) using Bio-Rad Fluorescent DNA Quantification Kits (Bio-Rad Laboratories CA, USA) according to the manufacturer´s manual.

II.1.6 Polymerase chain reaction (PCR)

PCR was carried out in 50 µl reaction volumes containing 20 - 50 ng DNA, 5 µl 10x PCR buffer (0.5 M potassium acetate, 9 mM Tween 20, 15 mM magnesium acetate, 0.5 M TrisHCl (Tris hydroxymethyl aminomethane hydrochloride) pH 7.5), 2.5 mM of each dNTP (Bio-Budget Technologies GmbH Krefeld, Germany), 50 µM of each oligonucleotide primer and 1U HOT FIREPol® DNA polymerase (Solis BioDyne, Tartu, Estonia). PCRs were run in a Biometra T1 Thermocycler (Biometra GmbH, Göttingen, Germany), software version v5.01tl, using the following cycle conditions: denaturation at 95°C for 5 min, 35 cycles of 45 s at 95°C, 45 s at annealing temperatures of 50 – 65°C (depending on the different primer pairs), followed by 1 min of elongation at 72°C and a final extension at 72°C for 10 min. Amplification products were analyzed by gel electrophoresis on 1.3% (w/v) agarose gels in TAE-buffer (40 mM Tris hydroxymethyl aminomethane, 2 mM EDTA, 20 mM acetic acid; pH 8.0) with ethidium bromide staining. As loading buffer, 6X DNA Loading Dye for sample DNA (Fermentas, St. Leon-Rot, Germany) was used. GeneRuler™ 1 kb Plus DNA Ladder (Fermentas, St. Leon-Rot, Germany) was used as molecular weight standard. Amplification products were purified using the High Pure PCR Product Purification Kit™ from Roche (Mannheim, Germany) according to the manufacturer´s instruction.

II.1.7 Sequence analysis

For sequence analysis 2.5 – 10 ng DNA was used. BigDye® terminator cycle sequencing was done on an Applied Biosystem 3730XL sequencer by Syngenta Biotechnology Inc. (NC, USA). CodonCode Aligner (CodonCode Corporation, MA, USA) and DNASTAR Lasergene® v5 (DNASTAR, Inc., Madison Wisconsin, USA) were used to manually edit and analyze the sequences. For multiple alignment of the sequences ClustalW (Thompson *et al.*, 1994) was used. For pairwise alignment, especially the alignment of the complete sequences to the exon regions, the computer algorithm Needleman – Wunsch (Needleman and Wunsch, 1970) was chosen. Polymorphisms were determined by comparing each amplicon with the reference sequence from the rapeseed cultivars “Express” and “Tapidor”, respectively.

II.1.8 Sequencing strategy

For the candidate genes, one to five fragments per locus were initially analyzed by sequencing a screening set of eight exotic genotypes (indicated in Table II.1 & Table II.2). The fragment that showed the most polymorphisms in the coding region was then sequenced in the rest of the 32 exotic genotypes. If no polymorphism was detected in the screening set, no more sequence analyses were done at this locus.

To check if the detected polymorphisms are located in the exon region, information about intron and exon structure provided by Renate Schmidt´s groups (see II 1.3 in the materials section) was used. For the candidate genes *FIE*, *FUS3*, *IKU2*, *LEC2*, *PKP1* and *PKP3* intron/exon structures were determined by comparing the reference sequences of “Express” to the homologous *Arabidopsis* genes, for which this structure is known (*Arabidopsis* Genome Initiative, 2000). The gene number of the corresponding *A. thaliana* gene (The *Arabidopsis* Information Resource,

www.arabidopsis.org) is shown in Table II.6. After a first comparison on DNA sequence level the approximated exon regions in the rapeseed genes were translated and the inferred intron/exon structure was refined by a comparison with the respective *Arabidopsis* protein sequences.

Table II.6: Corresponding *Arabidopsis thaliana* genes for the evaluation of the coding region in candidate gene loci for oil content in rapeseed (www.Arabidopsis.org)

Gene abbreviation	Loci	gene number of the corresponding A. thaliana gene
PKP3	*123P16, 123J24*	*At1g32440*
PKP1	*PPK17, PPK196, PPK214*	*At3g22960*
LEC2	*L65, L83*	*At3g26790*
FUS3	*F161, F183*	*At3g26790*
FIE	*FIE67*	*At3g20740*
IKU2	*I103, I124*	*At3g19700*

If no polymorphism was detected in the coding region, then the fragment with the most polymorphisms was chosen for further investigation.

II.1.9 Evaluation of the degree of polymorphisms at the different loci

The higher the sequence length for which information is available, the more conclusive data can be obtained for polymorphisms frequency at a locus. Therefore the sequence information of the exotic screening set (indicated in Table II.1) for all analyzed fragments has been used. Also the sequence information of the winter rapeseed screening set (indicated in Table II.2) was integrated, to achieve a more meaningful result by increasing the number of analyzed varieties per locus.

To be able to compare the polymorphisms frequency between the different loci, the value “degree of polymorphism” was calculated. The degree of polymorphism was defined as the number of polymorphisms per 1000 base pairs (polym./kb). This was necessary because the final edited sequences that could be evaluated for

polymorphisms in the screening set ranged from 650 bp to 3340 bp depending on the locus. Therefore, the number of polymorphisms is not immediately comparable between different loci. To also the sequence information of the eight winter rapeseed varieties, that were used as screening set) in the conventional winter rapeseed material

II.1.10 Determining novel haplotypes

If a polymorphism has not been detected in the set of conventional winter rapeseed breeding materials, it was defined as novel polymorphism. Haplotypes that showed novel polymorphisms or a new combination of polymorphisms that did not occur in the set of the conventional winter rapeseed materials were defined as novel haplotypes or novel alleles. The sequence information of the exotic rapeseed material of all fragments listed in Table II.7 and the sequence information of the corresponding fragments of the conventional winter rapeseed materials have been integrated in the total analysis of haplotypes. It was assumed that the genotype, for which only sequence information of the selected fragment was available, has the same haplotype as the genotype with the same polymorphisms in the selected fragment, for which additional sequence information was available. Genotypes showing haplotypes not present in the screening set or the conventional winter rapeseed materials were classified as single haplotype without any speculation which polymorphisms could have been in the other fragments. For the candidate genes *FIE*, *FUS3*, *IKU2*, *LEC2*, *PKP1* and *PKP3,* the novel haplotypes have been named in alphabetical order in which they have been detected. The haplotype of the reference sequence of the cultivars “Express” and “Tapidor” was always named “A”. At the other candidate genes, the haplotypes have been named in the same way under consideration of the already used letters for haplotypes detected and named in the

previous GABI BRIDGE project (Lange, personal communication).

II.1.11 Characterization of polymorphisms in the different haplotypes

The polymorphisms in coding region were checked whether they cause no amino acid exchanges (silent polymorphisms), conservative or non-conservative amino acid exchanges (according to Schwartz and Dayhoff, 1978), frame shifts or stop codons. This was done by comparing the amino acid sequences of the reference cultivars “Express” and “Tapidor” to the amino acid sequences of the exotic genotypes and the amplicons of the conventional winter rapeseed materials. The amino acid sequences were retrieved by translating the DNA sequence of the coding regions (DNASTAR lasergene).

The Conserved Domain Database (CDD) of NCBI was used to locate polymorphisms in protein domain regions. Protein domains are defined as functional and/or structural units of a protein. A functional protein domain is determined for example through substitution of amino acids followed by measurements of the activity of the altered protein like the catalytic activity of an enzyme. Structural domains are regions with an independent secondary structure. Functional protein domains and structural domains coincide rather often. Domains are detected in polypeptide sequences on the basis of conserved sequence patterns or motifs (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd_help.shtml). Protein information of the different candidate genes was aligned to the sequences available through CDD, using the CD-Search tool on the NCBI platform (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer *et al.,* 2013). Mapped conserved domains without additional functional information were not taken into account due to the fact that one cannot assume a special effect of amino acid exchanges in these regions.

II.2 Results

II.2.1 Genetic diversity at the different candidate gene loci

As described in the methods section, the degree of polymorphism was evaluated, using the sequence data of the screening set of the winter rapeseed breeding materials and the screening set of the exotic materials. The evaluated degree of polymorphism per locus is listed in Table II.7. Locus *H81* of the *PDH* showed the highest degree of polymorphism of all candidate gene loci with 60 polymorphisms/ kb. At the candidate gene loci, showing a degree of polymorphism of zero, no polymorphism at all was detected in the screening sets.

Table II.7: Degree of polymorphism across both screening sets at the different loci

candidate gene, locus	analyzed fragments	sequence length (bp) of the manually edited sequence	coding region (%)	number of detected polymorphisms	degree of polymorphism (polym./kb)
BCCP2, B51	2, 3	1490	41	7	4.70
BCCP2, B62	1, 2, 3	1800	43	39	21.67
BCCP2, B114	1, 2, 3	2090	36	1	0.48
BCCP2, B173	1, 2, 3	2190	29	37	16.89
WRI, W2	1, 2, 5	1510	60	14	9.27
WRI, W9	1, 4, 5	1994	41	0	0
WRI, W80	1, 4, 5	2080	30	20	9.61
WRI, W102	1, 2, 5	2180	43	8	3.67
PKP2, K14	1, 2, 3	2660	61	8	3.01
PKP2, K48	2, 3, 4	2210	51	7	3.17
PKP2, K141	1	650	0	9	13.85
PKP3, 123P16	1, 2, 3	2610	59	58	22.22
PKP3, 123J24	1, 2, 3	2690	53	0	0
PKP1, PPK17	1, 2, 3	2220	78	18	8.11
PKP1, PPK196	1, 2, 3, 4	2970	63	23	7.74
PKP1, PPK214	1, 2, 3, 4, 5	3340	53	31	9.28

explanation for gene abbreviations see Table II.3; Fragment position (bp) in the reference sequences of “Express” or “Tapidor” see Table A.3 in the appendix

Table II.7: Degree of polymorphism across both screening sets at the different loci (continued)

candidate gene, locus	analyzed fragments	sequence length (bp) of the manually edited sequence	coding region (%)	number of detected polymorphisms	degree of polymorphism (polym./kb)
PDHK, P12	1, 2, 3	2210	49	1	0.45
PDHK, P33	1, 2, 3	1950	60	0	0
PDHK, P57	1, 2, 3	2070	53	18	8.70
PDHK, P133	1, 2, 3	2230	48	37	16.59
PDH, H25	1, 2, 3	2107	54	0	0
PDH, H29	1, 2, 3	1460	51	28	19.18
PDH, H40	1, 2, 3	1990	64	17	8.54
PDH, H71	1, 2, 3	2110	61	0	0
PDH, H81	1, 2, 3	2030	57	60	29.56
LEC2, L65	2, 3, 4	2460	25	10	4.07
LEC2, L83	1, 2a, 3a	2560	43	9	3.52
FUS3, F161	1, 2	1580	51	1	0.63
FUS3, F183	1b, 2b, 3	2030	41	29	14.29
FIE, FIE67	1, 2, 3b	1700	45	1	0.59
IKU2, IKU103	1b, 2, 4	1830	100	22	12.02
IKU2, IKU124	1, 2, 4	1940	98	8	4.12
∑		66941		521	7.78

explanation for gene abbreviations see Table II.3; Fragment position (bp) in the reference sequences of "Express" or "Tapidor" see Table A.3 in the appendix

II.2.2 Polymorphisms in the exotic genotypes in the selected fragments of the different candidate gene loci

At *PDHK P12* and *P33*, *PDH H25* and *H71*, *PKP3 123J24* and *WRI W9* no polymorphisms at all were detected in the exotic screening set. Therefore for these loci no further sequence analyses were done. For the remaining candidate gene loci a fragment was selected as described in the methods and sequenced in the rest of the exotic genotype set. An overview of the selected fragments and the number of detected polymorphisms is shown in Table II.8. This table also gives an overview of the detected polymorphisms in the full set of exotic genotypes. Not for all loci further

polymorphisms have been detected in the whole set of exotic genotypes compared with the screening set. This was true for loci *I124* and *P57*. At the four loci *K141*, *PPK17*, *PPK196* and *FIE67* more polymorphisms were detected, but all outside of the coding region. For all other investigated loci at least one more polymorphism has been detected in the coding region.

All in all, a total of 420 polymorphisms (21.8 polym./kb) were detected in the full set of exotic genotypes in the chosen fragments of the 26 loci in the eleven investigated candidate genes. Of these, a total of 175 polymorphisms (8.48 polym./kb) were detected in the coding region of the different candidate genes. The highest number of polymorphisms (38) per locus in the full set of exotic genotypes was found at locus *K14* for the selected 1084 bp fragment, followed by locus *P133* exhibiting 38 polymorphisms. The highest number of polymorphisms (22) in coding region was detected in the analyzed 985 bp fragment at locus *I103*, where the whole sequenced fragment lies in the coding region of the candidate gene *IKU2*. With four, the lowest number of polymorphisms was detected at locus *FIE67* and locus *K48*.

Table II.8: Polymorphisms in the exotic genotypes in the selected fragments

candidate gene, locus	**fragment**		**detected polymorphisms**	
	no.	**coding region %**	**exotic screening set** (tn/cr)	**full set of exotic genotypes** (tn/cr)
BCCP2, B51	2	62	2/1	6/4
BCCP2, B62	1*	27	15/4	26/5
BCCP2, B114	2*	63	1/0	13/5
BCCP2, B173	2	51	9/6	19/10
WRI, W2	5*	66	7/5	8/6
WRI, W80	5	65	8/4	15/7
WRI, W102	5*	67	5/4	12/9
PKP2, K14	1	72	4/1	47/5
PKP2, K48	2	71	3/3	4/4

***** for this fragment sequence information of the whole conventional winter rapeseed materials was available; tn: total number; cr: in coding region'; explanation for gene abbreviations see Table II.3; Fragment position (bp) in the reference sequences of "Express" or "Tapidor" see Table A.3 in the appendix

Table II.8: Polymorphisms in the exotic genotypes in the selected fragments (continued)

candidate gene, locus	fragment		detected polymorphisms	
	no.	**coding region %**	**exotic screening set** (tn/cr)	**full set of exotic genotypes** (tn/cr)
PKP2, K141	1*	0	7/0	9/0
PKP3, 123P16	1*	76	21/9	24/12
PKP1, PPK17	2*	40	1/0	6/0
PKP1, PPK196	4	86	5/4	6/4
PKP1, PPK214	4	25	27/0	27/0
PDHK, P57	3*	60	16/8	16/8
PDHK, P133	3*	63	32/6	38/11
PDH, H29	1*	56	16/11	23/12
PDH, H40	1*	31	8/3	11/5
PDH, H81	2	80	11/10	28/27
LEC2, L65	3	27	5/1	17/2
LEC2, L83	2a	25	3/0	15/1
FUS3, F161	1*	49	1/0	6/1
FUS3, F183	1b*	45	6/2	10/6
FIE, FIE67	2	53	1/1	4/1
IKU2, I103	1b	100	15/15	22/22
IKU2, I124	1*	100	8/8	8/8
∑			237/106	420/175

***** for this fragment sequence information of the whole conventional winter rapeseed materials was available; tn: total number; cr: in coding region'; explanation for gene abbreviations see Table II.3; Fragment position (bp) in the reference sequences of "Express" or "Tapidor" see Table A.3 in the appendix

II.2.3 Distribution of haplotypes in the candidate gene loci and plant genotypes

Over all candidate gene loci, altogether 149 different haplotypes were found in the exotic genotypes and the reference set of conventional rapeseed lines. Of these, 123 haplotypes were observed in the exotic genotypes, with 70 alleles being specific to these materials (novel alleles). With 52, most of the specific haplotypes were found in the twenty resynthesized rapeseed lines. Nine of those were actually specific to "MOY4", as shown in Table A.4 to Table A.14 in the appendix. The eight spring rapeseed varieties and the eight rapeseed varieties of Chinese origin each showed

four haplotypes specific to their group.

As shown in Table II.9, the highest total number of different haplotypes per locus was found in the candidate gene *IKU2*, locus *I103*, with 9 different haplotypes. This locus also showed the highest number of different haplotypes in exotic genotypes, as well as haplotypes specific for exotic genotypes. Novel haplotypes were not detected at the loci *123J24* of *PKP3*, *W9* of *WRI*, *H25* and *H71* of the *PDH*, and *P33* and *P12* of *PDHK*. In Table A.4 to Table A.14 in the appendix the exact haplotypes per locus and genotype are listed. These tables also show detailed information about the total number and types of polymorphisms per haplotype.

Table II.9: Number of haplotypes in the exotic genotypes and winter rapeseed cultivars

Gene	locus	No. of haplotypes			
		all	in exotic genotypes	specific to exotic genotypes	in winter rapeseed cultivars
PKP2	*K14*	8	6	5	3
	K48	4	3	1	3
	K141	5	4	2	3
PKP3	*123P16*	7	5	4	3
	123J24	1	1	-	1
PKP1	*PPK17*	4	3	1	3
	PPK196	3	3	1	2
	PPK214	6	4	1	5
LEC2	*L65*	7	5	4	3
	L83	5	4	3	2
FUS3	*F161*	4	4	2	2
	F183	6	5	3	3
FIE	*FIE67*	3	3	1	2
IKU2	*I103*	9	9	7	2
	I124	3	3	1	2
WRI	*W2*	6	6	4	1
	W9	1	1	-	1
	W80	4	3	2	3
	W102	8	7	4	4
BCCP2	*B51*	4	3	2	2
	B62	7	7	3	4
	B114	5	5	3	2
	B173	5	4	3	2

explanation for gene abbreviations see Table II.3

Table II.9: Number of haplotypes in the exotic genotypes and winter rapeseed cultivars (continued)

Gene	locus	No. of haplotypes			
		all	in exotic genotypes	specific to exotic genotypes	in winter rapeseed cultivars
PDH	*H25*	1	1	-	1
	H29	5	3	2	3
	H40	7	5	4	3
	H71	1	1	-	1
	H81	7	5	3	4
PDHK	*P12*	3	1	-	3
	P33	1	1	-	1
	P57	5	4	2	3
	P133	4	4	2	2

explanation for gene abbreviations see Table II.3

For the genotypes representing the screening set, all in all 32 loci have been analyzed. The rest of the exotic genotype set has been analyzed only at 26 loci, because in loci *123J24*, *H25*, *H71*, *P12*, *P33*, *W9* no polymorphism was detected at DNA sequence level in the screening set. Due to unsuccessful amplification in the PCR or sequencing errors, at some loci sequence information is not available for all the different genotypes. For the resynthesized genotypes “H44” and “MOY4” even for nineteen loci sequence information could not be obtained. An overview of the number of detected haplotypes per exotic genotype is shown in Table II.10.

