



# A candidate gene-based association study of tocopherol content and composition in rapeseed (*Brassica napus*)

Steffi Fritsche<sup>1</sup>, Xingxing Wang<sup>2</sup>, Jinqian Li<sup>3</sup>, Benjamin Stich<sup>3</sup>, Friedrich J. Kopisch-Obuch<sup>1</sup>, Jessica Endrigkeit<sup>1</sup>, Gunhild Leckband<sup>4</sup>, Felix Dreyer<sup>4</sup>, Wolfgang Friedt<sup>5</sup>, Jinling Meng<sup>2</sup> and Christian Jung<sup>1\*</sup>

<sup>1</sup> Faculty of Agricultural and Nutritional Sciences, Plant Breeding Institute, Christian-Albrechts-University, Kiel, Germany

<sup>2</sup> National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China

<sup>3</sup> Quantitative Crop Genetics, Max Planck Institute for Plant Breeding Research, Cologne, Germany

<sup>4</sup> Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, Hohenlieth, Germany

<sup>5</sup> Faculty of Agricultural Sciences, Nutritional Sciences and Environmental Management, Institute of Agronomy and Plant Breeding I, Justus-Liebig-University, Giessen, Germany

## Edited by:

Xiaowu Wang, Chinese Academy of Agricultural Sciences, China

## Reviewed by:

Antoni Rafalski, Pioneer Hi-Bred International, A DuPont Business, USA

Jianbing Yan, Huazhong Agricultural University, China

## \*Correspondence:

Christian Jung, Plant Breeding Institute, Christian-Albrechts-University, Olshausenstrasse 40, 24118 Kiel, Germany.  
e-mail: c.jung@plantbreeding.uni-kiel.de

Rapeseed (*Brassica napus* L.) is the most important oil crop of temperate climates. Rapeseed oil contains tocopherols, also known as vitamin E, which is an indispensable nutrient for humans and animals due to its antioxidant and radical scavenging abilities. Moreover, tocopherols are also important for the oxidative stability of vegetable oils. Therefore, seed oil with increased tocopherol content or altered tocopherol composition is a target for breeding. We investigated the role of nucleotide variations within candidate genes from the tocopherol biosynthesis pathway. Field trials were carried out with 229 accessions from a worldwide *B. napus* collection which was divided into two panels of 96 and 133 accessions. Seed tocopherol content and composition were measured by HPLC. High heritabilities were found for both traits, ranging from 0.62 to 0.94. We identified polymorphisms by sequencing selected regions of the tocopherol genes from the 96 accession panel. Subsequently, we determined the population structure (Q) and relative kinship (K) as detected by genotyping with genome-wide distributed SSR markers. Association studies were performed using two models, the structure-based GLM + Q and the PK-mixed model. Between 26 and 12 polymorphisms within two genes (*BnaX.VTE3.a*, *BnaA.PDS1.c*) were significantly associated with tocopherol traits. The SNPs explained up to 16.93% of the genetic variance for tocopherol composition and up to 10.48% for total tocopherol content. Based on the sequence information we designed CAPS markers for genotyping the 133 accessions from the second panel. Significant associations with various tocopherol traits confirmed the results from the first experiment. We demonstrate that the polymorphisms within the tocopherol genes clearly impact tocopherol content and composition in *B. napus* seeds. We suggest that these nucleotide variations may be used as selectable markers for breeding rapeseed with enhanced tocopherol quality.

**Keywords:** *Brassica napus*, tocopherol (vitamin E), candidate genes, association study, SNP identification

## INTRODUCTION

Together with soybean and oil palm, rapeseed (*Brassica napus*, genome AACC,  $2n = 38$ ) belongs to the most important oil crops in the world. Because of its high-quality nutritional composition it is a common source of edible oil. Recently, the focus in rapeseed breeding has turned to improving and altering the content and composition of salutary oil constituents such as carotenoids (Shewmaker et al., 1999; Yu et al., 2008a; Wei et al., 2010), sterols (Amar et al., 2008; Hamama and Bhardwaj, 2011), oleic acid and linolenic acid contents (Rücker and Röbbelen, 1996; Schierholt and Becker, 2001; Zhang et al., 2004; Wittkop et al., 2009), and tocopherols (Marwede et al., 2004; Endrigkeit et al., 2009), the latter being also known as vitamin E.