Table II.10: Number of haplotypes detected in the exotic genotypes

Exotic genotype	Gene pool	Haplotypes shared with the reference set	Novel haplotypes
R99	Resyn.	23	4
H44	Resyn.	5	8
G50	Resyn.	18	5
G43	Resyn.	20	2
R76	Resyn.	20	4
S17	Resyn.	16	4
G39	Resyn.	20	9
H149	Resyn.	14	8
K29	Resyn.	24	6

Resyn.: resynthesized rapeseed

Table II.10: Number of haplotypes detected in the exotic genotypes (continued)

Exotic genotype	Gene pool	Haplotypes shared with the reference set	Novel haplotypes
RS1/2	Resyn.	21	6
K332	Resyn.	16	6
R140	Resyn.	13	7
H40	Resyn.	26	4
R8	Resyn.	19	4
H128	Resyn.	12	2
L122	Resyn.	19	2
S29	Resyn.	19	6
H365	Resyn.	9	12
R54	Resyn.	17	5
MOY4	Resyn.	0	13
Italy	Chinese	21	3
Xiangyou 11	Chinese	15	5
Ganyu 3	Chinese	15	2
87-50182	Chinese	23	3
Zhenyou 11	Chinese	15	3
Linyou 5	Chinese	22	2
Tanto	Spring	18	3
Barossa	Spring	22	1
Oro	Spring	21	9
Altex	Spring	19	7
Shiralee	Spring	14	4
Roy 12	Spring	15	4

Resyn.: resynthesized rapeseed

For six exotic genotypes eight or more haplotypes that are specific to the exotic genotypes were found. With 13 novel haplotypes, the highest allelic variation was observed in the resynthesized rapeseed “MOY4”. Interestingly, this genotype never shared a haplotype with the current winter rapeseed breeding materials at any locus (see Table A.4 to Table A.14 in the appendix).

II.2.4 Distribution of non-silent polymorphisms in the different haplotypes

An overview of the detected polymorphisms in the different haplotypes per candidate gene locus is presented in Table A.4 to Table A.14 in the appendix. Here, also the corresponding genotypes are listed. The observed polymorphisms between the haplotypes ranged from Insertions/Deletions (InDels) to Single Nucleotide Polymorphisms (SNPs) in introns and exons. Among the SNPs in coding regions some were found to cause non-conservative respectively conservative amino acid exchanges. An overview of amino acid exchanges caused by single nucleotide polymorphisms in coding regions is given in Table II.11. Some of these changes in proteins are located in conserved domains of proteins, which are functional and/or structural units of a protein. The haplotypes carrying SNPs causing amino acid exchanges in conserved protein domain regions are shown in Table II.12.

Table II.11: SNPs causing amino acid exchanges with respect to the reference sequence from ”Express” or “Tapidor” in candidate gene loci for oil content

Gene	locus	haplotype	SNP position[#]	amino acid exchange
PKP2	*K14*	D	1119	glutamate to glutamine (c)
		B, E, D	1314	glutamate to glycine (nc)
	K48	C	3511	proline to leucine (nc)
PKP3	*123P16*	B	615	alanine to serine (c)
		D	653	serine to tyrosine (nc)
		D	713	threonine to methionine (nc)
		G	862	serine to isoleucine (nc)
		D	1010	isoleucine to asparagine (nc)
		G	2801	methionine to isoleucine (c)
PKP1	*PPK17*	D	523+524	serine to phenylalanine (nc)
		D	525+527	threonine to serine (c)
		D	529+530	proline to leucine (nc)
		D	532+533	serine to tyrosine (nc)
		D	535+536	arginine to proline (nc)
		D	538	threonine to isoleucine (nc)
	PPK196	C	558	histidine to asparagine (nc)
		B, C	3344	proline to serine (c)

[#] position in the “Express” or ”Tapidor” reference sequence; explanation for gene abbreviations see Table II.3; (c) conservative amino acid exchange; (nc) non conservative amino acid exchenage

Table II.11: SNPs causing amino acid exchanges with respect to the reference sequence from "Express" or "Tapidor" in candidate gene loci for oil content (continued)

Gene	locus	haplotype	SNP position[#]	amino acid exchange
PKP1	*PPK196*	B, C	3685	lysine to glutamine (nc)
		B	3841	valine to isoleucine (c)
PKP1	*PPK214*	E	553	histidine to proline (nc)
		E	950	glutamate to lysine (nc)
BCCP2	*B51*	C	3259	valine to alanine (nc)
		C	3345	asparagine to aspartate (c)
		C	3550	lysine to glutamine (nc)
		B	3761	glycine to alanine (c)
	B62	E, G, J	814	arginine to proline (nc)
		D, F	826	alanine to valine (nc)
		D	843	aspartate to asparagine (c)
	B114	B, C, G	2577	isoleucine to methionine (c)
		B, C, G	2819	glutamate to alanine (nc)
		B, C	2917	proline to serine (c)
	B173	B	2053	lysine to glutamine (nc)
		C	2384	threonine to alanine (c)
		B, C	2692	serine to alanine (c)
PDH	*H29*	H, C	1555	serine to phenylalanine (nc)
		H, C	1645	phenylalanine to serine (nc)
		H, C	1662	isoleucine to valine (c)
		H,C	1669	serine to tyrosine (nc)
		H	3021	glutamine to glutamate (c)
	H40	H, I, J	688	serine to proline (c)
	H81	C	1866	aspartate to glutamate (c)
PDHK	*P57*	C	4954	phenylalanine to leucine (nc)
		D	4605	asparagine to lysine (nc)
WRI	*W2*	B, D	4275	glutamate to glutamine (c)
		C, D	4334	tryptophan to arginine (nc)
	W80	B	1547	arginine to proline (c)
		B, D	4653	glutamate to glutamine (c)
		B, D	4662	isoleucine to valine (c)
		B	4891	valine to glycine (nc)
		D	4917	threonine to alanine (c)
		D	4990	threonine to isoleucine (nc)
	W102	H, E	6688	alanine to threonine (c)
		E, F, G	6718	phenylalanine to leucine (nc)
		G	6617	alanine to glycine (c)
		H	6446	arginine to histidine (c)
		H	6531	glutamate to arginine (nc)

[#] position in the "Express" or "Tapidor" reference sequence; explanation for gene abbreviations see Table II.3; (c) conservative amino acid exchange; (nc) non conservative amino acid exchenage

Table II.11: SNPs causing amino acid exchanges with respect to the reference sequence from "Express" or "Tapidor" in candidate gene loci for oil content (continued)

Gene	locus	haplotype	SNP position[#]	amino acid exchange
	W102	H	6670	alanine to threonine (c)
IKU2	*I103*	C, D, F	533	glutamine to arginine(nc)
		C, D	605	serine to asparagine (nc)
		B, C, D	608	arginine to glutamine (nc)
		C, D	618	asparagine to lysine (c)
		B	625	aspartate to thyrosine (nc)
		D, E, F, G, H , I	668	asparagine to aspartate (c)
		B	748	arginine to serine (nc)
		B	828	phenylalanine to leucine (nc)
		E, I	842	alanine to valine (nc)
		C	968	glycine to alanine (c)
		B, C, E, I	1378	glutamate to lysine (nc)
		C	1582	methionine to valine (c)
		C	1594	glycine to arginine (nc)
		C	1614	leucine to asparagine (nc)
		E, I	2371	isoleucine to valine (c)
	I124	B	580	alanine to serine (c)
		B	652	aspartate to asparagine (c)
		B	737	threonine to lysine (nc)
		B	782	isoleucine to asparagine (nc)
		C	1157	threonine to asparagine (nc)
		B, C	1318	isoleucine to phenylalanine (nc)
		B	1456	tyrosine to aspartate (nc)
FUS3	*F161*	B, D	1067	methionine to isoleucine (c)
	F183	F	714	proline to serine (c)
		B, C	1013	lysine to arginine (nc)
		B	1907	phenylalanine to leucine (nc)
		B	2053+2054+2055	glutamate to serine (nc)
		B	2131	phenylalanine to tyrosine (c)
FIE	*FIE67*	B, C	1191	valine to glutamate (nc)
LEC2	*L65*	B, C	600	alanine to threonine (c)
	L83	D	692	aspartate to valine (nc)
		B	2484	aspartate to glutamate (c)

[#] position in the "Express" or "Tapidor" reference sequence; explanation for gene abbreviations see Table II.3; (c) conservative amino acid exchange; (nc) non conservative amino acid exchenage

Table II.12: Non-silent polymorphisms in conserved protein domain (cd) regions (critical polymorphisms)

Gene	locus	haplotype	SNP[#]	protein position*	conserved domain (cd)	ID most important mapped cd-sequences
PKP2	*K48*	C	3511	201	Pyruvate kinase, barrel domain	cd00288; pfam00224
PKP3	*123P16*	D	1010	143		
PKP1	*PPK196*	B, C	3685	202	Pyruvate kinase, alpha/beta domain	pfam02887
LEC2	*L65*	B, C	600	243	Plant-specific B3-DNA binding domain	cd10017
	L83	B	2484	196		
FUS3	*F161*	B, D	1067	109		
	F183	B	1907	183		
FIE	*FIE67*	B, C	1191	84	WD40 domain	cd00200
IKU2	*I103*	E,I	842	155	Leucine-rich repeats (LRRs), ribonuclease inhibitor (RI)-like subfamily	cd00116
		C	968	197		
			1378	334		
			1582	402		
			1594	412		
			1614	418		
PDH	*H81*	C	1866	216	Dehydrogenase E1 component	cd02000; pfam00676

[#] position in the "Express" or "Tapidor" reference sequence; explanation for gene abbreviations see Table II.3; (c) conservative amino acid exchange; (nc) non conservative amino acid exchenage

As shown in Table A.4 to Table A.14, haplotypes with InDels in coding regions were found at six different candidate gene loci. A deletion (position 572 bp - 574 bp in the "Express" reference) in the coding region at locus *H40* of candidate gene *PDH* causes a deletion of serine in the amino acid sequence (haplotypes "B","D","I","J"). An insertion observed at this locus (position 688 bp in the "Express" reference) causes a stop codon in consequence of a frame shift (haplotype "F"). The second locus for which a frame shift was detected is locus *PPK17* of the candidate gene *PKP1* (haplotype "D", position 542 bp in the "Express" reference: deletion of cysteine). An insertion of the nucleotides "cgc" at locus *B173* of the candidate gene *BCCP2* (position 2694 bp in the "Express" reference) results in an additional alanine and an amino acid exchange from glutamate to arginine (haplotypes "B" and "C"). At

locus *F183* of the candidate gene *FUS3*, an insertion of three bases (position 526 bp in the “Express” reference) causes an additional cysteine in the amino acid sequence (haplotype “B”). Due to an insertion of nine bases at locus *W102* of the candidate gene *WRI* (position 6749 bp in the “Express” reference) the three amino acids threonine, threonine and isoleucine are added to the amino acid sequence (haplotype “H”). Three base pair deletions cause the elimination of an alanine at locus *W80* (position 4983 bp – 4986 bp in the “Express” reference; haplotypes “B”,”D”) and an elimination of lysine at locus *W2* (position 4259 bp – 4262 bp in the “Express” reference; haplotypes “B”,”C”) of the candidate gene *WRI*. None of these detected InDels was found to be located in a protein domain region.

III. Evaluation of effects on oil content of novel alleles in candidate genes for oil content

III.1 Materials and Methods

III.1.1 Investigated candidate gene loci for oil content

F_2 populations of nineteen selected genotypes, which carried novel alleles (defined in section III 1.2) in candidate genes for oil content, were analyzed for phenotypic effects on oil content. The candidate gene loci investigated in this study as well as the function of the encoded proteins are shown in Table III.1.

Table III.1: Candidate gene loci for oil content investigated in F_2 populations

Gene abbreviation	Gene name	Loci	Function of the protein
PKP2	Plastidial pyruvate kinase 2	*K14, K48, K141*	Plastidial pyruvate kinase subunit β1: Catalyzes the reaction from phosphoenolpyruvate to pyruvate
PKP3	Plastidial pyruvate kinase 3	*123P16*	Plastidial pyruvate kinase subunit β2: Catalyzes the reaction from phosphoenolpyruvate to pyruvate
PKP1	Plastidial pyruvate kinase 1	*PPK196*	Plastidial pyruvate kinase subunit α: Catalyzes the reaction from phosphoenolpyruvate to pyruvate
LEC2	Leafy cotelydon2	*L65, L83*	Transcription factor: Control of seed maturation and regulation of developmental gene expression
IKU2	IKU2	*I103, I124*	LEUCINE RICH REPEAT KINASE: Control of seed size
WRI	Wrinkled	*W2, W80, W102*	Transcription factor: Control of storage compound biosynthesis
CAC2	Acetyl Co-enzyme A carboxylase biotin carboxylase subunit	*A78*	Biotin carboxylase: Catalyzes the reaction from Acetyl-CoA to Malonyl-CoA
CAC3	Acetyl Co-enzyme A carboxylase carboxyltransferase alpha subunit	*C10*	Alpha -carboxyltransferase: Catalyzes the reaction from Acetyl-CoA to Malonyl-CoA
BCCP2	Biotin carboxyl carrier protein 2	*B62, B173*	Biotin carboxyl carrier protein 2: Catalyzes the reaction from Acetyl-CoA to Malonyl-CoA
KAS III	3-Ketoacyl-acyl carrierprotein synthase III	*S13*	Short chain 3-Ketoacyl carrierprotein synthase: Catalyzes the condensation reaction of acetyl-CoA and malonyl-ACP

Table III.1: Candidate gene loci for oil content investigated in F2 populations (continued)

Gene abbreviation	Gene name	Loci	Function of the protein
PDH	Pyruvate dehydrogenase E1α	*H40*	Pyruvate dehydrogenase: Catalyzes the reaction from pyruvate to Acetyl-CoA
PDHK	Pyruvate dehydrogenase kinase	*P57, P133*	Pyruvate dehydrogenase kinase: Inhibition of PDH-complex by phosphorylation

III.1.2 Plant materials

For the field trials 2008/09, genotypes carrying novel haplotypes in candidate genes for oil content, detected in the exotic materials analyzed in II, have been chosen as donors for novel alleles. The sequences from the exotic rapeseed materials were compared with the reference sequences from conventional breeding materials. Haplotypes, that showed novel polymorphisms or a new combination of polymorphisms that did not occur in the reference set, were defined as novel haplotypes or novel alleles. For the F_2 populations in the field trials 2007/08 genotypes carrying novel alleles detected in the "core-set" of the earlier GABI BRIDGE project have been chosen (Ecke, personal communication). Here, alleles with a frequency less than ten in the "core-set" were defined as novel alleles. In four resynthesized genotypes (H48, H226, S4, Yellow II), in four of the winter rapeseed genotypes (DH7AxM, Duell, Erox, Wotan) and in one forage rape (Licapo) novel alleles have been detected. An overview of the selected genotypes with the analyzed novel alleles at the different candidate gene loci is shown in Table III.2 for the field trial 2007/08 and Table III.3 for the field trial 2008/09. With two exceptions the labels of the novel alleles conform with the names given in the diversity study of this project and in the previous GABI BRIDGE project (Lange, personal communication), respectively. For the novel alleles detected in the previous GABI BRIDGE project in the resynthesized rapeseeds "H48", locus *W102* (allele I) and "H226", locus *P133*

(allele “E”) new names were given because in a reanalysis of the sequence data analyzed in GABI BRIDGE additional polymorphisms were detected, resulting in additional alleles.

Table III.2: F_2 populations, candidate gene loci and novel alleles analyzed for effects on oil content in the field trial 2007/08

F_2 population	candidate gene[1], locus	novel allele	environments[2]
Express x H48	*PKP2, K141*	E	3
	KAS III, S13	B	2
	CAC2, A78	D	1
	WRI, W102	I	1
Sollux x H226	*PDHK, P133*	E	3
	BCCP2, B173	E	3
	CAC2, A78	A	1
	CAC3, C10	B	1
Express x DH7AxM	*KAS III, S13*	B	3
Sollux x S4	*PDHK, P57*	B	3
	WRI, W80	C	3
Express x Yellow II	*PDHK, P133*	B	3
	PDH, H40	C	3
	PKP2, K48	D	2
Express x Wotan	*PDHK, P57*	B	3
	PKP2, K48	D	3
Express x Licapo	*PKP2, K141*	B	3
	KAS III, S13	B	2
	CAC2, A78	B	1
	WRI, W102	C	1
Sollux x Erox	*KAS III, S13*	B	3
	BCCP2, B62	E	2
Express x Duell	*WRI, W80*	C	3
	PKP2, K141	D	3
	WRI, W102	B	3

[1] abbreviations see Table III.1; [2] Number of locations from which F_2 plants were included in the analysis

Table III.3: F_2 populations, candidate gene loci and novel alleles analyzed for effects on oil content in the field trial 2008/09

F_2 population	candidate gene[1], locus	novel allele	environments[2]
Express x Altex	*PKP1, PPK196*	D	2
	WRI, W80	B	3
	IKU2, I124	B	2
	BCCP2, B173	B	2
Express x G39	*WRI, W2*	B	3
	PKP2, K14	E	3
Sollux x H40	*PKP2, K141*	G	3
	LEC2, L65	C	1
	PKP3, 123P16	C	2
Express x H149	*PKP3, 123P16*	D	2
Sollux x H365	*PKP2, K14*	D	2
	PKP2, K48	C	2
	PKP2, K141	C	2
	PDHK, P57	C	2
	LEC2, L65	D	2
Sollux x K29	*PKP3, 123P16*	B	3
	IKU2, I103	C	3
	PDHK, P57	B	3
Sollux x K332	*WRI, W2*	C	3
	LEC2, L83	B	3
	IKU2, I103	D	1
Express x L122	*WRI, W2*	D	3
Express x R76	*WRI, W102*	F	3
Express x Xiangyou 11	*PDH, H40*	F	2
	WRI, W102	G	2
	PKP2, K14	B	2

[1] abbreviations see Table III.1; [2] Number of locations from which F_2 plants were included in the analysis

For the field trials, the genotypes carrying novel alleles were crossed with “Express” or “Sollux” depending on their erucic acid content. Genotypes with less than 23% erucic acid content were crossed with “Express”, which is a zero erucic acid quality winter rapeseed cultivar. “Sollux”, which is a winter rapeseed cultivar with high erucic acid content, was crossed with genotypes with more than 23% erucic acid. This

approach was used to minimize segregation for erucic acid content, which would increase the variance in oil content (Ecke *et al.*, 1995).

Segregating F_2 populations were generated out of one F_1 plant to avoid heterogeneity. So, using only one F_2 population per cross across all locations should give the highest power of detection of phenotypic effects of novel alleles for candidate genes segregating in the F_2 populations.

III.1.3 Field experiments

In the field trials 2007/08, 300 seeds per F_2 population were grown per location with the aim to harvest at least 100 plants per F_2 population per location. The segregating F_2 populations were sown in the fields at Göttingen Reinshof, Germany (nursery of the Georg-August-Universität Göttingen) and at Biemsen, Germany (nursery of Syngenta Seeds GmbH). At Thüle, Germany (nursery of Deutsche Saatveredelung AG) they were germinated in the greenhouse, than planted in the field.

In 2008/09, the field trials were performed at Göttingen Reinshof, Germany, at Einbeck, Germany (nursery of KWS SAAT AG) and at Rauischholzhausen, Germany (nursery of the Justus-Liebig-Universität Gießen). To ensure the survival of at least 100 plants per F_2 population per location, in Göttingen and Rauischholzhausen 400 seed per F_2 populations were germinated in the greenhouse and planted in the field, not sown. In the field trial at Thüle, where the plants had been planted in 2007, survival had been much better than in Göttingen in that year. To minimize loss through blackleg disease, in Göttingen for each F_2 population additional 400 plants were germinated in the greenhouse and planted in the field after treatment with “Rapool-Premiumbeizung plus Metconazol + DMM” (RAPOOL-Ring GmbH, Isernhagen, Germany). At the third location Einbeck, 400 seeds per F_2 population were sown in the field. At Rauischholzhausen, the whole F_2 population of “Sollux” x

“H365” and “Express” x “Xiangyou 11” suffered heavy losses from *Rhizoctonia* disease in the greenhouse, where all F_2 populations have been germinated before planting in the field. To get at least 300 plants per F_2 population, more plants were sampled per F_2 population at Göttingen Reinshof for these crosses.

III.1.4 Analysis of oil content and marker analysis

Leaf material was harvested at the beginning of flowering to avoid contamination by foreign pollen on the leaves. Based on the sequence information available from the sequencing, Saaten-Union Biotec GmbH and KWS SAAT AG developed markers for the novel alleles and analyzed DNA from the leaf samples of the F_2 populations to follow the segregation of the novel alleles in these populations.

All plants were manually harvested as single plants and oil content was determined by near-infrared reflectance spectroscopy (NIRS) (Tillmann, 1997) with a near infrared spectrometer 6500 from Foss GmbH (Hamburg, Germany) using year specific calibrations of the “Verband Deutscher Landwirtschaftlicher Untersuchungs- und Forschungsinstitute” (VDLUFA, Kassel, Germany).

III.1.5 Statistical analysis

Orthogonal multivariate analysis of variance (MANOVA) and all further statistical analyses to check the basic requirements for a variance analysis were performed with STATISTICA, version 8 (StatSoft Europe GmbH, Hamburg, Germany). As basic requirements for using variance analysis, normal distribution as well as variance homogeneity have to be given in the experimental data. To test for effects of the novel alleles on oil content, interaction of the environment with the novel alleles as well as general environmental effects on oil content, the allele combination per candidate gene was investigated in combination with the corresponding oil content. For the allele combination at a candidate gene locus the term “genotype” was used in

this analysis. To distinguish between the interactions of the different genotypes at candidate gene loci within F_2 populations and environment, and genotype x environment interaction in the sense of classical quantitative genetics, the former are described as single locus genotype x environment interaction.

Data of the field experiments were checked for normal distribution with the Kolmogorov - Smirnoff test (interpretation of Lilliefors) for all factor levels (oil content per genotype per location) and visual check of normal distribution of the residual sum of squares. Variance homogeneity was tested with the Levine-test for all factor levels. If normal distribution or homogeneity of variance was not given after the elimination of outliers, which were defined as values that differ more than three times the standard deviation from the mean value (offered as “normal” outlier test in STATISTIKA), log10 or BoxCox - transformation was used to reach normal distribution and variance homogeneity. If the data met the basic requirements, MANOVA was performed at a significance level of $\alpha = 0.05$. If no transformation resulted in normal distribution and homogeneity of the variance, significance was tested at a more stringent level ($\alpha = 0.02$), according to Stevens, 1990.

If a MANOVA result showed a significant influence on oil content (α level 0.05 or 0.02), “HSD–test for unbalanced n” (STATISTIKA, version 8) was used as post hoc test. A significant result of the MANOVA only shows that a significant difference in oil content exists within two of the factor levels at least. The posthoc test reveals significant differences among the factor steps. The chosen post hoc test is a modified Tukey-HSD-Test, also usable for unbalanced samples (n). As all posteriori methods, the chosen test also controls for multiple testing to avoid false positive results.

III.2 Results

III.2.1 Performance of the tested F_2 populations in the field trials

Originally, 100 plants were to be sampled per cross and location, but due to frost damages and diseases, especially blackleg disease (*Leptosphaeria maculans* infection), less than 100 plants per location had survived from most crosses in the year 2007/08. In total, seed lots of 154 to 288 plants could be harvested and analyzed by NIRS per F_2 population in this year. In the field trials 2008/09 the aim to harvest 100 plants per F_2 population per location could also not always be reached despite our further efforts described in the materials section. The F_2 populations suffered from heavy losses due to *Rhizoctonia* disease in the greenhouse in Gießen as well as lodging problems at the different field trial locations. In total, seed lots of 260 to 363 plants could be analyzed per F_2 population. The F_2-populations with the highest average oil contents are “Sollux” x “Erox” in the field trials 2007/08 and “Sollux” x “H149” as well as “Sollux” x “K29” in the field trials 2008/09. In the F_2 populations of the field trials the average oil content varied from 43 – 51%. An overview of the number of harvested plants and the distribution of oil content in the F_2 populations is shown in Table III.4. The number of plants and the average oil content per genotype per location is listed in detail in Table A.15 and Table A.16 in the appendix.

Table III.4: Distribution of oil content (%) in the F_2 populations

F_2 population	No. of harvested plants	mean	max	min	standard deviation
Field trials 2007/08					
Express x DH7AxM	165	44.88	48.46	37.40	2.0
Express x Duell	244	46.76	50.93	37.69	2.2
Express x Yellow II	154	45.18	51.37	34.56	2.8
Express x Licapo	263	45.16	49.54	35.55	2.0
Express x H48	288	44.84	49.87	35.23	2.3
Express x Wotan	274	46.61	52.37	32.98	2.7
Sollux x Erox	181	50.84	60.02	38.12	4.2
Sollux x H226	240	49.28	56.72	36.28	3.5
Sollux x S4	249	48.39	57.20	37.72	3.8
Field trails 2008/09					
Express x Altex	344	44.52	50.72	36.47	2.5
Express x G39	363	42.80	49.28	32.18	3.0
Sollux x H40	276	44.83	50.91	36.69	2.5
Sollux x H149	274	46.18	52.83	34.58	3.2
Sollux x H365	264	44.86	51.80	37.65	2.6
Sollux x K29	272	46.16	52.49	34.75	3.1
Sollux x K332	261	45.69	51.25	35.67	2.6
Sollux x L122	273	45.99	52.17	35.39	2.9
Express x R76	276	45.09	50.32	34.52	2.5
Express x Xiangyou 11	260	42.98	48.92	36.61	2.2

III.2.2 Analysis of phenotypic effects

Twenty candidate gene loci of 12 candidate genes for oil content have been analyzed in one to five populations, resulting in a total of 51 analyzed candidate gene loci/ population combinations shown in Table III.2 and Table III.3. It has been tested for influence on oil content due to environment, single locus genotype x environment interaction and genotype. As described in the materials and methods section, to reach the basic requirements for performing a variance analysis, transformations or α-level adjustments were required in some cases. An overview of the applied transformations as well as the tested α-levels after transformation is given in Table III.5.

Table III.5: Applied transformations and tested α-levels after transformation

F_2 population	Candidate gene[1], locus	novel allele	transformation	tested α level
Sollux x H226	*PDHK, P133*	E	log10	0.02
	BCCP2, B173	E	log10	0.02
	CAC2, A78	A	BoxCox	0.05
Express x DH7AxM	*KAS III, S13*	B	log10	0.02
Sollux x S4	*PDHK, P57*	B	BoxCox	0.02
	WRI, W80	C	BoxCox	0.02
Express x Yellow II	*PDHK, P133*	B	log10	0.05
	PDH, H40	C	log10	0.05
	PKP2, K48	D	log10	0.02
Express x Wotan	*PDHK, P57*	B	log10	0.02
	PKP2, K48	D	log10	0.02
Sollux x Erox	*KAS III, S13*	B	log10	0.02
Express x Duell	*WRI, W80*	C	BoxCox	0.05
	PKP2, K141	D	BoxCox	0.02
Express x Altex	*PKP1, PPK196*	D	BoxCox	0.02
	WRI, W80	B	BoxCox	0.05
	IKU2, I124	B	BoxCox	0.02
	BCCP2, B173	B	BoxCox	0.05
Sollux x H365	*PDHK, P57*	C	log10	0.05
Sollux x K29	*PKP3, 123P16*	B	BoxCox	0.02
Express x L122	*WRI, W2*	D	BoxCox	0.02

[1] abbreviations see Table III.1

III.2.2.1 Environmental effects on oil content

In the field trials of the year 2007/08 in all populations that were evaluated over more than one location, with the exception of “Express” x “Wotan”, significant effects of the environment on oil content were observed. Also in the field trials 2008/09 significant environmental effects were detected for almost all F_2 populations tested over more than one location. Only in the populations “Express” x “Altex” and “Sollux” x “H365” no environmental effect on oil content was observed.

III.2.2.2 Candidate gene loci with significant genotype and/ or single locus genotype x environment interaction effects on oil content

Altogether, three candidate gene loci, *S13* of *KAS III*, *K48* of *PKP2* and *L65* of *LEC2*, showed significant genotype effects on oil content in one population each. Significant single locus genotype x environment interactions were observed for loci *S13* of *KAS III*, *P57* of *PDHK*, *W2* of *WRI* and *K141* of *PKP2*. An overview of the significant effects on oil content due to single locus genotype x environment interaction and genotype, as well as the ranking of the genotypes, is given in Table III.6.

KAS III, locus *S13*

The candidate gene locus *S13* of *KAS III* showed a significant genotype effect on oil content in the population “Express” x “Licapo”. *S13* also showed significant single locus genotype x environment interactions in this population and in the two populations “Express” x “H48” and “Express” x “DH7AxM”. Only for the population “Sollux” x “Erox” no significant effect on oil content was observed. In all investigated populations the novel allele was haplotype “B”. In the population “Express” x “Licapo” for the heterozygote genotype, significant higher oil content was detected in comparison to the homozygous genotype of “Express” at Göttingen Reinshof. Analysis in the population “Express” x “DH7AxM” revealed the reverse case at the location Thüle. Here, the homozygous genotype for “Express” had significant higher oil content in comparison to the heterozygote genotype. For the genotype homozygous for allele “B” no significant difference in oil content compared with the other genotypes has been revealed at any single location in these two populations. However, in the population “Express” x “H48” even a dominant gene action of the novel allele is indicated because the heterozygote genotype and the genotype homozygous for allele “B” show significant higher oil content in comparison to the homozygous genotype of “Express”.

LEC2, locus *L65*

In addition to haplotype “C”, where no significant effect was observed, also the novel haplotype “D” at locus *L65* of *LEC2* was tested for effects on oil content. This haplotype was tested in the population “Sollux” x “H365”, where a significant genotype effect was observed. However, this significant effect was not confirmed by the post hoc test.

PKP2, locus *K48 and K141*

At locus *K48* of *PKP2* the possible effect of the novel haplotypes “C” and “D” on oil content was analyzed in the corresponding populations. Only in the population segregation for the novel allele “C” a significant genotype effect was observed. At locus *K141*, the novel alleles “E”, “B”, ”D”, “G” and “C” were tested in the five corresponding populations. Only in the population “Sollux” x “H365”, segregating for the novel allele “C”, a significant single locus genotype x environment interaction was observed. Unfortunately, the post hoc test did not reveal any significant differences in oil content between the genotypes within and over all locations.

WRI, locus *W2*

The haplotypes “B”, “C” and “D” at locus *W2* of *WRI* were tested for significant effects on oil content. Significant single locus genotype x environment interaction was observed in the population segregating for the novel allele “B”. For the genotype homozygous for allele “B” no significant difference in oil content compared to the other genotypes has been revealed at any single location.

PDHK, locus *P57*

Haplotype “B”, detected at locus *P57* of *PDHK*, was analysed in three different populations. Significant single locus genotype x environment interaction was observed in two of the three investigated populations. But only for one population segregating for “B” a significant difference in oil content was detected at one location. In the population “Express” x “Wotan” significant higher oil content was observed for the novel allele at Thüle. At the two additional locations, no significant difference in oil content was observed for the genotypes. The population “Sollux” x “H365” is segregating for the novel allele “C” at the same locus. Here, also single locus genotype x environment interaction was observed, but the post hoc test did not reveal any significant differences between the genotypes within and over all locations.

Table III.6: Summary of significant genotype and single locus genotype x environment (sIG x E) effects at candidate gene loci for oil content

Candidate gene, locus	Population	novel haplotype	Significant effects		Significant differences in oil content between the genotypes at the different locations (HSD test)[1]		
			genotype	sIG x E (P)			
					Thüle	Göttingen	Biemsen
KAS III, S13	Express x H48	B	n.s.	0.001712	n.s.	H,B>A	-
KAS III, S13	Express x DH7AxM	B	n.s.	0.011170	A>H	n.s.	n.s.
KAS III, S13	Express x Licapo	B	0.043846	0.000236	n.s.	H>A	-
PDHK, P57	Express x Wotan	B	n.s.	0.001131	B>A,H	n.s.	n.s.
					Einbeck	Rauischholzhausen	Göttingen
PDHK, P57	Sollux x K29	B	n.s.	0.017172	n.s.	n.s.	n.s.
PDHK, P57	Sollux x H365	C	n.s.	0.002974	n.s.	n.s.	n.s.
WRI, W2	Express x G39	B	n.s.	0.045910	n.s.	n.s.	n.s.
PKP2, K48	Sollux x H365	C	0.003523	n.s.	n.s.	n.s.	n.s.
PKP2, K141	Sollux x H365	C	n.s.	0.048475	n.s.	n.s.	n.s.
LEC2, L65	Sollux x H365	D	0.046602	n.s.	n.s.	n.s.	n.s.

For further information of the candidate genes see Table III.1; [1]A: genotypes homozygous for the "Express" or "Sollux" allele, depending on the cross, B: genotypes homozygous for the novel allele, H: heterozygous genotypes

III.2.2.3 Evaluation of critical polymorphisms in the novel haplotypes at candidate gene loci investigated in the field trials

Critical polymorphisms are polymorphisms like InDels in coding regions causing frame shifts or SNPs causing amino acid exchanges in conserved protein domain regions. These polymorphisms have a high probability to cause functional changes in the protein. None of the novel haplotypes evaluated in the field trials 2007/08 carries a critical polymorphism. In the field trials 2008/09 five haplotypes carrying critical polymorphisms, shown in Table II.12, were tested for possible effects on oil content in the corresponding F_2 populations. In haplotype “C” at locus *K48* of the candidate gene *PKP2* one SNP, causing a non-conservative amino acid exchange in the coding region of the conserved barrel domain of the protein kinase, was observed. This haplotype was investigated in the F_2 population “Sollux” x “H365”, where a significant genotype effect was observed but not confirmed by the post hoc test. Also haplotype “D” at locus *123P16* of the candidate gene *PKP3* carries one SNP causing a non-conservative amino acid exchange, located in the coding region of the conserved barrel domain of the protein kinase. No significant single locus genotype x environment interaction or genotype effect was observed in the F_2 population “Express” x “H149”. For the candidate gene *LEC2* in haplotype “C” at locus *L65* and in haplotype “B” at locus *L83* a critical polymorphism causing a conservative amino acid exchange located in the plant-specific B3-DNA binding domain was detected, respectively. These novel alleles were tested for significant effects on oil content in the F_2 populations “Sollux” x “H40” and “Sollux” x “K332”, respectively. No significant single locus genotype x environment interaction or significant genotype effect has been observed for these loci in the tested F_2 populations. Also no significant effect on oil content was observed for the novel allele “C” at locus *I103* of the candidate gene *IKU2*, tested in the F_2 population “Sollux” x “K29”. This allele exhibits with five the

most amino acid exchanges within a conserved protein domain region. Three of them are non-conservative and two are conservative amino acid exchanges.

With allele “F” at locus *H40* of *PDH*, also an allele for which a stop codon was detected in the analyzed sequence was tested for a possible influence on oil content. Also for this allele no significant effect on oil content was observed.

IV. General Discussion and Conclusions

IV.1 Polymorphism frequency at the differnet candidate gene loci

In a tailed multiplex 454 amplicon sequencing approach, Gholami *et al.* (2012) investigated the genetic diversity in 9 candidate genes for oil content (including *PKP2*, *PKP1*, *FIE*, *FUS3*, *IKU2* and *LEC2*) for 20 genotypes included in the exotic materials selected for this study. Taken all investigated sequences together, Gholami *et al.* (2012) calculated an overall frequency of 1.2 SNP per 100 bases (12 SNPs/ kb) in a total of 24.67 kb sequence information of the investigated loci. This is higher than the calculated frequency of polymorphisms (SNPs and InDels) in the total of 66.94 kb analyzed in the screening sets over all candidate gene loci in this study, which is 7.78 polym./kb. Westermeier *et al.* (2009) analyzed the genetic diversity in candidate genes for oil content in six rapeseed cultivars of common winter rapeseed breeding materials. He detected 93 polymorphisms in a total of 21.4 kb (4.35 polym./kb), which is lower than the evaluated degree of polymorphism over all analyzed loci in this study. Due to the fact, that Gholami *et al.* (2012) used different names for the loci in comparison to the loci investigated in this study, it is not possible to compare the frequency of polymorphisms at individual loci. By comparing the overall frequency determined by Gholami *et al.* (2012) with the degree of polymorphism at individual loci in this study, it is in agreement with the evaluated degree of polymorphism at eleven loci (8.11 to 16.89 polym./kb). At four loci, a higher degree of polymorphism (19.18 to 29.56 polym./kb) has been detected, while the rest of the 32 loci investigated in this study showed a significant lower frequency of polymorphisms (< 5 polym./kb), compared to the study of Gholami *et al.* (2012). Finally, all these comparisons point out that the polymorphism frequency depends on the analyzed material and the analyzed candidate gene loci.

Interestingly, if one compares the detected numbers of polymorphisms between the different loci of the same candidate genes, very different polymorphism frequencies

can be observed. For example in the candidate gene PDH, five loci have been investigated. At locus H81, with 29.56 polymorphisms per 1000 bp, the highest degree of polymorphism has been detected over all investigated loci, while for loci H71 and H25 no polymorphism at all has been detected. For the other two loci, a degree of polymorphism of 8.54 (polym./kb) and 19.18 (polym./kb) has been obtained. Also by comparing the polymorphism frequency obtained in the chosen fragment for which all exotic genotypes have been analyzed, this phenomen can be observed. For example, at candidate gene *PKP2*, the highest number of polymorphisms (47) in the full set of exotic genotypes was found at the selected fragment of locus *K14*. Locus *K48* of the same candidate gene was one of the loci with the lowest number of polymorphisms (4). This big difference is not attributable to a big difference in the analyzed sequence length or a different amount of coding region. The investigated sequences at both loci have a length difference of just 162 bp and with 72 % (locus *K14*) and 71 % (locus *K48*) the same amount of coding region in the analyzed fragments. Also with regard to the successful amplification and sequencing of the different exotic genotypes, there is no explanation for this result. At locus *K14* information was available for 15 resynthesized rapeseed genotypes, five Asian cultivars and five spring rapeseed cultivar sequence. At locus *K48* sequence information of three more resynthesized rapeseed genotypes and one more Asian cultivar could be analyzed. These very different SNP frequencies between different homologues of the same gene were also detected by Gholami *et al.* (2012). They suggest that these discrepancies may reflect the preferential conservation of specific homologues during polyploidization.