Rapeseed oil contains high amounts of vitamin E, an essential component in human nutrition and health. A sufficient uptake of vitamin E can help to prevent neurological disorders,

atherosclerosis, cataracts, and cancer (Witztum, 1993; Öhrvall et al., 1996; Schuelke et al., 1999; Sheehy et al., 2000; Schneider, 2005). Vitamin E is synthesized by plants and other photosynthetic organisms. The name is a generic term, which encompasses a group of fat-soluble compounds with antioxidant activity also called tocochromanols (Grusak and DellaPenna, 1999; DellaPenna and Pogson, 2006). The basic structure of the tocochromanols is characterized by a polar chromanol ring and a hydrophobic polyprenyl side chain, products of the shikimate and 1-deoxy-D-xylulose 5-phosphate (DOXP) pathways. Tocochromanols with a fully saturated tail are termed tocopherols, whereas those with an unsaturated tail are termed tocotrienols. The number of methyl groups on the chromanol ring define the four natural occurring tocopherol and tocotrienols forms ( $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ ; Munné-Bosch and Alegre, 2002). With regard to the vitamin E activity,  $\alpha$ -tocopherol has the highest activity and therefore, is the most

important vitamin E form for human nutrition (DellaPenna and Last, 2006).

The tocopherol biosynthetic pathway has been elucidated several years ago (Soll et al., 1980) and genes (*VTE*, loci 1–5; *PDS1*) encoding the respective enzymes of this pathway have been isolated and characterized in *Arabidopsis thaliana* and *Synechocystis* sp. PCC3903 (Norris et al., 1998; Porfirova et al., 2002; Bergmüller et al., 2003; Collakova and DellaPenna, 2003; Van Eenennaam et al., 2003; Valentin et al., 2006). In plants, tocopherols are mainly synthesized in plastids except for the first step which is catalyzed in the cytosol. The major tocopherol form in rapeseed oil is  $\gamma$ -tocopherol followed by  $\alpha$ - and  $\delta$ -tocopherol (Pongracz et al., 1995). The total tocopherol content (TTC) of 87 winter rapeseed genotypes ranged from 182 to 367 mg kg<sup>-1</sup> and was significantly affected by genotype and environment (Goffman and Becker, 1999, 2002). Recently, genetic dissection of tocopherol biosynthesis in crop plants has been done for maize (Wong et al., 2003; Chander et al., 2008), soybean (Li et al., 2010), tomato (Almeida et al., 2011), and sunflower (Haddadi et al., 2011). In rapeseed, between five and seven QTL with additive and/or epistatic effects were mapped for  $\alpha$ -,  $\gamma$ -, and TTC and composition ( $\alpha/\gamma$  ratio) on six linkage groups in a segregating DH population (Marwede et al., 2005). The first gene from *B. napus* involved in tocopherol biosynthesis was cloned by using sequence information of *VTE4* orthologs of *A. thaliana* (Endrigkeit et al., 2009). In that study, the authors verified the function of the cloned *B. napus* gene by an *A. thaliana* transgenic approach leading to a shift in the tocopherol composition in seeds of *BnaA.VTE4.a1* overexpressing plants. Finally, the gene was mapped on *B. napus* chromosome A02 to the position of two QTLs controlling  $\alpha$ -tocopherol content (ATC; Wang et al., in preparation). Linkage mapping is a well-established approach in rapeseed and has become the main tool for identifying genomic regions which contribute to the variation of quantitative traits (Snowdon et al., 2006; Long et al., 2007; Radoev et al., 2008; Zhao et al., 2008; Mei et al., 2009; Chen et al., 2010; Yin et al., 2010; Smooker et al., 2011; Zhang et al., 2011).

In recent years, association studies have become a valuable tool in plant genetics to study the correlation between genetic variants and trait differences based on linkage disequilibrium (LD; Thornsberry et al., 2001; Gupta et al., 2005; Zhu et al., 2008; Hall et al., 2010; Rafalski, 2010). Association studies benefit from the use of genetically diverse germplasm allowing the examination of the total allelic diversity derived from historical and evolutionary recombination events, whereas linkage mapping studies simply exploit the genetic diversity present between two parental genotypes. In rapeseed, marker-trait associations have been identified in several studies using a genome-wide approach for which a large number of markers had to be screened to reach the required density (Hasan et al., 2008; Honsdorf et al., 2010; Zou et al., 2010; Jestin et al., 2011; Rezaeizad et al., 2011). So far, only one candidate gene-based study has been carried out in *B. napus*, investigating the effect of *BnaA.FRI.a* haplotypes on flowering time (Wang et al., 2011).