IV.2 Identification of different haplotypes and novel alleles

There are two reasons for a possible underestimation of the total number of haplotypes in the investigated plant materials. First, due to the fact that a set of two spring, two Asian and four resynthesized rapeseed lines were used as a screening set to decide whether the rest of the genotypes should be sequenced at this locus or not, we cannot exclude, that additional haplotypes in the rest of the genotypes were missed because the screening set did not show any polymorphisms. Second, potential null alleles have not been taken into account. There are at least two possible reasons why amplification fails reproducibly. First, polymorphisms at the primer binding site do prevent primer annealing. Second, the locus is absent in the examined genotype. This means that there exists a null-allele, which could be defined as one more “different” haplotype. According to Gholami *et al.* (2012), presence/ absence (equivalent to a null allele) variation (PAV) is expected to be wide spread in *B. napus*. Gholami *et al.* (2012) interpreted amplicon absence as a presence/ absence variation. This was conclusively substantiated because of the fact that relatively non-stringent PCR conditions are necessary using tailed primers. Because of that, small variations at the primer binding side did not prevent primer annealing under their conditions. This is not the case in this study because to ensure the locus specificity of the primers in the Sanger sequencing approach, very stringent PCR conditions were chosen. To distinguish between the two possibilities for a reproducible failure in amplification in this study, one could check all fragments of the locus for a positive amplification. A failed amplification caused by polymorphisms at the primer binding site is not likely when amplification fails reproducibly for three to five fragments per locus. However, this was not checked for all candidate gene loci, so one cannot exclude polymorphism at the primer binding site as the reason for the absence of the PCR-product. Therefore, a potential null allele has not been taken

into account by the analysis of novel haplotypes. However, taken time and cost efficiency into account, this possible underestimation of alleles was accepted.

In the winter rapeseed materials the fragment that showed the most polymorphisms in the screening set was sequenced in the rest of the winter rapeseed materials. In the exotic genotype set, the fragment that showed the most polymorphisms in the coding region was sequenced. Due to different selection criteria, in some cases, not the same fragment was selected for sequence analysis of the whole genotype sets of the different plant materials. That means that at some loci the whole exotic genotype set consisting of 32 genotypes was compared to the whole winter rapeseed set of 24 genotypes. At other loci the sequences of the fragment selected in the exotic genotype set were compared only to the screening set of the winter rapeseed materials consisting of eight genotypes. At locus *FIE67* and *W2* the comparison was done only to the winter rapeseed screening set for all fragments, because in the winter rapeseed screening set no polymorphism was detected and accordingly no further sequence analysis was done in the rest of the winter rapeseed genotypes. So there could be an overestimation of novel alleles at the loci where the whole exotic genotype set was compared only to the winter rapeseed screening set. On the other hand, if one compares the number of novel haplotypes detected in exotic genotypes outside the exotic screening set for the loci without the same selected fragment in both materials to the number of novel alleles detected at loci with the same fragment selected in both materials, there is no hint for an overestimation of novel alleles.

At least seventy novel haplotypes, which are specific for the exotic genotypes, were detected. Fifty-two novel alleles were found in the 20 resynthesized rapeseed lines. This high number of novel alleles was expected because it has been shown in many studies that the species used for the artificial resynthesis of rapeseed show large

genetic and phenological variability (Becker *et al.,* 1995 and 1999; Seyis *et al.,* 2001 and 2003; Girke *et al.,* 2012a; Jesske *et al.,* 2013a). Nineteen of the twenty investigated resynthesized rapeseed genotypes were selected on the fact that they clustered in six different subgroups, with average genetic distances (GD) of 0.31 to 0.42 to winter rapeseed (Girke *et al.*, 2012a), to ensure a large genetic diversity in the selected set of resynthesized rapeseed lines. Nine of the haplotypes specific to resynthesized rapeseed have been found exclusively in “MOY4”. This genotype never shared a haplotype with the current winter rapeseed breeding materials at any candidate gene for oil content. One reason could be that for 19 loci no successful analysis of polymorphisms was possible. But in most cases reproducible failure of amplification in the PCR was the reason for this result, which leads to the conclusion that there were polymorphisms at the primer binding site or that the locus is absent. Most likely the reason for this result is attributable to the special type of resynthesized rapeseed line that is “MOY4”. In comparison to the other investigated rapeseed lines, which are resynthesized lines from the two domesticated species *B. rapa* and *B. oleracea*, “MOY4” is a hybridization out of “Yellow Sarson” (*Brassica rapa var. trilocularis, A-genome)* and the wild species *Brassica montana*. That means that *Brassica montana* is the C-genome donor in this resynthesized rapeseed line, while in the other investigated resynthesized lines in most cases *B. oleracea* is the C-genome donor, as shown in Table II.1. Studies of Allender *et al.* (2007) in chloroplast genomes, revealed that there is only a low diversity in *B. oleracea*, compared to wild species. This could be attributable to the fact, that *B. oleracea* has only a single center of domestication in the Eastern Mediterranean region, as described by Allender *et al.* (2007). Also, analysis of the genetic distance of the resynthesized rapeseed “MOY4” to other genotypes carried out by Jesske *et al.* (2013a), gives a further lead to the diversity of this resynthesized rapeseed in comparison to winter

rapeseed and resynthesized rapeseed lines from domesticated species. Jesske *et al.* (2013a) clearly showed that resynthesized lines with wild species provide a genetic diversity absent from resynthesized lines derived from *B. rapa* and *B. oleracea* and the current breeding materials. Moreover, due to the fact that “MOY4” was clustered together with other resynthesized rapeseed lines derived from interspecific crosses with “Yellow Sarson“ in a group with an comparatively high average genetic distance of GD = 0.62 to common breeding winter rapeseed materials, the author hypothesizes that the turnip rape type “trilocularis” also could have a special status due to a divergent evolution in comparison to other turnip rape types (Jesske 2011).

Altogether these results show that the potential of the gene pool of rape seed has not yet been reached to increase the genetic diversity in candidate genes for oil content and can even be enlarged with wild relatives.

IV.3 Significant effects on oil content of novel alleles in candidate genes for oil content in F_2 populations

It is well known that environment has a great impact on oil content in rapeseed. For example, Si *et al.* (2003) analyzed the influence of genotype and environment on oil and protein concentrations of rapeseed in different environments in Australia. The field trials were located in medium-high, medium and low rainfall areas of southern Australian cropping zones. The author concludes that environment had a much larger impact than genotype on oil content. For rainfall and temperature a significant assoziation with oil content was found (Pritchard *et al.,* 2000; Si *et al.,* 2003). With regard to the reported importance of photosynthesis on oil accumulation in green seeds and the positive correlation between oil content and light intensities in *Arabidopsis thaliana* and *Brassica napus* (Ruuska *et al.,* 2004; Goffman *et al*., 2005; Li *et al.,* 2006), the factor sunlight hours could also have an important role in determining oil content at different locations. Also, in field trials conducted in regions

with less strong climate differences than the regions in the study of Si *et al*. (2003), significant effects of the environment on oil content in rapeseed were observed. For example Girke *et al*. (2012b) and Jesske *et al*. (2013b) detected significant effects of the different environments on oil content in resynthesized rapeseed lines tested in field trials that were conducted within Germany, respectively within Germany and the UK. In this study the field trials were actually conducted only in a small region of Germany, comparable to the experiments of Girke *et al.* (2012b). Here significant effects of the environment on oil content were found in almost all F_2 populations. Only the populations “Express” x “Wotan”, “Express” x “Altex” and “Sollux” x “H365” did not show significant environmental effects.

In the earlier project GABI BRIDGE and in the study of Würschum *et al*., (2013) a number of loci form candidate genes for oil content were identified where the allelic differences significantly affected oil content (Ecke and Lange, personal communication; Würschum *et al*., (2013)). Würschum *et al.* (2013) analyzed nine candidate genes for oil content for significant effects on oil content in association studies investigating 685 diverse elite rapeseed inbred lines. The nine investigated candidate genes were genes coding for plastidial pyruvate kinase beta 1 subunit (*PKP2*, called *PK* in Würschum *et al.,* 2013), plastidial pyruvate dehydrogenase (*PDH*), pyruvate dehydrogenase kinase (*PDHK*), biotin carboxyl carrier protein 2 (*BCCP2*), acetyl Co-enzyme A carboxylase biotin carboxylase subunit (*CAC2*, called *BT* in Würschum *et al*., 2013), acetyl co-enzyme A carboxylase carboxyltransferase alpha subunit (CAC3, called BCT in Würschum *et al.,* 2013), wrinkled (*WRI*) and diacylglycerol acyltransferase 1 (*DGAT)*. They studied 17 SNPs derived from these candidate genes for oil content. Unfortuately, it is not clear if these SNPs are locus specific SNPs and represent therefore a specific locus. However, the results and

conclusions in the study of Würschum *et al.* (2013) lead to the conclusion that this is the case. Würschum *et al.* (2013) detected main effect QTL for at least one locus of eight candidate genes for oil content. Seven of these candidate genes for oil content have also been analyzed in this study: *BCCP2*, *PDHK*, *PDH* and *PKP2*. Unfortunately, it is not clear which loci were investigated by Würschum *et al*., (2013), therefore a comparison of the results is not possible. Nevertheless, the results of GABI BRIDGE (personal communication) and Würschum et al., (2013) lead to the conclusion that genetic diversity can be exploited by the selection of favorable alleles.

Haplotypes harboring SNPs that cause amino acid exchanges in conserved protein domain regions of the investigated candidate genes, together with the haplotypes exhibiting InDels causing frame shifts, are defined as haplotypes with critical polymorphisms. Those haplotypes seem to be the most promising haplotypes that entail the possibility to influence oil content because they have a high probability to cause functional changes in the protein. In 16 haplotypes, amino acid exchanges have been found to be located in conserved protein domain regions. For example, three haplotypes at locus *I103* exhibit amino acid exchanges within the conserved protein domain region of the leucine-rich repeat kinase, encoded by the candidate gene *IKU2*. Also in the conserved plant-specific B3 DNA-binding domain at the investigated loci of the transcription factors *LEC2* and *FUS3* SNPs causing amino acid exchanges have been detected. Here, a possible influence on fatty acid biosynthesis can be taken into consideration due to the fact that *LEC1*, *LEC2* and *FUS3* also regulate the transcription factor *WRI* (Yamamoto *et al.,* 2010). The haplotypes harboring the SNPs located in the protein domain region of the leucine-rich repeat kinase at locus *I103* even may show an influence on the seed size, which is of interest with regard to space for storage components. Or in the case of *LEC2*

and *FUS3* one could speculate on an influence on seed maturation (Luerssen *et al.,* 1998; Stone *et al.,* 2001). Therefore, it could be interesting to evaluate the thousand kernel weight and the maturing time for plant genotypes harboring these novel alleles. However, this was not tested within the framework of this study.

Frame shift mutations have been detected at two loci of the candidate genes for oil content. Even a stop codon in consequence of a frame shift was detected for haplotype “F” at locus *H40* of *PDH*. None of these InDels have been found in a functional protein domain. However, a frame shift leads to an altered polypeptide chain or in the case of a stop codon to a shorter protein which can have consequences on the protein function or even lead to loss of protein function. This can have influence on oil content. A shorter protein is usually degraded quickly, therefore, the detected InDel leading to a stop codon is anticipated to have negative effects on oil content, because the candidate gene *PDH* is involved in substrate channeling for the biosynthesis in fatty acid biosynthesis. However, also no significant effect on oil content can be considered for the observed critical polymorphisms, due to the fact that there exist more than one locus coding for the same candidate gene. In line with this, by analyzing variation for specific fatty acids among homozygous lines of diploid *B. oleracea* segregation populations, Barker *et al.* (2007) found evidence that allelic combinations tend to mask the effect of null or active alleles. To validate these theoretic possibilities of such effects, six of the novel alleles carrying critical polymorphisms have been tested for phenotypic effects in F_2 populations: allele “F” at locus *H40*, allele “C” at locus *L65*, allele “B” at locus *L83*, allele “C” at locus *I103*, allele “D” at locus *123P16* and allele “C” at locus *K48*. With one exception none of the haplotypes harboring critical polymorphisms showed significant effects on oil content. Only in the population segregating for the novel allele “C” at locus *K48* of *PKP2*, which harbours a SNP causing a non-conservative

amino acid exchange in the pyruvate kinase barrel domain, a significant gene effect was observed but not confirmed by the post hoc test. On the other hand significant effects on oil content have been detected in F_2 populations segregating for novel alleles where no critical polymorphisms have been detected, as shown in Table III.6. Also, in the earlier GABI BRIDGE project, a locus of a candidate gene, where polymorphisms had been detected exclusively in the intron region, showed a significant association with oil content (Ecke, personal communication). A possible explanation for this could be that such an allele exhibits an amino acid exchange which is located in a functional motif outside the scope of structurally conserved domains. Finally, one also should keep in mind that a very limited region of the candidate genes has been analyzed. The evaluated degree of polymorphism points out, that there could be many more polymorphisms per locus which have not been detected in this study. This indicates that all of the novel haplotypes could be interesting with regard to the search for haplotypes influencing oil content positively.

When significant differences in oil content between contrasting genotypes in the F_2 populations have been observed at several locations, the ranking with regard to the average oil content of alleles often was not consistent between locations. An overview of the average oil content of the alleles in the field trials is shown in Table A.15 and Table A.16 in the appendix. At locus *P57*, for example, the novel allele conveyed higher average oil content at the location Göttingen but lower oil content at the location Einbeck in comparison to the haplotype of “Sollux” in the population “Sollux” x “H365”. These changes in the ranking of alleles at different locations probably reflect the significant single locus genotype x environment interactions that have been observed. This in turn indicates that genotype x environment interactions are based on single locus genotype x environment interactions caused by different

effects of alleles on the phenotype in different environments. Two novel alleles were tested in more than one population. The novel haplotype “B” at locus *P57* of the *PDHK* was tested in two different populations. In both populations a significant single locus genotype x environment effect was observed. When the novel allele “B” at locus *S13* of *KASIII* was tested in more than one population, effects on oil content differed between the four populations. For three populations significant single locus genotype x environment interacting was detected. At one of the populations even an additional significant genotype effect on oil content was observed. But while significantly higher oil content for the heterozygote genotype in comparison to the genotype of the “Express” allele was detected in the population “Express” x “Licapo”, a significant effect of oil content in favour of the allele of “Express” was observed in the second population and a significantly higher oil content was observed for the heterozygote genotype and the genotype with the novel allele in comparison to the “Express” allele in the third population. In the fourth population where the effect of the novel allele “B” was tested, no significant effect on oil content was detected. These results indicate that for complex quantitative traits such as oil content, phenotypic differences between alleles may depend on the genetic background. Also Würschum *et al.* (2013) found evidence for strong effects on the trait oil content based on epistasis, which referes to interactions between alleles at different loci of the same candidate genes as well as alleles of different genes involved in fatty acid biosynthesis. In line with these results, Osborn *et al.* (2007) also observed variable results depending on the genetic background and environment when analyzing significant effects of alleles in different genetic backgrounds of rapeseed on seed yield, which is also a complex quantitative trait. This in turn indicates that extensive testing in different genetic backgrounds and environments is required to determine the value of novel alleles.

A further reason for the result that in most cases no significant effects of genotypes on oil content and single locus genotype x environment interactions were detected, and when significant effects on oil content have been detected, only weak significance was observed could ly in the chosen filed trial approach: First the number of tested genotypes was too low to reach the necessary statistical power to observe further significant effects of novel alleles in candidate genes for oil content on the complex trait oil content. Second, for analyzing phenotypic effects of novel alleles in such a complex trait as oil content, F_2 populations are not optimal because field trials require plots of genetically uniform individuals for reaching the highest statistical power of detection. This could be reached through the construction of a panel of substitution lines, in which segments of the genome showing novel alleles from an exotic genotype are introgressed into the genetic background of a conventional variety using marker assisted selection. Another convenient approach is chromosome doubling of haploid individuals to generate a genetically fixed doubled haploid (DH) for the production of DH populations (Qiu *et al.*, 2006). Nevertheless, generating substitution lines or DH populations is very time intensive and in the time frame of this project was not possible to implement.

In most cases where significant gene effects or single locus genotype x environment interactions have been detected in the variance analysis, the post hoc test did not show significant differences between the genotypes, as shown in Table III.6. This could be attributable to the fact, that the chosen post hoc test is an “A-posteriori” test, which is much more conservative than an “A-priori” test like the LSD test. Usually, one would like to choose a less conservative “A-priori” test to check significant differences between the different factor levels. Less conservative means that at the same α-level more significant effects can be revealed. But this is only possible in an

experiment with so called “defined” multiple comparison tests, which is not the case in this experiment. The other limitation is, that the maximal permitted number of planned multiple comparison tests is restricted by the degrees of freedom of the variance analysis in an “A–priori” test. The maximal permitted number in this approach is not high enough for comparing all genotypes with each other. Therefore, the more conservative Tukey-HSD test was used to check for significant differences between the genotypes in oil content. This test also controls the multiple levels of testing to avoid false positive results, which is not the case in the LSD test. This test normally can only be used in experiments with balanced samples. Due to the expected segregation ratio of 1:2:1 in a F_2 population this condition can never be met. For unbalanced samples one usually would choose the even more conservative “A-priori” test of Scheffé. But STATISTICA provides the opportunity to use a modified test of Tukey-HSD, also usable for testing data from unbalance samples. This test is not as conservative as the Scheffé-test (Köhler *et al.,* 2007; STATISTIKA hand-out).

IV.4 Conclusions

Allelic diversity was evaluated on the DNA sequence level for candidate genes for oil content in exotic genotypes represented by Chinese, spring and resynthesized rapeseed in comparison to reference sequences from the winter rapeseed varieties “Express” or “Tapidor”. Novel alleles were identified due to comparing the sequences of the exotic materials to sequences of winter rapeseed materials. All in all, the results show that exotic germplasm is indeed a source of untapped genetic variation, even for genes involved in a biosynthetic pathway as basic as fatty acid biosynthesis. Confirming the results of Jesske *et al*. (2013a) that wild relatives enhance the genetic diversity in comparison to resynthesized rapeseed lines from domesticated species, the genotype with the largest number of unique alleles was “MOY4”. The diversity study also indicates that the analyzed materials may contain functionally different alleles that may influence oil content.

Some of the detected novel alleles in the exotic materials as well as novel alleles detected in an earlier GABI BRIDGE project have been tested for effects on oil content in F_2 populations. Due to the lower statistical power of detection in F_2 populations compared to DH populations and substitution lines, it is possible that effects of novel alleles on oil content remained undetected. On the other hand this does mean that a significant effect of a novel allele on oil content detected by this approach can be considered as relatively "true". Based on this thought, observed significant single locus genotype x environment interactions or a significant gene effect on oil content can be seen as a strong hint for the importance of the observed locus, respectively, the allele for oil content, even if the post hoc test did not reveal any significant differences in oil content between the genotypes. Four candidate gene loci showed significant single locus genotype x environment interactions, two of them actually in three different populations, giving a strong indication that allelic diversity at

these loci leads to variation in oil content. In addition, hints for dominant gene actions were found at single locations.

A problem in using the novel alleles of exotic genotypes is that the exotic genotypes are agronomically weak and that they are not adapted to the cultivation in Germany (Girke *et al*., 2012b; Jesske *et al*., 2013a and 2013b). Moreover, the observed effects on oil content showed that novel alleles which appear to have a positive effect in one genetic background may not have the same effect in all backgrounds. Therefore, a longer pre-breeding would be required for introgressing the alleles in an adapted and agronomically potent genetic background. Because of this, the observed significant effects first should be confirmed in field trials with substitution lines or DH populations in different environments. If such an approach confirms for example the positive effect on oil content of the novel allele "B" at locus *P57* of *PDHK*, this allele could eventually be used to broaden the genetic diversity for the trait oil content.