Up to now, the tocopherol forms  $\alpha$ ,  $\gamma$ , and  $\lambda$  have been determined by high-performance liquid chromatography (HPLC) analysis, an invasive, laborious, and expensive method, which

is not considered to be suitable as routine selection procedure. Therefore, a marker-assisted strategy would be a substantial step forward toward the selection of rapeseed varieties with enhanced tocopherol content and composition and therefore, facilitate the breeding process immensely.

In the present work, we conducted a candidate gene-based association approach to identify and assess the role of polymorphisms in *B. napus* tocopherol biosynthesis genes on tocopherol content and composition. We developed gene-specific primers and sequenced fragments of the candidate genes in a diverse set of rapeseed accessions. By identifying those allelic variations associated with either tocopherol content or composition, promising candidates for the development of molecular markers were detected, verified in a second rapeseed set and can now be used for the selection of rapeseed varieties with enhanced tocopherol qualities.

## MATERIALS AND METHODS

### PLANT MATERIAL AND FIELD EXPERIMENTS

We investigated 229 accessions from a worldwide *B. napus* collection which were divided into two panels of 96 and 133 accessions. The 96 accessions of panel 1 are part of a core collection, established during a European project on genetic diversity in *Brassica* crop species (<http://documents.plant.wur.nl/cgn/pgbr/brasedb/brasresgen.htm>, Table A1 in Appendix). In 2007/2008 panel 1 was grown over winter near the city of Giessen (University of Giessen) in central Germany and near Holtsee in Northern Germany (NPZ Lembke Company, Hohenlieth, Germany). The experiments were performed as a randomized complete block design (RCBD) with two replications and 1.75 m × 2.50 m plots with 100–120 seeds per plot or in case of limited seed availability 50–60 seeds per plot. Seeds were harvested from six to eight open pollinated plants per plot and used for tocopherol and seed quality measurements. Phenotypic data of panel 1 were obtained from 91 *B. napus* accessions grown at both locations. The accessions “Wolynski,” “Ridana,” and “Ramon,” were grown only in Giessen whereas “Tapidor” and “Ningyou 7” were planted only in Holtsee (Table A1 in Appendix).

The second panel 2 consisted of 133 of the 140 *B. napus* accessions which were assessed by Wang et al. (in preparation) and represented a worldwide collection of rapeseed accessions including spring, semi-winter type, and winter type rapeseed cultivars. This panel was grown in 2008/2009 and 2009/2010 at Jingzhou, China (Hubei Province) as a RCBD with three replications and 3 m<sup>2</sup> plots with 30 plants. For panel 2, phenotypic data were obtained for 133 *B. napus* accessions, grown in 2008/2009, and for 109 *B. napus* accessions grown in season 2009/2010.

### TOCOPHEROL AND SEED QUALITY TRAITS MEASUREMENTS

Contents of  $\alpha$ -,  $\gamma$ -, and  $\lambda$ -tocopherol in seeds were determined by HPLC (Schledz et al., 2001; Dähnhardt et al., 2002; Falk et al., 2003). For the extraction, 30–80 mg seeds were disaggregated in 1500  $\mu$ l *n*-heptane. The solution was incubated at –20°C for 2 h and 20  $\mu$ l was used for HPLC analysis. Separation of tocopherols was performed on a silica gel column (5  $\mu$ M LiChrospher® Si 60, Merck) using a mobile phase consisting of

an *n*-heptane/isopropanol-mixture (99 + 1; v + v). Quantification of tocopherols was done by fluorescence detection (excitation at  $\lambda = 290$  nm, emission at  $\lambda = 328$  nm). To identify specific tocopherol forms, the retention times were compared with standards of Merck's tocopherol kit (Merck, Darmstadt, Germany) and for each tocopherol form a calibration was conducted by correlating the concentration of the single forms with the signal output. The concentrations of the analyzed samples in this study were within the linear range of the calibration. Only minor traces of  $\beta$ -tocopherol were obtained, which were not further analyzed during this study. TTC was calculated as the sum of ATC,  $\gamma$ -tocopherol content (GTC), and  $\delta$ -tocopherol content (DTC) and the tocopherol composition was expressed as the ratio of  $\alpha$ - and  $\gamma$ -tocopherol (AGR).

Glucosinolate (GSL), seed oil (SOC), and seed protein (SPC) contents of all 96 panel 1 accessions were measured by near-infrared spectroscopy (NIRS). From each field plot, two subsamples were analyzed. For NIRS measurements 3–5 mg of intact seeds were used. Individual seed spectra from 1100 to 2500 nm were obtained with a NIRSystem 5000 Autocup sampler (Foss, Rellingen, Germany). Internal seed standards were used as control and analyses were done according to the VDLUFA (Kassel, Germany) calibration equation. The tocopherol content in the oil (OTR) was calculated as the ratio of oil and TTC for which the means of each accession was used.

#### DNA EXTRACTION AND GENOTYPIC ANALYSIS

DNA was extracted from panel 1 accessions grown at the location Holtsee from one single plant per plot using the NucleoSpin<sup>®</sup> 96 Plant (4 × 96) kit (Macherey and Nagel, Düren, Germany). The DNA concentration was adjusted to  $5 \text{ ng } \mu\text{l}^{-1}$  using a TECAN-Freedom EVO 150<sup>®</sup> robot (Männedorf, Switzerland).

The 13 tocopherol candidate genes (*BnaX.VTE1.a*, *BnaX.VTE1.b*, *BnaA.VTE2.a*, *BnaX.VTE2.b*, *BnaX.VTE3.a*, *BnaX.VTE3.b*, *BnaA.VTE4.a*, *BnaX.VTE4.b*, *BnaX.VTE4.c*, *BnaC.VTE5*, *BnaX.PDS1.a*, *BnaX.PDS1.b*, and *BnaA.PDS1.c*) were identified by BAC library screening and characterized by functional and mapping approaches (Fritsche et al., in preparation; Wang et al., in preparation). We chose different methods for genotyping each *B. napus* panel. First, we sequenced fragments of the 13 tocopherol candidate genes in panel 1 accessions to identify polymorphisms within these genes. Therefore, extracted DNA of panel 1 accessions was used as PCR template. Gene locus specific primer pairs were developed and tested for different regions of the candidate genes. After amplification, fragments displaying the expected lengths on 1% agarose gels were sequenced by Sanger sequencing (Institute for Clinical Molecular Biology, Kiel, Germany). Only primer pairs producing a single PCR fragment were used for genotyping panel 1, which resulted also in high-quality sequence trace files for each fragment (Table A2 in Appendix). Using DNASTar Lasergene SeqMan Pro 7.2.1 software (Madison, WI, USA), fragments were assembled and the quality of the ABI trace files was analyzed and edited manually by visual examination. We used the TASSEL software (Bradbury et al., 2007) to identify single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) within the sequences of panel 1 accessions. Alignments were constructed with CLC main workbench

5 (CLC bio, Aarhus, Denmark) or the multiple alignment tool CLUSTALW2 (Larkin et al., 2007; Goujon et al., 2010). Comparisons to publicly available sequences were done with the Basic Local Alignment Search Tool (BLAST) from the NCBI website (Altschul et al., 1997).

Second, panel 2 was genotyped with cleaved amplified polymorphic site (CAPS) markers derived from sequence information of panel 1. Restriction enzymes were selected with the restriction site analysis tool implemented in the software CLC main workbench 5 (Table A3 in Appendix).

The population structure of panel 1 was determined by using 31 publicly available genome-wide microsatellite markers (Cheng et al., 2009). For amplification of the microsatellites M13-tailed primers were used (Schuelke, 2000). PCR reactions were performed with four primers: SSR forward primer with M13F-tail, SSR reverse primer with M13R-tail, IRD700-labeled M13F, and un-labeled M13R primers (MWG Biotech, Inc., Ebersberg, Germany). The PCR products were separated on the LI-COR 4300 DNA analyzer system (LI-COR Biosciences, Lincoln, NE, USA). Due to multiple loci amplification or multiple allelic genotypes of the SSR markers, bands were scored as 1 or 0.

#### STATISTICAL ANALYSIS

Mean values of each accession were calculated for each field trial and each trait. For each panel, an analysis of variance (ANOVA) was performed with SAS PRO MIXED version 9.2 (SAS Institute, 2009) to examine the effect of genotype, environment, and genotype × environment interaction on the respective traits and to estimate the variance components.