Summary

Novel alleles in genes for oil content can entails the possibility for new approaches to positively influence the oil content in rape seed (Osborn *et al*., 2007; Würschum *et al.,* 2013; Jesske *et al.,* 2013a). In this study, the allelic diversity of a large number of candidate gene loci for oil content has been analyzed on the DNA sequence level in a broad set of exotic rapeseed materials represented by Chinese, spring and resynthesized rapeseed and has been compared to current winter rapeseed breeding materials. Many alleles have been observed that were specific to the exotic genotypes. The highest number of novel alleles was detected in a resynthesized rapeseed out of “Yellow Sarson” (*Brassica rapa var. trilocularis, A-genome)* and *Brassica montana (C-genome)*. Many of the detected alleles showed non-silent SNPs and InDels in coding regions of the analyzed candidate gene loci. Some of these polymorphisms even were found to be located in protein domain regions. These results indicate that some of the detected alleles may influence the oil content. To test a possible influence on oil content, novel alleles detected in the diversity study as well as novel alleles detected in the earlier GABI BRIDGE project have been tested in segregating F_2 populations at three different locations, respectively. Significant gene effects on oil content were detected at three candidate gene loci, *S13* of *KAS III*, *K48* of *PKP2* and *L65* of *LEC2* in one population each. Single locus genotype x environment interaction affecting oil content was significant at six candidate gene loci, including the candidate gene locus *S13* of *KAS III*. These significant effects on oil content can be seen as a strong indicator for the importance of the observed locus respectively the allele with regard to oil content.

Zusammenfassung

Durch eine Erhöhung der Diversität in Kandidatengenen für Ölgehalt könnten sich neue Ansätze zur Erhöhung des Ölgehalts ergeben (Osborn *et al*., 2007; Würschum *et al*., 2013). Im Rahmen dieser Arbeit wurde die allele Diversität auf der Basis von DNA-Sequenzen in einer großen Anzahl von Kandidatengenloci für Ölgehalt in exotischem Rapsmaterial (Sommerrapssorten, chinesische Rapssorten und Resynthesen) im Vergleich zu Winterraps untersucht. Viele Allele wurden exklusiv in den exotischen Genotypen aufgefunden (neue Allele). Die höchste Anzahl an neuen Allelen wurde in der Resynthese "MOY4" (*Brassica rapa var. trilocularis* x *Brassica montana)* entdeckt. Viele der Allele wiesen SNP, die zu Aminosäureaustauschen führen, sowie InDel im kodierenden Bereich der untersuchten Kandidatengenloci auf. Einige dieser Polymorphismen konnten sogar Bereichen von Proteindomänen zugeordnet werden. Im Großen und Ganzen konnte in der durchgeführten Diversitätsstudie gezeigt werden, dass die untersuchten exotischen Genotypen genutzt werden können, um die allele Diversität in Kandidatengenen für Ölgehalt zu erhöhen. Zudem weist die relativ hohe Anzahl an nicht stillen Polymorphismen in den kodierenden Bereichen der verschiedenen Kandidatengenloci darauf hin, dass einige aufgefundene Allele den Ölgehalt beeinflussen könnten. Um einen möglichen positiven Einfluss von neuen Allelen in Kandidatengenen für Ölgehalt auf den Ölgehalt zu testen, wurden diese in spaltenden F_2-Populationen an jeweils drei verschiedenen Standorten untersucht. In diesem Versuch wurden neue Allele, die in der durchgeführten Diversitätsstudie entdeckt wurden, sowie neue Allele, welche bereits in dem vorigen GABI BRIDGE Projekt aufgefunden wurden, mittels Varianzanalyse auf eine Auswirkung auf den Ölgehalt getestet. In einzelnen Populationen konnten signifikante Geneffekte auf den

Ölgehalt an drei verschiedenen Kandidatengenloci ermittelt werden: *S13* (Kandidatengen *KAS III)*, *K48* (Kandidatengen *PKP2*), *L65* (Kandidatengen *LEC2*). Einzel-Locus-Genotyp x Umweltinteraktionen mit einem signifikanten Effekt auf Ölgehalt wurden an sechs Kandidatengenloci inklusive *S13* von *KAS III* aufgefunden. Diese aufgefundenen signifikanten Effekte auf den Ölgehalt können als deutlicher Hinweis auf die Wichtigkeit der untersuchten Loci beziehungsweise des Alleles auf den Ölgehalt interpretiert werden.

References

Abbadi, A., Brummel, M., Spener, F. „Knockout of regulatory site of 3-ketoacyl-ACP synthase III enhances short- and medium acyl-ACP synthesis“ The Plant Journal (2000) 24, 1-9

Allender, C. J., Allainguillaume, J., Lynn, J., King, G. J. “Simple sequence repeats uneven distribution of genetic diversity in chloroplast genomes of Brassica oleracea L. and (n=9) wild relatives” Theor Appl Genet (2007) 114, 609-618

Allender, C. J. and King, G. J. “Origin of the amphiploid species *Brassica napus L.* investigated by chloroplast and nuclear molecular markers” BMC Plant Biology (2010) 10:54 doi:10.1186/1471-2164-14-120

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

Ambasht, P. K. and Kayastha, A. M. “Plant pyruvate kinase” Biologia plantarum (2002) 45, 1-10

Andersson, C., Brimer, L., Cottrill, B., Fink-Gremmels, J., Jaroszewski, J., Sørensen, H. “Glucosinolates as undesirable substances in animal feed: Scientific Opinion of the Panel on Contaminants in the Food Chain” The EFSA Journal (2008), no. 590

Andre C. and Benning C. “*Arabidopsis* seedlings deficient in a plastidial pyruvate kinase are unable to utilize seed storage compounds for germination and establishment” Plant Physiol (2007) 145, 1670-1680

Andre, C., Froehlich, J. E., Moll, M. R., Benning, C. “A heteromeric plastidial pyruvate kinase complex involved in seed oil biosynthesis in *Arabidopsis*” The Plant Cell (2007) 19, 2006-2022

Arabidopsis Genome Initiative “Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*” Nature (2000) 408, 796-815

Arumuganathan K. and Earle E. D. “Nuclear DNA content of some important plant species” Plant Mol Biol Report (1991) 9, 208-218

Babula, D., Kaczmarek, M., Barakat, A., Delseny, M., Quiros, C. F., Sadowski, J. “Chromosomal mapping of Brassica oleracea based on ESTs from *Arabidopsis* thaliana: Complexity of the comparative map” Mol. Genet. Genomics (2003) 268, 656-665

Bach, K. E. “Duplizierte Gene in *Brassica napus* – genetische Vielfalt in Kandidatengenen für Ölgehalt. (Dissertation), Universität Potsdam, Potsdam (2007)

Bancroft, I., Morgan, C., Fraser, F., Higgins, J., Wells, R., Clissold, L., Baker, D., Long, Y., Meng, J., Wang, X., Liu, S., Trick, M. “Dissecting the genome of polyploidy crop oilseed rape by transcriptome sequencing” Nature Biotechnology (2011) 29, 762-768

Barker, G. C., Larson, T. R., Graham, I. A., Lynn, J. R., King, G. J. “Novel Insights into Seed Fatty Acid Synthesis and Modification Pathways from Genetic Diversity and Quantitative Trait Loci Analysis of the Brassica C Genome” Plant Physiol (2007) 144, 1827-1842

Baud S. and Graham I. “A spatiotemporal analysis of enzymatic activities associated with carbon metabolism in wild-type and mutant embryos of *Arabidopsis* using insitu histochemistry” The Plant Journal (2006) 46, 155-169

Baud, S., Santos-Mendoza, M., To, A., Harscoët, E., Lepiniec, L., Dubreucq, B. “WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in *Arabidopsis*” The Plant Journal (2007a) 50, 825-838

Baud, S., Wuilleme, S., Dubreucq, B., de Almeida, A., Vuagnat, C., Lepiniec, L., Miquel, M., Rochat, C. „Function of plastidial pyruvate kinases in seeds of *Arabidopsis thaliana*” The Plant Journal (2007b) 52, 405-419

Baud, S., Wuilleme, S., To, A., Rochat, C., Lepiniec, L “Role of Wrinkled1 in the transcriptional regulation of glycolytic and fatty acid biosynthesis genes in *Arabidopsis*” The Plant Journal (2009) 60, 933-947

Beare, J. L., Gregory, R. W., Campbell, J. A. “The effects of different varieties of rapeseed oil on weight gain, and of golden rapeseed oil on the reproduction of the rat” Can J Biochem Physiol (1959) 37, 1191 1195

Becker, H. C., Engqvist, G. M., Karlsson, B. “Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers” Theor Appl Genet (1995) 91, 62-67

Becker, H. C., Löptien, H., Röbbelen, G. “Breeding of the *Brassica*: An overview” In: Gomez-Campo C (ed) Biology of Brassica Coenospecies. Elsevier, Science BV, Amsterdam (1999) 413-460

Becker, H. C. „Nutzung genetischer Ressourcen in der Pflanzenzüchtung“ In: Hammer, K. and Gladis, T. (Hrsg.): Schriften zu Genetischen Ressourcen, (2001) 16, 79 – 90

Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R., Mathews, S. “Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*” Proc. Natl. Acad. Sci. USA (2010) 107, 18724-18728

Bouyer, D., Roudier, F., Heese, M., Andersen, E. D., Gey, D., Nowack, M. K., Goodrich, J., Renou J.-P., Grini, P. E., Colot, V., Schnittger, A. “Polycomb Repressive Complex 2 Controls the Embryo-to-Seedling Phase Transition” PLoS Genet (2011) 7(3): e1002014

Braybrook, S. A., Stone, S. L., Park, S., Bui, A. Q., Le, B.H., Fischer, R. L., Goldberg, R. B., Harada, J. J. “Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis” Proc. Natl. Acad. Sci. USA (2006) 103, 3468-3473

Brunel, D., Froger, N., Pelletier, G. “Development of amplified consensus genetic markers (ACGM) in Brassica napus from Arabidopsis thaliana sequences of known biological function” Genome (1999) 42, 387-402

Budde, R. J. A.; Fang T. K.; Randall, D. D. ”Regulation of the phosphorylation of mitochondrial pyruvate dehydrogenase complex in situ” Plant Physiol. (1988) 88, 1031-1036

Butruille, D. V., Guries, R. P., Osborn T. C. “Increasing Yield of Spring Rape Hybrids through Introgression of Winter Germplasm”; Crop Sci. (1999) 39, 1491-1496

Cao, H. X. and Schmidt, R. “Screening of Brassica napus bacterial artificial chromosome library using highly parallel single nucleotide polymorphism assays” BMC Genomics (2013) 14:603; doi:10.1186/1471-2164-14-603

Cavell, A. C., Parkin, I. A. P., Dean, C., Trick, M. “Collinearity between a 30-centimorgan segment of Arabidopsis thaliana chromosome 4 and duplicated regions within the Brassica napus genome” Genome (1998) 41, 62-69

Cernac, A. and Benning, C. “WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in *Arabidopsis*” The Plant Journal (2004) 40, 575-585

Cernac, A., Andre, C., Hoffmann-Benning, S., Benning, C. “WRI1 is required for seed germination and seedling establishment” Plant Physiol (2006) 141, 725-757

Chanvivattana, Y., Bishopp, A., Schubert D., Stock, C., Moon Y.-H., Sung, Z. R., Goodrich, J. “Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*” Development (2004) 131, 5263-5276

Chaudhury, A. M., Ming, L., Miller, C., Craig, S., Dennis, E. S., Peacock, W. J. “Fertilization-independent seed development in *Arabidopsis thaliana*” Proc. Nat. Acad. Sci. USA (1997) 94, 4223-4228

Crowe, A. J., Abenes, M., Plant, A., Moloney, M. M. “The seed-specific transactivator, ABI3, induces oleosin gene expression” Plant Sci. (2000) 151, 171-181

Dehesh, K., Tai, H., Edwards, P., Byrne, J., Jaworski, J. G. “Overexpression of 3-Ketoacyl-Acyl-Carrier Protein Synthase IIIs in Plants Reduces the Rate of Lipid Synthesis” Plant Physiol (2001) 125, 1103-1114

Diers and Osborn “Genetic diversity of oilseed *Brassica napus* germplasm based on restricted length polymorphisms.” Theor Appl Genet (1994) 88, 662-668

Diers, W. B., McVetty P. B. E., Osborn T. C. “Relationship between Heterosis and Genetic Distance Based on Restriction Length Polymorphism Markers in Oilseed Rape (*Brassica napus L.*)”; Crop Sci (1996) 36,79-83

Ecke, W., Uzunova, M., Weißleder, K. „Mapping the genome of rapeseed (*Brassica napus L.*). II. Localization of genes controlling erucic acid synthesis and seed oil content” Theor Appl Genet (1995) 91, 972-977

Elborough, K. M., Winz, R., Deka, R. K., Markham, J. E., White, A. J., Rawsthorne, S., Slabas, A. R. “Biotin carboxyl carrier protein and carboxyl transferase subunits of the multi-subunit form of acetyl-CoA carboxylase from *Brassica napus*: cloning and analysis of expression during oilseed rape embryogenesis” Biochem J (1996) 315, 103-112

FAO statistics from November 2013 http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor

Focks, N. and Benning, C. “Wrinkled1: a novel, low-seed-oil mutant of *Arabidopsis* with a deficiency in the seed-specific regulation of carbohydrate metabolism” Plant Physiol (1998) 118, 91-101

Garcia D., Saingery V., Chambrier P., Mayer U., Jürgens G., Berger F. “*Arabidopsis haIKU* Mutants Reveal New Controls of Seed Size by Endosperm” Plant Physiol (2003) 131, 1661-1670

Gehringer, A., Snowdon, R., Spiller, T., Basunanda, P., Friedt, W. „New Oilseed Rape (*Brassica napus*) Hybrids with High Levels of Heterosis for Seed Yield under Nutrient-poor Conditions” Breeding Science (2007) 57, 315-302

Gholami, M., Bekele, W. A., Schondelmaier, J., Snowdon, R. J. „A tailed PCR procedure for cost-effective , two-order multiplex sequencing of candidate genes in polyploidy plants” Plant Biotechnology Journal (2012) 1-11

Giraudat, J., Hauge, B. M., Valon, S., Smalle, J., Parcy, F., Goodman, H. M. “Isolation of the *Arabidopsis ABI3* gene by positional cloning.” The Plant Cell (1992) 4, 1251-1261

Girke, A. „Neue Genpools aus resynthetisiertem Raps (*Brassica napus L*.) für die Hybridzüchtung“ PhD thesis, School of Agricultural Sciences, Georg-August Universität Göttingen (2002)

Girke A., Schierholt, A., Becker, H.C. „Extending the rapeseed genepool with resynthesized *Brassica napus L*. I: Genetic diversity” Genet Resour Crop Evol (2012a) 59, 1441-1447

Girke A., Schierholt, A., Becker, H. C. „Extending the rapeseed gene pool with resynthesized *Brassica napus L*. II: Heterosis” Theor Appl Genet (2012b) 124, 1017-1026

Goffman, F. D.; Alonso, A. P.; Schwender J.; Shachar-Hill, Y.; Ohlrogge, J. B. “Light enables a very high efficiency of carbon storage in developing embryos of rapeseed” Plant Physiol (2005) 138, 2269-2279

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. M., Coupland, G. ” A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*” Nature (1997) 386, 44-51.

Grant, M. R., McDowell, J. M., Sharpe, A. G., de Torres Zabala, M., Lydiate, D. J., Dangl, J. L. “Independent deletions of a pathogen-resistance gene in *Brassica* and *Arabidopsis*” Proc. Natl. Acad. Sci. USA (1998) 95, 15843-15848

Grossniklaus, U., Vielle-Calzada, J., Hoeppner, M. A., Gagliano, W. B. “Maternal Control of Embryogenesis by *MEDEA*, a *Polycomb* Group Gene in *Arabidopsis”* Science (1998) 280, 446-450

Hajduch, M.; Hearne, L. B.; Miernyk, J. A.; Casteel, J. E.; Joshi, T.; Agrawal, G. K.; Song, Z.; Zhou, M.; Xu, D.; Thelen, J. J. “Systems Analysis of Seed Filling in *Arabidopsis*: Using General Linear Modeling to Asses Concordance of Transcriptional and Protein Expression” Plant Physiol. (2010) 152, 2078-2087

Hasan, M.; Seyis, F.; Badani, A: G., Pons-Kühnemann, J.; Friedt, W.; Lühs, W:; Snowdon, R.J. “Analysis of genetic diversity in Brassica napus L. gene pool using SSR markers” Genetic resources and Crop Evolution (2006) 53, 793-802

Jaworski, J. G., Clough, R. C., Barnum, S. R. “A Cerulenin Intensive Short Chain 3-Keto-Acyl Carrier Protein Synthase in Spinacia oleracea Leaves” Plant Physiol. (1989) 90, 41-44

Jaworski, J. G., Post-Beittenmiller, D., Ohlrogge, B. „Acetyl-acyl carrier protein is not a major intermediate in fatty acid biosynthesis in spinach“ Eur. J. Biochem. (1993) 213, 981-987

Jesske, T. „Brassica - Wildarten als neue genetische Ressource für die Rapszüchtung” PhD thesis, School of Agricultural Sciences, Georg-August-Universität Göttingen, (2011)

Jesske, T., Olberg, B., Schierholt, A., Becker, H. C., „Resynthesized lines from domesticated and wild Brassica taxa and their hybrids with Brassica napus L.: Genetic diversity and hybrid yield” Theor Appl Genet (2013a) doi: 10.1007/s00122-012-2036-y

Jesske, T., Schierholt, A., Becker, H. C. “Glucosinolates and other seed quality traits of resynthesized Brassica napus L. derived” Genet Resopur Crop Evol (2013b) doi: 10.1007/s10722-013-9993-0

Jiang, D., Wang, Y., Wang, Y., He, Y. „Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the *Arabidopsis* Polycomb Repressive Complex 2 Components“ PLoS ONE (2008) 3(10), e3404

Johnston, M. L., Luethy, M. H., Miernyk, J. A., Randall, D. D. “Cloning and molecular analyses of the Arabidopsis *thaliana* plastid pyruvate dehydrogenase subunits” Biochim Biophys Acta (1997) 1321, 200–206.