All factors were treated as random effects using the model:  $Y_{ijkl} = \mu + l_i + b(l)_{ij} + g_k g_{lk} + e_{ijkl}$  where  $y$  is the respective seed trait of the  $k$ th accession tested in the  $j$ th block of the  $i$ th environment,  $\mu$  is the overall mean,  $l_i$  are the effect of the environment,  $b(l)_{ij}$  the block effect,  $g_k$  the effect of the accessions,  $g_{lk}$  the interaction effect between accession and environment and  $e_{ijkl}$  the random experimental error. Heritability was calculated as:  $h^2 = V_g / V_g + V_{g/l} + V_e / R$ , where  $h^2$  is the broad sense heritability,  $V_g$  is the genetic variance of the test panel,  $V_{g/l}$  is the variance of the genotype by environment interaction divided by the number of environments and  $V_e / R$  is the residual variance divided by the total number of replications.

Population structure of panel 1 was examined with the software STRUCTURE version 2.2.3. (Pritchard et al., 2000) using the admixture model and correlated allele frequencies. For between 1 and 10 subpopulations ( $K$ ) the burn-in length period of 100,000 iterations, followed by 100,000 Markov Chain iterations were selected (Figure A1 in Appendix).

Principal component analysis (PCA) was performed based on the above mentioned SSR markers, which were treated as dominant markers, band by band. The first and second principal component was used ( $D$  matrix) for the association analysis.

The kinship coefficient  $K_{ij}$  between inbreds  $i$  and  $j$  were calculated based on the SSR markers according to:  $K_{ij} = S_{ij-1} / 1 + T + 1$ , where  $S_{ij}$  was the proportion of marker loci with shared variants between inbreds  $i$  and  $j$  and  $T$  the average probability that a variant from one parent of inbred  $i$  and a variant from one parent of inbred  $j$  are alike in state, given that they

are not identical by descent (Bernardo, 1993). For the series of  $T$  values 0, 0.025, . . . , 0.975  $K$  matrices between all inbreds were calculated. Negative kinship values between inbreds were set to 0. The optimum  $T$  value was calculated according to Stich et al. (2008).

Population structure of panel 2 was evaluated by Wang et al. (in preparation) based on genotyping the 133 accessions with 41 SSR markers, and was provided as  $Q$ -matrix.

#### LINKAGE DISEQUILIBRIUM AND ASSOCIATION ANALYSIS

$R^2$  values of LD and corresponding  $p$ -values for all loci pairs were calculated using the software R. For LD decay analysis only SNPs with a minimum frequency of 0.05 were considered. Indels were regarded as one polymorphic site. A non-linear regression of  $r^2$  vs. the genetic map distance (cM) was performed (Heuertz et al., 2006).

Polymorphisms were analyzed for association with the following traits: ATC, GTC, TTC, AGR, SOC, GSL, TOC, and SPC. The two models, general linear model (GLM) and PK-mixed model, were used to analyze associations between polymorphic sites and the traits in panel 1. The first model was conducted with TASSEL using the implemented GLM. Analyses were conditioned with population structure estimates, by using the  $Q$ -matrix obtained from the STRUCTURE software. Only polymorphisms with a minor allele frequency of larger than 5% were included in the association analysis. For assuming an association an adjusted  $p$ -value (Bonferroni correction) of less than 0.05 was required. The PK-mixed model was constructed as  $M_{ip} = \mu + a_p + \sum_{u=1}^z D_{iu}v_u + g_i^* + e_{ip}$ , where  $M_{ip}$  was the entry mean of the  $i$ th entry carrying allele  $p$ ,  $a_p$  the effect of allele  $p$ ,  $e_{ip}$  the residual,  $v_u$  the effect of the  $u$ th column of the population structure matrix  $D$ , and  $g_i^*$  the residual genetic effect of the  $i$ th entry (Stich et al., 2008; Yu et al., 2008b). For panel 2, association analysis of polymorphic sites and tocopherol traits was performed using the PK-mixed model. SSR marker data were developed and provided by Wang et al. (in preparation) which were used for population structure and kinship calculations with the same method described before. Accessions with missing phenotypic or genotypic data were excluded from the analysis.

The R package EMMA (Kang et al., 2008) and the significance threshold of 0.05 was applied to perform the above outlined association analysis of all traits with the polymorphisms. Evaluation of the  $p$ -value distribution was done by generating a histogram plot (Figure A2 in Appendix). To test the global hypothesis, the Bonferroni correction was used (Pocock et al., 1987). The percentage of phenotypic variation explained by the significant SNPs was calculated by  $R_{LR}^2 = 1 - \exp(-\frac{2}{n}(\log L_M) - (\log L_0))$ , where  $\log L_M$  is the maximum log-likelihood of the model of interest,  $\log L_0$  the maximum log-likelihood of the intercept-only model, and  $n$  the number of observations (Magee, 1990).

## RESULTS

### PHENOTYPIC VARIATION OF TOTAL TOCOPHEROL CONTENT AND COMPOSITION

In rapeseed panel 1, TTC ranged from 234.63 to 379.10 mg kg<sup>-1</sup> with a mean of 304.14 mg kg<sup>-1</sup> (SD ± 29.17). The mean of TTC in panel 2 was 344.80 mg kg<sup>-1</sup> (SD ± 39.25) with a range of 197.54–460.07 mg kg<sup>-1</sup> (Figure 1A). AGR varied from 0.46 to

1.51 in panel 1 and from 0.33 to 2.14 in panel 2 (Figure 1B). In the ANOVA highly significant ( $p \leq 0.01$ ) effects of genotype and genotype × environment interaction were observed for all traits, except for the genotype × environment interaction effect for AGR in panel 1 (Table 1). High broad sense heritability values were estimated for all traits; from 0.62 to 0.78 for TTC and from 0.77 to 0.94 for AGR (Table 1). The heritability values for ATC and GTC ranged from 0.77 to 0.89.

Seed quality traits such as GSL, SOC, and SPC were measured with accessions from panel 1 in order to unravel any relationship with TTC or AGR. GSL contents ranged from 6.50 to 114.55 μmol g<sup>-1</sup> with a mean of 76.59 (SD ± 25.37). Phenotypic variation was also found for SOC (44.64–58.88% DW) and for SPC (17.25–24.70% DW). We observed high heritability values for all three characters, ranging from 0.63 to 0.98 (Table 2).

ATC and GTC were significantly ( $p < 0.01$ ) related with TTC and AGR (Table 3). Correlations between TTC and AGR as well as between ATC and GTC were not significant. Moreover, the correlation between tocopherol traits and SOC was not significant, whereas a negative correlation was detected between SOC and SPC ( $p < 0.01$ ). All tocopherol traits, except AGR, were significantly ( $p < 0.01$ ) negatively correlated with SPC. Apart from ATC and SOC, the GSL content was significantly ( $p < 0.01$ ) correlated with all other traits.