Jullien, P. E., Katz, A., Oliva, M., Ohad, N., Berger, F. „Polycomb Group Complexes Self-Regulate Imprinting of the Polycomb Gene MEDEA in *Arabidopsis*“ Current Biology (2006) 16, 486-492

Kagaya, Y., Toyoshima, R., Okuda, R., Usui, H., Yamamoto, A., Hattori, T. “LEAFY COTYLEDON1 controls seed storage protein genes through its regulation of FUSCA3 and ABSCISIC ACID INTENSIVE3.” Plant Cell Physiol. (2005) 46, 399-406

Katz A., Oliva, M., Mosquna, A., Hakim, O., Ohad, N. “FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development” The Plant Journal (2004) 37, 707-719

Ke, J., Behal, H. R., Back, S. L., Nikolau, B. J., Wurtele, E. S., Oliver, D. J. “The role of pyruvate dehydrogenase and acetyl-coenzyme A synthase in fatty acid synthesis in developing *Arabidopsis* seeds”; Plant Physiol. (2000) 123, 497-508

Kimber, D. S. and McGregor, D. I. “The species and their origin, cultivation and world production” In: Kimber D. S. and McGregor D. I. (eds), Brassica Oilseeds: production and utilization. CABI Publishing, Wallingford, UK (1995), pp. 1-9

Kinoshita, T., Harald, J. J., Goldberg, R. B., Fischer, R. L. „Polycomb repression of flowering during early plant development“ Proc. Natl. Acad. Sci. (2001) 98, 14156-14161

Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J.,Wells, D., Katz, A., Margossian, L., Harada, J. J., Goldberg, R. B., Fischer, R. L. “Control of fertilization-independent endosperm development by the MEDEA polycomb gene in *Arabidopsis*” Proc. Natl. Acad. Sci. USA (1999) 96, 4186–4191

Köhler, C., Henning, L., Bouveret, R., Gheyselinck, U., Grossniklaus, U., Gruissem, W. "*Arabidopsis* MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed Development" The EMBO Journal (2003a) 22, 4804-4814

Köhler, C., Henning, L., Spillane, C., Pien, S., Gruissem, W., Grossniklaus, U. „The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1" Genes and Development (2003b) 17, 1540-1553

Köhler, W., Schachtel, G., Voleske, P. „Biostatistik: Vierte, aktualisierte und erweiterte Auflage" (2007) ©Springer-Verlag Berlin Heidelberg

Kroj, T., Savino, G., Valon, C., Giraudat, J., Parcy, F. "Regulation of storage protein gene expression in *Arabidopsis*" Development (2003) 130, 6065-6073

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

Lan, T. H., DelMonte, T. A., Reischmann, K. P., Hyman, J., Kowaslki, S., McFerson, J., Kresovich, S., Raterson, A. H. "An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*" Genome Res. (2000) 10, 776-788

Lernmark, U. and Gardeström, P. "Distribution of pyruvate dehydrogenase complex activities between chloroplasts and mitochondria from leaves of different species" Plant Physiol. (1994) 106, 1633–1638

Li, Y., Beisson, F., Pollard, M., Ohlrogge, J. "Oil content of *Arabidopsis* seeds: The influence of seed anatomy, light and plant to plant variation" Phytochemistry (2006) 67, 904-915

Li, R. J., Hu, Z.Y., Zhang, H. S., Wang, H.Z., Hua, W. "Cloning and functions analysis of a pyruvate dehydrogenase kinase in Brassica napus" Plant cell reports (2011) 30, 1533-1540

Lombard, V. and Delourme, R. "A consensus linkage map for rapeseed (Brassica napus L.): construction and integration of three individual maps from DH populations" Theor Appl Genet (2001) 103, 491-507

Lonien, J. and Schwender, J. „Analysis of Metabolic Flux Phenotypes for Two *Arabidopsis* Mutants with Severe Impairment in Seed Storage Lipid Synthesis" Plant Physiol. (2009) 151, 1617-1634

Lotan, T., Ohto, M., Yee, K. M., West, M. A., Lo, R., Kwong, R. W., Yamagishi, K., Fischer, R. L., Goldberg, R. B., Harada, J. J. "*Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells" Cell (1998) 93, 1195-1205

Luerssen, H., Kirik, V., Herrmann, P., Misera, S. "FUSCA3 encodes a protein with a conserved VP1/AB13-like B3 domain which is of functional importance for the regulation of seed maturation in *Arabidopsis thaliana*" The Plant Journal (1998) 15, 755-764

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

Luo, M., Bilodeau, P., Koltunow, A., Dennis, E. S., Peacock, W. J., Chaudhury, A. M. "Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*" Proc. Natl. Acad. Sci. USA (1999) 96, 296-301

Luo, M., Bilodeau, P., Dennis, E. S., Peacock, W. J., Chaudhury, A. “Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds” Proc. Natl. Acad. Sci. USA (2000) 97, 10637-10642

Luo, M., Dennis, E. S., Berger, F., Peacock, W. J., Chaudhury, A. “*MINISEED3 (MINI3)*, a *WRKY* family gene, and *HAIKU2 (IKU2),* a leucine-rich repeat (*LRR) KINASE* gene, are regulators of seed size in Arabidopsis” Proc. Natl. Acad. Sci. USA (2005) 102, 17531-17536

Lysak, M. A., Koch, M. A., Pecinka, A., Schubert, I. “Chromosome triplication found across the tribe Brassicaceae” Genome Res. (2005) 15, 516-525

Maeo, K., Tokuda, T., Ayame, A., Mitsui, n., Kawai, T., Tsukagoshi, H., Ishiguro, S., Nakamura, K. “An AP2-type transcription factor, WRINKLED1, of *Arabidopsis* thaliana bind to the AW-box sequence conserved among proximal upstream regions of genes involved in fatty acid synthesis” The Plant Journal (2009) 60, 476-487

Marchler-Bauer, A., Panchenko, A.R., Shoemaker, A., Thiessen, P.A., Geer, L.Y., Bryant, S.H. “CDD: conserved domains and protein three-dimensional structure.” Nucleic Acids Res. (2013) 30, 281-283

Marillia, E.; Micallef, B. J.; Micallef, M.; Weninger, A.; Pedersen, K. K.; Zou, J.; Taylor, D. C. “Biochemical and physiological studies of *Arabidopsis thaliana* transgenic lines with repressed expression of the mitochondrial pyruvate kinase” Journal of Experimental Botany (2003) 54, 259-270

Masaki, T., Mitsu, N., Tsukagoshi, H., Nishii, T., Morikami, A., Nakamura, K. “ACTIVATOR of Spomin:LUC1/WRINKLED1 of *Arabidopsis* thaliana transactivates sugar-inducible promoters” Plant Cell Physiol. (2005) 46, 547-556

Mentzen, W. I., Peng, J., Ransom, N., Nikolau, B. J., Wurtele, E. S. „Articulation of three core metabolic processes in *Arabidopsis*: Fatty acid biosynthesis, leucine metabolism and starch metabolism“ BMC Plant Biology (2008) 8:76 doi:10.1186/1471-2229-8-76

Metzger, J. O. and Bornscheurer, U. “Lipids as renewable resources: current state of chemical and biotechnological conversion and diversification”; Appl Microbiol Biotechnol (2006) 71, 13-22

Mosquna, A., Katz, A., Decker, E. L., Rensing, S. A., Reski, R., Ohad, N. „Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution“ Development (2009) 136, 2433-2444

Mu, J., Tan, H., Zheng, Q., Fu, F., Zhang, Y., Yang, X., Wang, T., Chong, K., Wang, W., Zuo, J. „LEAFY COTELYDON1 is a key regulator of fatty acid biosynthesis in *Arabidopsis*“ Plant Physiol. (2008) 148, 1042-1054

Needleman, S. B. & Wunsch, C. D. “A. general method applicable to the search for similarities in the amino acid sequence of two proteins” J. Mol. Biol. (1970) 48, 443–453

Negm, F. B., Cornel, F. A., Plaxton, W. C. “Suborganellar Localization and Molecular Characterization of Nonproteolytic Degraded Leukoplast Pyruvate Kinase from Developing Castor oil Seeds” Plant Physiol. (1995) 109, 1461-1469

O´Neill, C. M. and Bancroft, I. “Comparative physical mapping of segments of the genome of *Brassica oleracea var. alboglabra* that are homologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*” Plant. J. (2000) 23, 233-243

Ohad, N., Margossian, L., Hsu, D. C., Williams, C., Repetti, P., Fischer, R. L. “A mutation that allows endosperm development without fertilization” Proc. Natl. Acad. Sci. USA (1996) 93, 5319-5324

Ohad, N., Yadegari,R., Margossian, L., Hannon, M., Michaeli, D., Harada, J. J., Goldberg, R. B., Fischer, R. L., “Mutations in FIE, a WD Polycomb Group Gene, Allow Endosperm Development without Fertilization” The Plant Cell (1999) 11, 407–415

Ohlrogge, J. and Browse, J. “Lipid biosynthesis” Plant Cell (1995) 7, 957–970

Osborn T. C., Kramer C., Graham E., Braun C. J. “Insights and Innovations from wide Crosses: Examples from canola to tomato”; Crop Sci. (2007) 47, Supplement 3, 228-237

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

Parkin, I. A. P., Sharpe, A. G., Lukens, L., Keith, D. J., Lydiate, D. J. “Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape)” Genome (1995) 38, 1122-1131

Parkin, I. A. P., Gulden, S. M., Sharpe, A. G., Lukens, L., Trick, M., Osborn, T. C., Lydiate, D. J. “Segemental structure of the Brassica napus genome based on comparative analysis with *Arabidopsis thaliana*” Genetics (2005) 171, 765-781

Plaxton, W. C., Smith, C. R., Knowles, V. L. “Molecular and regulatory properties of leucoplast pyruvate kinase from *Brassica napus* (Rapeseed) suspension cells” Archives of Biochemistry and Biophysics (2002) 400, 54-62

Plieske, J. and Struss, D. “Microsatellite markers for Genome analysis in *Brassica.* I. Development and abundance in *Brassica* species.” Theor Appl Genet (2001) 102,689-694

Post-Beittenmiller, D., Roughan, G., Ohlrogge, J. B., “Regulation of Plant Fatty Acid Biosynthesis” Plant Physiol. (1992) 100, 923-930

Pritchard, F. M., Eagles, H. A., Norton, R. M., Salsbury, P. A. Nicolas, M. “Environmental effects on seed composition of Victorian canola” Australian Journal of Experimental Agriculture (2000) 40, 679-685

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

Rana, D., van den Boogaart, T., O´Neill, C. M., Hynes, L., Bent, E., Macpherson, L., Park, J. Y., Lim, Y. P., Bancroft, I. “Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives” The Plant Journal (2004) 40, 725-733

Riaz, A., Li, G., Quresh Z., Swati, M. S., Quiros, C. F. “Genetic diversity of oilseed Brassica napus inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance” Plant breeding (2001) 120, 411-415

Roine, P. and Uksila, E. “Experiments on feeding rats with rapeseed oils” Acta Agric Fenn (1959) 94, 151-153

Roschzttardtz, H.; Fuentes, I.; Vasquez, M.; Corvalan, C.; Leon, G.; Gomez, I.; Araya, A.; Holuigue, L.; Vicente-Carbajosa, J.; Jordana, X. “A Nuclear Gene Encoding the Iron-Sulfur Subunit of Mitochondrial Complex II Is Regulated by B3 Domain Transcription Factors during Seed Development in *Arabidopsis*” Plant Physiol. (2009) 150, 84-95

Ruuska, S., Girke, T., Benning, C., Ohlrogge, J. B. “Contrapuntal Networks of Gene Expression during *Arabidopsis* Seed Filling” Plant Cell (2002) 14, 1191-1206

Ruuska, S. A., Schwender, J., Ohlrogge, J. B. „The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes” Plant Physiol. (2004) 136, 2700-2709

Sangwan R. S., Gauthier, D. A., Turoin, D. H., Pomeroy, M. K., Plaxton, W. C. “Characterization of competent cells and early events of Agrobacterium-mediated genetic transformation in *Arabidopsis thaliana*” Planta (1992) 187, 198-202

Santos-Mendoza, M., Dubreucq, B., Baud, S., Miquel, M., Caboche, M., Lepiniec, L. “LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in *Arabidopsis* leaves” FEBS Letters (2005) 579, 4666-4670

Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., Lepiniec, L. “Deciphering gene regulatory networks that control seed development and maturation in *Arabidopsis*” The Plant Journal (2008) 54, 608-620

Sasaki, Y., Hakamada, K., Suama, Y., Nagano, Y., Furusawa, I., Matsuno, R. “Chloroplast-encoded protein as a subunit of acetyl-CoA carboxylase in pea plant” J Biol Chem (1993) 268, 25118-25123

Sasaki, Y., Konishi, T., Nagano, Y. “The compartmentation of acetyl-coenzyme A carboxylase in plants” Plant Physiol (1995) 108, 445-449

Schmidt, R., Acarkan A., Boivin, K. “Comparative structural genomics in the Brassica family“ Plant Physiol. Biochem. (2001) 39, 253-262

Schwartz, R. M. and Dayhoff M. O., Atlas of Protein Sequence and Structure, 5 (1978), 353–358

Schwender, J. and Ohlrogge, J. B. “Probing in vivo metabolism by stable isotope labeling of storage lipids and proteins in developing *Brassica napus* embryos” Plant Physiol. (2002) 130, 347-361

Schwender, J., Ohlrogge, J. B., Shachar-Hill, Y. J. “A flux model of Glycolysis and the oxidative pentosephosphate pathway in developing *Brassica napus* embryos“ Biol.Chem (2003) 278, 29442-29453

Schwender, J.; Goffman, F.; Ohlrogge, J. B.; Shachar-Hill, Y. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds” Nature (2004) 432, 779-782

Schwender J.; Shachar-Hill, Y.; Ohlrogge, J. B. “Mitochondrial Metabolism in Developing Embryos of Brassica napus” Journal of Biological Chemistry (2006) 281, 34040-34047

Scott, R. J., Spielman, M., Bailey J., Dickinson H. G. “Parent-of-origin effects on seed development in *Arabidopsis thaliana*” Development (1998) 125, 3329-3341

Seyis, F., Friedt, W., Lühs, W. „Resynthese-Raps (Brassica napus L.) als genetische Ressource für die Qualitäts- und Ertragszüchtung“ In: Hammer, K. & Gladis, T. (Hrsg.): Schriften zu Genetischen Ressourcen, (2001) 16, 91–112

Seyis, F., Snowdon, R. J., Luhs, W., Friedt, W., “Molecular characterization of novel resynthesized rapeseed (*Brassica napus*) lines and analysis of their genetic diversity in comparison with spring rapeseed cultivars” Plant Breed. (2003) 122, 473-478

Shengwu, H., Ovesná, J., Kucera, V., Vyvadilová,M. “Evaluation of genetic diversity of *Brassica napus* germplasm from China and Europe assessed by RAPD markers”; Plant Soil Environ. (2003) 49, 106-113

Si, P., Mailer, R. J., Galwey, N., Turner, D. W. “Influence of genotype and environment on oil and protein concentrations of canola (*Brassica napus L.*) grown across southern Australia” Australian Journal of Agricultural Research (2003) 54, 397-407

Slocum, M. K., Fidgore, S. S., Kennard, W. C., Suzuki, J. Y., Osborn, T. C. “Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*” Theor Appl Genet (1990) 80, 57-64

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

Sørensen, M. B., Chaudhury, A. M., Robert, H., Bancharel, E., Berger, F. “Polycomb group genes control pattern formation in plant seed” Current Biology (2001) 11,277–281

Spillane, C., MacDougall, C., Stock C., Köhler, C., Vielle-Calzada, J.-P., Nunes, S. M., Grossniklaus, U., Goodrich, J. “Interaction of the *Arabidopsis* Polycomb group proteins FIE and MEA mediates their common phenotypes” Current Biology (2000) 10, 1535–1538

Stevens, James P. “INTERMEDIATE STATISTICS – A modern approach“ LAWERENCE ERLBAUM ASSOCIATES, PUBLISHERS Hillsdale, New Jersey (1990), 41-74

Stoll, C., Lühs, W., Zarhloul M. K., Brummel, m., Spener, F., Friedt, W. “Knockout of KASIII regulation changes fatty acid composition in canola (*Brassica napus*)“ Eur. J. Lipid Sci. Technol. (2006) 108, 277-286

Stone S. L., Kwong, L. W., Yee, K. M., Pelletier, J., Lepiniec, L., Fischer, R., Goldberg, R. B., Harada, J. J. “LEAFY COTYLEDON 2 encodes a B3 domain transcription factor, that induces embryo development” Proc. Natl. Acad. Sci. USA (2001) 98, 11806-11811

Stone, S. L.; Braybrook, S. A.; Paula, S. L.; Kwong, L. W.; Meuser, J.; Pelletier, J.; Hsieh, T.; Fischer, R. L.; Goldberg, R. B.; Harada, J. J. “*Arabidopsis* LEAFY COTELYDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis” Proc. Natl. Acad. Sci. USA (2008) 105, 3151-3156

Tanksley, S. D. and McCouch, S. R. “Seed banks and molecular maps: Unlocking genetic potential from the wild”; Science (1997) 277, 1063-1066

The Brassica rapa Genome Sequencing Project Consortium. “The genome of the mesopolyploid crop species *Brassica rapa*” Nature genetics (2011) 43, 1035-1040

Thelen, J. J., Mekhedov, S., Ohlrogge, J. B. „*Brassicaceae* express multiple isoforms of biotin carboxyl carrier protein in a tissue specific manner“ Plant Physiol (2001) 125, 2016-2028

Thelen, J. J. and Ohlrogge, J. B. „Metabolic engineering of fatty acid biosynthesis in plants” Metabolic Engineering (2002a) 4, 12-21

Thelen, J. J. and Ohlrogge, J. B. „Both antisense and sense expression of biotin carboxyl carrier protein isoform 2 inactivates the plastid acetyl-coenzyme A carboxylase in *Arabidopsis thaliana*” The Plant Journal (2002b) 32, 419-431

Thomasson, H. J. “The biological value of oils and fats. II. The growth-retarding substance in rapeseed oil” J Nutr (1955) 56, 469-475

Thompson, J. D., Higgins, D. G. and Gibson, T .J. “CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice”. Nucleic Acids Res. (1994) 22, 4673–4680.

Tillmann, P. “Recent experience with NIRS analysis of rapeseed” CGIRC Bull. (1997) 13, 84–87

To, A., Valon, C., Savino, G., Guilleminot, J., Devic, M., Giraudat, J., Parcy, F. “A network of local and redundant gene regulation governs *Arabidopsis* seed maturation” The Plant Cell (2006) 18, 1642-1651

Tovar-Mendez, A.; Miernyk, J. A.; Randall, D. D. “Regulation of pyruvate dehydrogenase complex activity in plant cells” Eur. J. Biochem (2003) 270, 1043-1049

Town, C. D., Cheung, F., Maiti, R., Crabtree, J., Haas, B. J., Wortman, J.R., Hine, E. E., Althoff, R., Arbogast, T. S., Tallon, L. J., Vigouroux, M., Trick, M., Bancroft, I. “Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy” The Plant Cell (2006)18, 1348-1359

Turnham, E. and Northcote, D. H. “Changes in the activity of acetyl-CoA carboxylase during rape-seed formation” Biochem J (1983) 212, 223-229

U “Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization” Japanese Journal of Botany (1935) 7, 389-452

United States Department of Agriculture “Oil crops year book” August 2013; http://www.ers.usda.gov/data-products/oil-crops-yearbook.aspx#.UoXzG-KSE1E

USDA report, from November 2013: “European Union: Oilseeds and Products Supply and Distribution”; http://www.fas.usda.gov/psdonline/

Valentini, G., Chiarelli, L., Fortin R., Speranze, M. L., Galizzi, A. and Mattevi, A. “The allosteric regulation of pyruvate kinase.” J. Biol. Chem. (2000) 275, 18145-18152

Westermeier, P., Wenzel, G., Mohler, V. “Development and evaluation of single-nucleotide polymorphism marker in allotetraploid rapeseed (*Brassica napus L*.)” Theor Appl Genet (2009) 119, 1301-1311

Wood, C. C., Robertson M., Tanner, G., Peacock, W. J., Dennis, E. S., Helliwell C. A. “The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INTENSIVE 3” Proc. Natl. Acad. Sci. USA (2006) 103, 14631-14636

Würschum, T., Maurer, H. P., Dreyer, F., Reif, J. C. “Effect of inter- and intragenic epistasis on the heritability of oil content in rapeseed (*Brassica napus* L.)” Theor Appl Genet (2013) 126, 435-441

Yadegari, R., Kinoshita, T., Lotan, O., Cohen, G., Katz, A., Choi, Y., Katz, A., Nakashima, K., Harada, J. J., Goldberg, R. B., Fischer, R. L., Ohad, N. “Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms” The Plant Cell (2000) 12, 2367-2381

Yamamoto, A.; Kagaya, Y.; Usui, H.; Hobo, T.; Takeda, S.; Hattori, T. “Diverse Roles and Mechanisms of Gene Regulation by the *Arabidopsis* Seed Maturation Master Regulator FUS3 Revealed by Microarray Analysis” Plant Cell Physiol. (2010) 51, 2031-2046

Yang, Y. W., Lai, K. N., Tai, P. Y., Li, W. H. “Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages” J. Mol. Evol. (1999) 48, 597-604

Zhou, Y., Zhang, X., Kang, X., Zhao, X., Zhang, X., Ni, M. „SHORT HYPOCOTYL UNDER BLUE1 Associates with MINISEED3 and HAIKU2 Promoters in Vivo to Regulate *Arabidopsis* Seed Development” The Plant Cell (2009) 21, 106-117

Ziolkowski, P. A., Kaczmarek, M., Babula, D., Sadowski, J. “Genome evolution in *Arabidopsis/Brassica*: conservation in divergence of ancient rearranged segments and their breakpoints” The Plant Journal (2006) 47, 63-74

Zou, J., Zhu, J., Huang, S., Tian, E., Xiao, Y., Fu, D., Tu, J., Fu, T., Meng, J. „Broadening the avenue of intersubgenomic heterosis in oilseed Brassica“ Theor Appl Genet (2010) 120, 283-290

Appendix

Table A.1: Assignment table for the primers published in Cao and Schmidt (2013)

primers corresponding to candidate gene, locus - fragment	correspond to oligonucleotides* shown in Cao and Schmidt (2013) for amplicon
PKP3, 123P16-3	*BnaA.PKp3.a*
PKP3, 123J24-3	BnaC.PKp3.a
WRI, W2-2	BnaA.WRI1.a
WRI, W80-5	BnaC.WRI1.a
WRI, W9-4	BnaA.WRI1.b
WRI, W102-1	BnaC.WRI1.b
PKP1, PPK214-2	BnaA.PKp1.a
PKP1, PPK196-3	BnaC.PKp1.a
PKP1, PPK17-3	BnaC.PKp1.b
FUS3, F161-2	BnaA.FUS3.b
FUS3, F183-3	BnaC.FUS3.b
BCCP2, B62-2	BnaA.BCCP2.a
BCCP2, B51-3	BnaC.BCCP2.a
BCCP2, B173-1	BnaA.BCCP2.b
BCCP2, B114-3	BnaC.BCCP2.b

* oligonucleotide 1 corresponds to the forward primer, oligonucleotide 2 corresponds to the reverse primer used in this study; explanation for gene abbreviations see Table II.3

Table A.2: Assignment table for the reference sequences published by Bach (2007)

candidate gene, locus	label of the corresponding reference sequence (ASS) and open ready frame (ORF), respectively, shown in Bach (2007)
BCCP2, B51	B51
BCCP2, B62	B62
BCCP2, B114	B114
BCCP2, B173	B173
PDH, H25	H25
PDH, H29	H29
PDH, H40	H40
PDH, H71	H71
PDH, H81	H81
PDHK, P12	P12
PDHK, P33	P33
PDHK, P57	P57
PDHK, P133	P133
PKP2, K14	K14
PKP2, K48	K48
PKP2, K141	K141-141
WRI, W2	W2
WRI, W9	W9
WRI, W80	W80
WRI, W102	W102

explanation for gene abbreviations see Table II.3

Table A.3: Fragment position (bp) in the reference sequences of "Express" or "Tapidor"

candidate gene, locus	analyzed fragment	position in the corresponding refrence sequence	candidate gene, locus	analyzed fragment	position in the corresponding refrence sequence
BCCP2, B51	2	3085 - 3858	*WRI, W2*	2	1364 - 1951
	3	3943 - 4670		5	4206 - 4877
BCCP2, B62	1	302 - 894	*WRI, W9*	1	3034 - 3514
	2	1083 - 1856		4	5856 - 6462
	3	1846 - 2571		5	6471 - 7232
BCCP2, B114	1	1659 - 2424	*WRI, W80*	1	1339 - 2157
	2	2389 - 3161		4	3194 - 3904
	3	3151 - 3833		5	4597 - 5251
BCCP2, B173	1	1557 - 2387	*WRI, W102*	1	2231 - 2988
	2	2348 - 3074		2	3835 - 4535
	3	3265 - 3991		5	6268 - 7037
PKP2, K14	1	957 - 2041	*PKP3, 123P16*	1	433 - 1556
	2	2018 - 2857		2	1507 - 2462
	3	2835 - 3719		3	2296 - 3088
PKP2, K48	2	3142 - 4064	*PKP3, 123J24*	1	366 - 1371
	3	4187 - 4788		2	1313 - 2383
	4	5246 - 5945		3	2333 - 3134
PKP2, K141	1	4501 - 5128	*PDHK, P12*	1	290 - 1089
PKP1, PPK17	1	479 - 1288		2	945 - 1745
	2	1056 - 1864		3	1660 - 2510
	3	1922 - 2759	*PDHK, P33*	1	1069 - 1861
PKP1, PPK196	1	481 - 1203		2	1726 - 2491
	2	1583 - 2198		3	2422 - 3049
	3	2114 - 2878	*PDHK, P57*	1	3590 - 4431
	4	2924 - 3969		2	4293 - 5060
PKP1, PPK214	1	388 - 1137		3	4717 - 5710
	2	1045 - 1833	*PDHK, P133*	1	1246 - 2062
	3	1650 - 2438		2	2403 - 3348
	4	2192 - 2912		3	2940 - 3893
	5	2753 - 3806	*PDH, H25*	1	4657 - 5397
LEC2, L65	2	957 - 1876		2	5211 - 6096

explanation for gene abbreviations see Table II.3

Table A.3: Fragment position (bp) in the reference sequences of "Express" or "Tapidor" (continued)

candidate gene, locus	analyzed fragment	position in the corresponding refrence sequence	candidate gene, locus	analyzed fragment	position in the corresponding refrence sequence
LEC2, L65	3	2155 - 2982	*PDH, H25*	3	5948 - 6968
	4	3449 - 4257	*PDH, H29*	1	1259 - 2246
LEC2, L83	1	394 - 1350		2	1751 - 3041
	2a	1955 - 3003		3	2074 - 3477
	3a	3281 - 4147	*PDH, H40*	1	93 - 793
FUS3, F161	1	532 - 1382		2	683 - 1639
	2	1486 - 2306		3	1455 - 2123
FUS3, F183	1b	240 - 790	*PDH, H71*	1	4419 - 5087
	2b	713 - 1638		2	4988 - 5902
	3	1567 - 2277		3	5762 - 6537
IKU2, IKU103	1b	485 - 1470	*PDH, H81*	1	818 - 1629
	2	891 - 1795		2	1272 - 2085
	4	2232 - 2820		3	2206 - 2985
IKU2, IKU124	1	494 - 1468	*FIE, FIE67*	1	541 - 1088
	2	1118 - 1883		2	922 - 1668
	4	2364 - 2989		3b	1472 - 2246
WRI, W2	1	914 - 1686			

explanation for gene abbreviations see Table II.3

Table A.4: Haplotypes of the genotypes in the candidate gene *PKP1*

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *PPK17*									
1, 3, 4, 9, 10, 15, 20, 24, 26, 27, 28, 29, 30	2, 5, 20, 27, 80	"A"	-	-	-	-	-	-	-
2, 5, 7, 11, 12, 13, 14, 16, 17, 19, 21, 22, 23, 25	4, 19, 22, 24, 25, 26, 28, 29, 30, 32, 34, 50, 52, 53, 67, 70	"B"	1	-	-	-	-	-	-
32	-	"C"	1	-	-	4	-	-	-
-	39, 79	"D"	14	1	5	3	-	1	-
locus *PPK196*									
1, 5, 6, 9, 10, 14, 17, 19	2, 4, 5, 19, 22, 24, 25, 26, 27, 29, 30, 32, 34, 39, 53, 70, 79, 80	"A"	-	-	-	-	-	-	-
8, 28, 29, 30, 31	-	"B"	5	1	1	-	-	-	-
11, 18, 26, 27	28, 50, 52, 67	"C"	20	-	2	1	-	-	-
-	20	"D"	16	-	-	-	-	-	-
locus *PPK214*									
26, 27, 30, 31	2, 5, 20, 24, 25, 26, 27, 28, 29, 32, 50, 52, 53, 70	"A"	-	-	-	-	-	-	-
1, 3, 4, 5, 6, 7, 8, 10, 11, 12, 14, 17, 18, 19, 20, 21, 22, 24, 25	-	"B"	22	-	-	5	-	-	-
23	3, 4, 22, 30, 79, 80	"C"	27	-	-	5	-	-	-
9, 28, 29	67	"D"	-	-	-	1	-	-	-
-	34, 39	"E"	37	-	2	14	-	-	-
-	19	"F"	1	-	-	2	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.5: Haplotypes of the genotypes in the candidate gene *LEC2*

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *L65*									
1, 4, 5, 8, 11, 15, 16, 17, 19, 20, 21, 23, 24, 25, 26, 28, 29, 30, 31	2, 5, 19, 20, 22, 25, 26, 27, 28, 29, 30, 32, 39, 50, 52, 53, 67, 79, 80	"A"	-	-	-	-	-	-	-
3	-	"B"	2	1	-	1	-	-	-
7, 12, 13, 14	-	"C"	5	1	-	3	-	-	-
18	-	"D"	7	-	-	5	-	-	-
9, 10	-	"E"	-	-	-	1	-	-	-
-	4, 24	"F"	1	-	-	-	-	-	-
-	34, 70	"G"	1	-	-	1	-	-	-
locus *L83*									
1, 4, 5, 6, 7, 8, 9, 13, 16, 17, 19, 20, 25, 26, 28, 29, 31	2, 4, 5, 20, 22, 24, 25, 26, 27, 28, 29, 30, 32, 34, 39, 50, 52, 53, 67, 70, 79, 80	"A"	-	-	-	-	-	-	-
2, 11, 12, 14, 21	-	"B"	2	1	-	2	-	-	-
3, 10, 15	-	"C"	11	-	-	5	-	-	-
23, 24, 30	-	"D"	4	-	1	2	-	-	-
-	3, 19	"E"	4	-	-	2	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.6: Haplotypes of the genotypes in the candidate gene *FIE*, locus *FIE67*

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
1, 5, 7, 9, 10, 12, 13, 14, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 30, 31	2, 4, 5, 19, 32, 34, 39, 79	"A"	-	-	-	-	-	-	-
3, 4, 18, 28	67	"B"	1	-	1	-	-	-	-
32	-	"C"	2	-	1	2	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.7: Haplotypes of the genotypes in the candidate gene *FUS3*

Exotic genotypes (ID)	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *F161*									
1, 3, 4, 5, 6, 7, 11, 12, 19, 20, 23	2, 4, 5, 20, 22, 24, 26, 27, 28, 29, 32, 34, 39, 50, 52, 53, 67, 70, 79, 80	"A"	-	-	-	-	-	-	-
9, 13, 21, 22, 25, 26, 27, 29, 31	7, 19, 25, 30	"B"	-	-	-	1	-	-	-
17	-	"C"	2	-	-	1	-	-	-
32	-	"D"	1	1	-	2	-	-	-
locus *F183*									
1, 3, 5, 7, 8, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 31	2, 5, 22, 25, 27, 28, 29, 52, 53, 80	"A"	-	-	-	-	-	-	-
6, 9, 11, 26, 27, 28	3, 19, 24, 26, 30, 32, 34, 39, 50, 67, 79	"B"	18	1	3	7	1	-	-
4, 10	-	"C"	3	-	1	2	-	-	-
32	-	"D"	2	-	-	1	-	-	-
2, 12	-	"E"	1	-	-	2	-	-	-
-	4, 20	"F"	1	1	-	-	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.8: Haplotypes of the genotypes in the candidate gene *IKU2*

Exotic genotypes (ID)	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *I103*									
3, 4, 5, 10, 12, 13	2, 4, 5, 20, 24, 25, 26, 27, 28, 29, 30, 34, 39, 50, 52, 53, 67, 70, 79, 80	"A"	-	-	-	-	-	-	-
32	-	"B"	14	1	5	-	-	-	-
6, 9	-	"C"	16	3	6	-	-	-	-
7, 11, 16	-	"D"	5	2	3	-	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.8: Haplotypes of the genotypes in the candidate gene *IKU2* (continued)

Exotic genotypes (ID)	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *I103*									
14, 27, 29, 30	19, 22, 32	"E"	14	2	2	-	-	-	-
18	-	"F"	2	1	1	-	-	-	-
31	-	"G"	4	1	-	-	-	-	-
1, 8, 15	-	"H"	4	1	-	-	-	-	-
28	-	"I"	12	2	2	-	-	-	-
locus *I124*									
1, 10, 14, 20, 21, 23, 24, 25	2, 4, 5, 20, 24, 27, 80	"A"	-	-	-	-	-	-	-
3, 5, 7, 9, 11, 12, 13, 16, 17, 18, 19, 26, 27, 28, 29, 30, 31	3, 19, 22, 25, 26, 28, 29, 30, 32, 34, 39, 50, 52, 53, 67, 70, 79	"B"	8	2	4	-	-	-	-
32	-	"C"	2	-	2	-	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.9: Haplotypes of the genotypes in the candidate gene *PKP3*

Exotic genotypes (ID)	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *123P16*									
1, 3, 4, 5, 7, 14, 16, 17, 19, 23, 26, 27, 29, 30, 31	2, 4, 22, 24, 29, 30, 34, 50, 52, 53, 79	"A"	-	-	-	-	-	-	-
9, 10, 20, 25	-	"B"	34	1	-	6	-	-	-
13	-	"C"	6	-	-	1	-	-	-
8	-	"D"	10	-	3	-	-	-	-
28	-	"E"	-	-	-	1	-	-	-
-	5, 26, 28, 32, 39, 67, 70, 80	"F"	34	1	-	7	-	-	-
-	19	"G"	28	1	3	5	-	-	-
locus *123J24*									
7, 9, 10, 13, 23, 25, 28, 29	2, 4, 5, 19, 32, 34, 39, 67, 79	"A"	-	-	-	-	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.10: Haplotypes of the genotypes in the candidate gene *WRI*

Exotic genotypes (ID)	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *W2*									
1, 2, 3, 4, 5, 6, 8, 10, 12, 13, 14, 15, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31	4, 5, 6, 19, 32, 34, 39, 67, 79	"A"	-	-	-	-	-	-	-
7	-	"B"	10	1	1	4	1	-	-
11	-	"C"	2	-	1	3	1	-	-
16, 32	-	"D"	2	1	-	2	-	-	-
18	-	"E"	2	-	-	-	-	-	-
9	-	"F"	4	-	-	1	-	-	-
locus *W9*									
1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28	4, 5, 19, 32, 34, 39, 67, 79	"A"	-	-	-	-	-	-	-
locus *W80*									
1, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 16, 17, 20, 22, 23, 24, 25, 27	2, 5, 6, 32, 67, 79	"A"	-	-	-	-	-	-	-
28, 29	24	"B"	11	2	1	8	1	-	-
-	4, 19, 34, 39	"C"	-	-	-	1	-	-	-
32	-	"D"	7	3	1	1	1	-	-
locus *W102*									
4, 6, 8, 10, 11, 12, 14, 15, 17, 18, 26, 27, 28, 29, 31	2, 3, 5, 25, 26, 27, 28, 29, 30, 34, 39, 52, 67, 79, 80	"A"	-	-	-	-	-	-	-
2, 13, 16	19, 20, 32, 50, 53, 70	"B"	4	1	1	-	-	-	-
3, 9, 22, 24, 25, 30	22, 24	"C"	8	2	1	-	-	-	-
-	4	"D"	1	-	-	-	-	-	-
1, 7	-	"E"	4	1	1	-	-	-	-
5, 19	-	"F"	2	-	1	-	-	-	-
20, 21	-	"G"	4	1	1	-	-	-	-
32	-	"H"	7	3	1	1	1	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.11: Haplotypes of the genotypes in the candidate gene *BCCP2*

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *B51*									
1, 3, 6, 7, 8, 9, 10, 11, 13, 16, 17, 18, 20, 21, 22, 23, 24, 25, 27, 30	4, 5, 6, 22, 32, 34, 39, 67, 79	"A"	-	-	-	-	-	-	-
12, 19, 26, 28, 29, 31	-	"B"	2	1	-	-	-	-	-
32	-	"C"	4	1	2	-	-	-	-
-	19	"D"	4	-	-	1	-	-	-
locus *B62*									
1, 4, 8, 10, 13, 14, 19	2, 5, 6, 24, 25, 26, 27, 29, 52, 79, 80	"A"	-	-	-	-	-	-	-
3	-	"B"	-	-	-	1	-	-	-
6	-	"D"	8	1	1	5	-	-	-
7, 17	19, 28, 30, 32, 50, 53, 67	"E"	23	-	1	7	-	-	-
9, 18	34	"F"	7	-	1	2	-	-	-
11, 20, 23	4, 14, 20, 22, 39, 70	"G"	3	-	1	3	-	-	-
25, 28, 29	-	"J"	8	-	1	2	-	-	-
locus *B114*									
1, 5, 6, 7, 9, 10, 12, 14, 15, 16, 20, 21, 22, 23, 25, 26, 27, 28, 30, 31	2, 4, 5, 6, 19, 20, 22, 24, 26, 27, 28, 32, 34 39, 50, 52, 53, 67, 70, 79, 80	"A"	-	-	-	-	-	-	-
2, 18	-	"B"	4	2	1	3	1	-	-
8, 11	-	"C"	4	2	1	3	1	-	-
4, 13	25, 29, 30	"E"	-	-	-	1	-	-	-
32	-	"G"	5	1	1	1	-	-	-
locus *B173.*									
1, 3, 4, 5, 10, 11, 12, 13, 14, 15, 16, 17, 19, 21, 27	2, 4, 5, 6, 19, 20, 22, 24, 25, 27, 28, 29, 30, 32, 39, 50, 52, 53, 67, 70, 79	"A"	-	-	-	-	-	-	-
7, 26, 28, 29	-	"B"	26	1	1	10	1	-	-
2, 8, 18	-	"C"	10	2	-	4	1	-	-
-	26, 34, 80	"D"	12	-	-	6	-	-	-
9, 23	-	"E"	12	-	-	5	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.12: Haplotypes of the genotypes in the candidate gene *PDHK*

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *P12*									
7, 9, 13, 23, 25, 28, 29	2, 3, 4, 6, 19, 20, 24, 30, 39, 50, 53, 67, 70, 79	"A"	-	-	-	-	-	-	-
-	5, 26, 27, 28, 29, 32, 34, 52, 80	"B"	-	-	-	1	-	-	-
-	22	"D"	6	-	-	-	-	-	-
locus P33									
7, 9, 10, 13, 23, 25, 28	2, 4, 5, 6, 19, 20, 22, 24, 25, 26, 27, 28, 29, 30, 32, 34, 39, 50, 52, 53, 67, 70, 79	"A"	-	-	-	-	-	-	-
locus *P57*									
1, 3, 7, 8, 13, 17, 27, 28, 29	2, 6, 19, 20, 25, 26, 28, 30, 39, 53, 70, 79, 80	"A"	-	-	-	-	-	-	-
9, 20, 23, 24	32, 52, 67	"B"	1	-	-	-	-	-	-
2, 10, 11, 18, 32	-	"C"	15	-	1	1	-	-	-
-	4, 34, 39	"D"	2	-	1	-	-	-	-
25	-	"E"	1	-	-	-	-	-	-
locus *P133*									
2, 15, 11, 16, 18	2, 19	"A"	-	-	-	-	-	-	-
1, 3, 4, 5, 6, 8, 9, 13, 14, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31	4, 5, 6, 20, 22, 24, 25, 26, 27, 28, 29, 30, 32, 34, 39, 50, 52, 53, 67, 70, 79, 80	"B"	23	-	-	9	-	-	-
32	-	"C"	19	-	-	4	-	-	-
7	-	"D"	25	-	-	11	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.13: Haplotypes of the genotypes in the candidate gene *PDH*