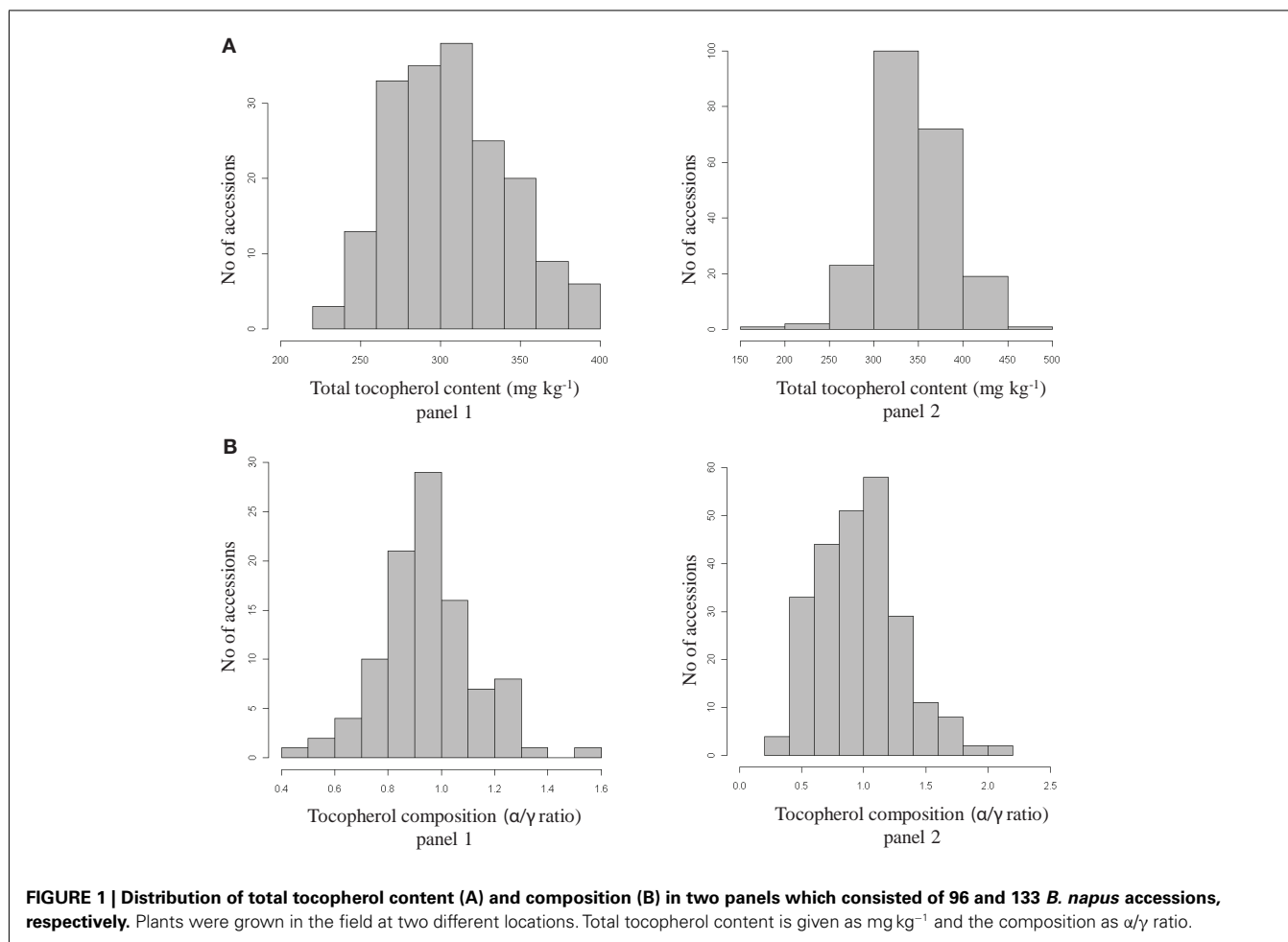
### IDENTIFICATION OF POLYMORPHISMS WITHIN TOCOPHEROL GENES

Single PCR products of the expected fragment size were detected for all 13 candidate genes which were amplified in the panel 1 accessions. However, specific primer pairs yielding high-quality sequences were developed for at least one region in nine genes (Table A2 in Appendix). The fragments of the remaining four candidate genes had poor sequence quality and were not further investigated in the present study. The amplified regions covered between 24.0 and 72.8% of the genes and included exons as well as introns (Table 4).

In summary, the sequencing of fragments of nine candidate genes with a total length of 6640 bp revealed 51 SNPs and 5 indels (Table 5). Taking monomorphic gene fragments into account we observed a density of 1 SNP/130 bp and 1 indel/1328 bp.

The identified polymorphisms were classified according to their minor allele frequency which displayed the frequency at which the less common allele of a polymorphism occurred in the accessions of panel 1. Setting a threshold of 5%, we found polymorphic sites in two candidate genes (*BnaA.PDS1.c*, *BnaX.VTE3.a*) whereas low polymorphic sites (frequency < 5%) were detected in three genes. We found no polymorphisms in the amplified fragments of the remaining four genes (Table A5 in Appendix).

For the gene *BnaA.PDS1.c* we identified in two amplified fragments in total 25 SNPs and three indels within 1033 bp, equivalent to an average density of 1 SNP/41 bp and an indel density of 1 indel/344 bp. Of these, 13 polymorphic sites were located in exons and 15 polymorphic sites within the only intron of this gene. LD with a mean  $r^2$  value of 0.74,  $p < 0.001$  was observed for the *BnaA.PDS1.c* polymorphisms (Figure 2). A LD block (mean  $r^2$  within LD block = 0.92,  $p < 0.001$ ) between SNP 996 and SNP 1250 was found, spanning 254 bp and including the insert region of the gene (Figure 2).