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *H25*									
13, 23, 25, 28, 29	2, 4, 5, 19, 20, 24, 26, 29, 30, 34, 70, 79, 80	"A"	-	-	-	-	-	-	-
locus *H29*									
1, 4, 5, 6, 9, 10, 11, 13, 14, 15, 16, 19, 20, 21, 23, 24, 25, 27	2, 4, 5, 6, 19, 25, 26, 27, 28, 29, 30, 32, 34, 39, 50, 52, 53, 67, 70, 79, 80	"A"	-	-	-	-	-	-	-
-	24	"C"	12	1	3	2	-	-	-
-	22	"D"	1	-	-	-	-	-	-
2, 12, 17, 18	-	"G"	6	-	-	3	-	-	-
7, 26, 28, 29	-	"H"	23	2	3	5	-	-	-
locus *H40*									
1, 4, 5, 6, 7, 12, 13, 14, 16, 17, 20, 22, 23, 25, 26, 27, 28, 29, 31	2, 4, 5, 6, 20, 22, 24, 26, 27, 28, 29, 30, 32, 34, 39, 50, 52, 53, 70, 79, 80	"A"	-	-	-	-	-	-	-
-	19	"B"	9	1	-	6	1	-	-
-	67	"D"	7	-	-	4	1	-	-
19, 21, 24	-	"F"	1	-	-	2	-	1	1
8	-	"H"	3	1	-	3	-	-	-
9	-	"I"	4	1	-	4	1	-	-
30	-	"J"	5	1	-	4	1	-	-
locus *H71*									
7, 9, 10, 13, 28, 29	2, 5, 19, 20, 22, 24, 25, 27, 28, 29, 30, 32, 34, 39, 52, 53, 67, 70, 79, 80	"A"	-	-	-	-	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.13: Haplotypes of the genotypes in the candidate gene *PDH* (continued)

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *H81*									
1, 3, 4, 5, 6, 7, 8, 10, 13, 14, 15, 16, 17, 19, 22	2, 5, 19, 22, 24, 25, 26, 27, 28, 29, 30, 32, 34, 39, 50, 52, 53, 67, 70, 79, 80	"A"	-	-	-	-	-	-	-
28, 29, 31	-	"B	40	-	-	6	-	-	-
27	-	"C"	7	1	-	-	-	-	-
11, 20, 21, 24, 26	6	"D"	5	-	-	-	-	-	-
2, 12, 18	-	"E"	16	-	-	-	-	-	-
-	20	"F"	14	-	-	2	-	-	-
-	4	"G"	12	-	-	2	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.14: Haplotypes of the genotypes in the candidate gene *PKP2*

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *K14*									
20	2, 5, 22, 25, 26, 27, 28, 29, 30, 32, 34, 50, 52, 52, 67, 70, 79, 80	"A"	-	-	-	-	-	-	-
1, 3, 17, 21, 22, 30	-	"B"	2	-	1	2	-	-	-
5, 6, 14, 19	-	"C"	-	-	-	1	-	-	-
8, 18	-	"D"	37	1	1	7	-	-	-
7, 15, 23, 28, 29	-	"E"	6	-	1	2	-	-	-
-	4, 19, 20, 39	"F"	4	-	1	-	-	-	-
9, 10, 13	-	"G"	2	-	1	2	-	-	-
-	24	"I"	1	-	1	-	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.14: Haplotypes of the genotypes in the candidate gene *PKP2* (continued)

Exotic genotypes (ID)*	Winter rapeseed varieties (ID)	Haplotype	Polymorphisms						
			SNP			InDel			
			all	cons AAE	non cons AAE	all	+/- AA	frame shift	stop codon
locus *K48*									
5, 9, 10, 13, 14, 16, 20, 21, 22, 23, 24, 25, 26, 27, 31	2, 3, 22, 29, 34, 53	"A"	-	-	-	-	-	-	-
1, 2, 3, 4, 6, 7, 8, 11, 12, 19, 28, 29	4, 5, 6, 19, 20, 24, 25, 26, 27, 28, 30, 39, 50, 52, 67, 70, 79, 80	"B"	6			-	-	-	-
17, 18	-	"C"	3	-	1	-	-	-	-
-	32	"D"	7	-	-	-	-	-	-
locus *K141*									
3, 4, 6, 8, 10, 11, 12, 15, 16, 21, 22, 26, 30, 31	2, 5, 6, 20, 25, 27, 28, 29, 30, 32, 39, 70, 79	"A"	-	-	-	-	-	-	-
1, 9, 20, 27	4, 19, 24, 26, 34, 80	"B"	2	-	-	-	-	-	-
2, 17, 18	-	"C"	6	-	-	-	-	-	-
-	53	"D"	1	-	-	-	-	-	-
5, 13	-	"G"	7	-	-	-	-	-	-

* genotype names see Table II.1 & Table II.2; cons: conservative; non cons: non-conservative; AAE: amino acid exchange; AA: amino acid

Table A.15: Number of plants and average oil content (%) per genotype per location in the field trials 2007/08

population[1]	candidate gene, locus	Thüle			Biemsen			Göttingen		
		A[2]	B	H	A	B	H	A	B	H
E x H48	*KASIII, S13*	23 (45.47)[3]	17 (44.65)	43 (45.18)	-	-	-	13 (41.84)	21 (43.51)	39 (43.67)
	PKP2, K141	22 (44.78)	15 (45.43)	47 (45.23)	22 (44.99)	21 (44.65)	28 (45.03)	13 (42.98)	17 (44.09)	43 (43.45)
	CAC2, A78	26 (45.45)	19 (44.78)	39 (45.16)	-	-	-	-	-	-
E x H226	*PDHK, P133*	21 (49.60)	21 (49.14)	49 (49.26)	13 (46.37)	10 (46.71)	25 (47.07)	12 (46.68)	15 (46.97)	49 (48.07)
	BCCP2, B173	23 (49.55)	18 (49.82)	50 (49.01)	17 (49.20)	12 (47.56)	33 (47.24)	24 (48.34)	15 (47.20)	40 (47.41)
	KASIII, S13	20 (49.78)	27 (48.52)	45 (49.53)	-	-	-	-	-	-
	CAC2, A78	18 (49.70)	29 (49.79)	44 (48.80)	-	-	-	-	-	-
	CAC3, C10	28 (49.03)	19 (49.93)	41 (49.38)	-	-	-	-	-	-
E x DH7AxM	*KASIII, S13*	29 (45.91)	10 (44.36)	38 (44.16)	5 (44.62)	14 (44.55)	22 (44.12)	10 (39.47)	5 (40.28)	18 (40.96)
S x S4	*PDHK, P57*	23 (48.35)	25 (48.48)	51 (48.35)	22 (44.44)	15 (45.81)	33 (46.68)	23 (48.87)	13 (49.66)	33 (48.48)
	WRI, W80	21 (49.38)	26 (48.22)	45 (48.03)	22 (45.65)	15 (43.87)	32 (46.26)	18 (48.42)	13 (48.71)	29 (49.69)
E x Yellow II	*PDHK, P133*	11 (45.57)	8 (45.70)	23 (44.82)	16 (43.39)	8 (44.85)	32 (45.14)	13 (44.73)	14 (42.40)	31 (43.87)
	PDH, H40	6 (46.10)	8 (46.69)	27 (44.56)	13 (44.74)	13 (43.89)	25 (44.50)	13 (43.31)	11 (43.82)	24 (42.77)
	PKP2, K48	-	-	-	11 (45.11)	14 (43.86)	27 (44.71)	13 (42.71)	8 (43.13)	29 (43.24)
	CAC2, A78	7 (44.58)	8 (46.28)	26 (45.02)	-	-	-	-	-	-
E x Wotan	*PKP2, K48*	36 (47.83)	21 (46.29)	38 (45.89)	17 (45.63)	24 (45.75)	44 (46.04)	33 (45.93)	17 (45.81)	36 (45.97)
	PDHK, P57	21 (45.83)	38 (47.86)	36 (46.03)	19 (46.66)	25 (45.72)	41 (45.61)	29 (45.55)	19 (45.57)	38 (46.39)

[1]E: "Express", S: "Sollux"; [2]A: genotypes homozygous for the "Express" or "Sollux" allele, depending on the cross, B: genotypes homozygous for the novel allele, H: heterozygous genotypes; [3]average oil content in %

Table A.15: Number of plants and average oil content (%) per genotype per location in the field trials 2007/08 (continued)

population[1]	candidate gene, locus	Thüle			Biemsen			Göttingen		
		A[2]	B	H	A	B	H	A	B	H
E x Licapo	*CAC2*, *A78*	26 (45.45)[3]	19 (44.78)	39 (45.16)	-	-	-	-	-	-
	PKP2, *K141*	22 (44.78)	15 (45.43)	47 (45.23)	22 (44.99)	20 (44.99)	28 (45.03)	13 (41.84)	17 (43.85)	41 (43.67)
	KASIII, *S13*	23 (45.47)	17 (44.65)	43 (45.18)	-	-	-	13 (41.84)	20 (43.85)	39 (43.67)
	WRI, *W102*	18 (44.81)	24 (45.56)	42 (45.20)	-	-	-	-	-	-
S x Erox	*KASIII*, *S13*	9 (51.49)	7 (50.01)	24 (50.71)	9 (47.29)	14 (44.55)	22 (44.88)	23 (48.88)	18 (49.45)	41 (49.04)
	BCCP2, *B62*	-	-	-	13 (45.74)	10 (46.17)	19 (45.58)	25 (48.35)	22 (48.72)	43 (49.85)
E x Duell	*PKP2*, *K141*	25 (46.52)	22 (46.66)	49 (47.22)	16 (45.66)	10 (45.34)	19 (46.06)	15 (43.80)	18 (43.34)	35 (44.59)
	WRI, *W102*	23 (46.35)	19 (46.86)	49 (47.15)	18 (45.71)	17 (45.72)	32 (46.50)	17 (43.40)	13 (44.64)	39 (44.10)
	WRI, *W80*	12 (47.57)	27 (47.04)	57 (46.71)	16 (46.03)	21 (45.36)	30 (46.62)	22 (43.92)	19 (43.54)	29 (44.35)

[1]E: "Express", S: "Sollux"; [2]A: genotypes homozygous for the "Express" or "Sollux" allele, depending on the cross, B: genotypes homozygous for the novel allele, H: heterozygous genotypes; [3]average oil content in %

Table A.16: Number of plants and average oil content per genotype per location in the field trials 2008/09

population[1]	candidate gene, locus	Rauischholzhausen			Einbeck			Göttingen		
		A[2]	B	H	A	B	H	A	B	H
E x Altex	*IKU2, I124*	-	-	-	21 (43.95)[3]	25 (44.33)	38 (44.09)	24 (44.38)	35 (44.51)	53 (44.93)
	BCCP2, B173	-	-	-	26 (43.15)	23 (44.91)	34 (44.07)	42 (44.70)	22 (45.02)	81 (44.85)
	WRI, W80	20 (44.06)	24 (44.20)	44 (44.65)	23 (44.51)	24 (43.44)	39 (43.89)	41 (43.99)	28 (45.34)	74 (44.60)
	PKP1, PPK196	-	-	-	14 (44.69)	12 (42.60)	38 (44.51)	22 (44.62)	31 (45.81)	78 (44.63)
E x G39	*PKP2, K14*	17 (42.78)	21 (42.03)	53 (41.77)	48 (42.16)	40 (41.22)	86 (41.70)	26 (45.08)	21 (44.90)	41 (44.50)
	WRI, W2	23 (42.18)	22 (42.07)	46 (41.99)	51 (41.73)	40 (41.03)	83 (42.01)	19 (44.50)	16 (46.23)	45 (44.42)
S x H40	*PKP2, K141*	25 (45.68)	23 (45.70)	46 (45.38)	17 (43.32)	23 (43.33)	46 (44.45)	22 (45.68)	12 (44.10)	50 (44.78)
	LEC2, L65	25 (45.44)	19 (46.34)	45 (45.49)	-	-	-	-	-	-
	PKP3, 123P16	-	-	-	24 (43.08)	24 (44.64)	38 (43.95)	15 (44.12)	23 (45.13)	40 (45.00)
E x H149	*PKP3, 123P16*	-	-	-	24 (45.95)	18 (44.16)	31 (45.28)	15 (47.24)	21 (44.53)	48 (45.56)
S x H365	*PKP2, K14*	-	-	-	19 (45.35)	21 (43.85)	49 (44.94)	68 (45.33)	31 (44.65)	72 (44.60)
	PKP2, K48	-	-	-	18 (44.53)	28 (43.85)	43 (45.47)	42 (44.80)	78 (44.53)	37 (45.59)
	PKP2, K141	-	-	-	7 (46.18)	21 (43.97)	62 (44.79)	36 (44.20)	37 (45.10)	95 (45.08)
	LEC2 L65	-	-	-	22 (45.34)	22 (43.59)	37 (45.02)	37 (44.72)	46 (44.63)	88 (45.15)
	PDHK, P57	-	-	-	34 (44.93)	9 (44.15)	47 (44.65)	52 (44.76)	26 (44.96)	93 (44.95)
S x K29	*PKP2, 123P16*	12 (46.52)	20 (46.59)	55 (47.05)	15 (45.59)	19 (45.76)	37 (45.48)	20 (56.51)	21 (45.04)	45 (46.60)
	IKU2, I103	25 (45.76)	21 (47.10)	47 (47.22)	23 (45.49)	17 (44.47)	47 (45.57)	24 (46.60)	18 (46.20)	41 (45.73)

[1]E: "Express", S: "Sollux"; [2]A: genotypes homozygous for the "Express" or "Sollux" allele, depending on the cross, B: genotypes homozygous for the novel allele, H: heterozygous genotypes; [3]average oil content in %

Table A.16 (continued): Number of plants and average oil content per genotype per location in the field trials 2008/09

population[1]	candidate gene, locus	Rauischholzhausen			Einbeck			Göttingen		
		A[2]	B	H	A	B	H	A	B	H
S x K29	*PDHK, P57*	15 (45.68)[3]	26 (47.36)	52 (47.09)	26 (46.76)	18 (45.29)	42 (44,49)	27 (45.89)	20 (46.32)	38 (46.42)
S x K332	*IKU2, I103*	13 (46.43)	18 (45.45)	48 (46.39)	-	-	-	-	-	-
	IKU2, I103	13 (46.43)	18 (45.45)	48 (46.39)	-	-	-	-	-	-
	LEC2, L83	27 (46.75)	16 (46.49)	34 (45.46)	24 (45.82)	26 (45.19)	32 (45.64)	20 (45.69)	32 (44.96)	34 (45.45)
	WRI, W2	3 (48.67)	22 (46.08)	60 (46.09)	22 (45.27)	16 (46.00)	32 (45.51)	23 (44.70)	18 (45.68)	46 (45.43)
E x L122	*WRI, W2*	24 (46.17)	21 (48.07)	40 (46.80)	17 (45.09)	21 (44.92)	30 (44.46)	17 (45.56)	34 (46.57)	33 (47.13)
E x R76	*WRI, W102*	27 (46.69)	28 (45.50)	36 (45.75)	23 (44.94)	24 (44.80)	41 (44.93)	21 (44.58)	20 (43.44)	46 (44.35)
E x Xiangyou 11	*WRI, W102*	-	-	-	34 (43.26)	12 (44.06)	39 (42.91)	45 (42.62)	34 (42.77)	70 (42.61)
	PK β1, K14	-	-	-	27 (43.29)	20 (43.31)	37 (43.12)	41 (42.38)	47 (43.06)	78 (42.82)
	PDH, H40	-	-	-	29 (43.20)	16 (42.32)	42 (43.55)	61 (43.11)	28 (42.83)	77 (42.66)

[1]E: "Express", S: "Sollux"; [2]A: genotypes homozygous for the "Express" or "Sollux" allele, depending on the cross, B: genotypes homozygous for the novel allele, H: heterozygous genotypes; [3]average oil content in %

Danksagung

Herrn Prof. Dr. Heiko Becker danke ich für die Überlassung des interessanten Themas und die Geduld.

Herrn PD. Dr. Wolfgang Ecke danke ich für das Vertrauen, die Diskussionsbereitschaft und die Geduld.

Rosi Clemens danke ich für ihren tatkräftigen Einsatz bei der ersten DNA-Ernte und ihre moralische Unterstützung.

Birgit Olberg, Gunda Asselmeyer und der „Feld- und Gewächshauscrew“ danke ich für die Hilfe im Gewächshaus und auf dem Feld.

Den Mitarbeitern der Pflanzenzüchtung und der Graslandwissenschaften danke ich für alle gegebenen Tipps und die freundliche Aufnahme am Institut.

Den Partnern im GABI OIL Projekt danke ich für die gute Zusammenarbeit.

Herrn Prof. Dr. Finkeldey danke ich für die Übernahme des Korreferates.

Frau Prof. Dr. Pawelzik, Herrn Prof. Dr. Karlovsky und Herrn Prof. Dr. Feußner danke ich für die Bereitschaft in der Prüfungskommission mitzuwirken.

Ein besonderes Dankeschön gilt meiner Familie für Ihren Glauben an mich, die Unterstützung und Geduld.

Diese Arbeit wurde vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen von GABI FUTURE „OIL - Omics basierte Strategien zur Erhöhung des Ölgehaltes im Raps“ finanziert (Förderkennzeichen 0315053).

Lebenslauf

Persönliche Daten

Name	Daniela Katja Weis, geborene Zeltner
Geburtsdatum	19. April 1982
Geburtsort	Wertheim am Main

Ausbildung/ Beruflicher Werdegang

seit 11/2011	Regierungspräsidium Darmstadt
11/2010 – 11/2011	Landesbetrieb Hessisches Landeslabor
09/2007 – 07/2014	Promotionsstudium am DNPW, Abteilung Pflanzenzüchtung an der Georg- August-Universität Göttingen
09/2007 – 08/2010	Wissenschaftliche Mitarbeiterin am DNPW, Abteilung Pflanzenzüchtung an der Georg- August-Universität Göttingen
05/2006 - 04/2007	Praktisches Jahr im Landesamt für Verbraucherschutz Sachsen-Anhalt Halle (Saale)
	Staatsexamen in Lebensmittelchemie
10/2001 - 04/2006	Studium der Lebensmittelchemie an der Martin-Luther-Universität Halle (Saale)
	Diplom in Lebensmittelchemie
09/1998 - 06/2001	Technisches Gymnasium Wertheim
	Allgemeine Hochschulreife
09/1992 – 06/1998	Realschule Wertheim
	Mittlere Reife
09/1988 – 06/1992	Grundschule Wertheim - Nassig