**Table 1 | Ranges, means, ANOVA statistics (components of variance of genotype, genotype  $\times$  environment interaction, and residual error), and heritability estimates of two *B. napus* panels, consisting of 96 (panel 1) and 133 accessions (panel 2), evaluated in field trials for seed  $\alpha$ - and  $\gamma$ -tocopherol content, total tocopherol content ( $\text{mg kg}^{-1}$ ), and tocopherol composition ( $\alpha/\gamma$  ratio).**

Statistics	ATC		GTC		TTC		AGR	
	Panel 1	Panel 2	Panel 1	Panel 2	Panel 1	Panel 2	Panel 1	Panel 2
Min	85.48	83.39	107.76	84.99	234.63	197.54	0.46	0.33
Max	208.29	286.30	196.39	280.83	379.10	460.07	1.51	2.14
Mean	146.74	162.45	157.29	177.66	304.14	344.80	0.94	0.97
SD	21.99	37.39	18.12	36.50	29.17	39.25	0.17	0.35
$V_g$	435.11**	833.01**	246.93**	926.30**	621.13**	840.27**	29,150**	69,390**
$V_{gi}$	66.14**	471.21**	42.34**	487.11**	191.56**	950.03**	1135	40,150**
$V_e$	89.98	27.09	131.79	27.06	318.63	95.21	4983	180.26
$h^2$	0.89	0.77	0.82	0.78	0.78	0.62	0.94	0.77

*F*-statistic: \*significant at  $p \leq 0.05$ ; \*\*significant at  $p \leq 0.01$ .

SD, standard deviation; ATC,  $\alpha$ -tocopherol content; GTC,  $\gamma$ -tocopherol content; TTC, total tocopherol content; AGR,  $\alpha/\gamma$  ratio;  $V_g$ , genetic variance;  $V_{gi}$ , variance of genotype  $\times$  environment interaction;  $V_e$ , residual error;  $h^2$ , broad sense heritability.

In *BnaX.VTE3.a* six SNPs and no indels were identified within 753 bp, which corresponds to an average SNP density of 1 SNP/125 bp. The LD between pairs of SNPs ranged from

0.04 to 1 (Figure 2) with an average  $r^2 = 0.39$  ( $p < 0.001$ ). Two LD blocks were observed from SNP 657 to SNP 741, comprising 84 bp, and from SNP 342 to SNP 359, comprising

**Table 2 | Statistics of the NIRS analysis of 96 *B. napus* accessions (panel 1) with several parameters (ranges, means, components of variance of genotype, genotype × environment interaction and residual error, and broad sense heritability) of seed glucosinolate ( $\mu\text{mol g}^{-1}$ ), protein (% DW), oil content (% DW), and tocopherol in oil (oil/total tocopherol ratio).**

	GSL	SOC	SPC	OTR
Min	6.50	44.68	17.25	1192.57
Max	114.55	58.88	24.70	2300.23
Mean	76.59	51.26	20.77	1694.30
SD	25.37	2.72	1.63	210.37
$V_g$	622.90**	1.21**	4.97**	n.a.
$V_{gi}$	13.49**	0.81**	1.41**	n.a.
$V_e$	22.71	1.18	1.58	n.a.
$h^2$	0.81	0.63	0.98	n.a.

\*Significant at  $p \leq 0.05$ ; \*\*significant at  $p \leq 0.01$ ; n.a., not available; GSL, glucosinolate; SPC, seed protein content; SOC, seed oil content; OTR, oil-tocopherol ratio; SD, standard deviation;  $V_g$ , genetic variance;  $V_{gi}$ , variance of genotype × environment interaction;  $V_e$ , residual error;  $h^2$ , broad sense heritability.

**Table 3 | Correlation coefficients of  $\alpha$ - and  $\gamma$ -tocopherol, total tocopherol content, tocopherol composition, glucosinolate, oil, and protein content of 96 accessions in panel 1.**

Correlation coefficients	GTC	TTC	AGR	SOC	SPC	GSL
ATC	0.13	0.76*	0.67*	0.04	-0.24*	-0.01
GTC		0.74*	-0.59*	0.03	-0.26*	-0.32*
TTC			-0.06	0.04	-0.32*	-0.22*
AGR				0.01	-0.02	0.22*
SOC					-0.51*	0.08
SPC						0.31*

\*Significant at  $p \leq 0.01$ . ATC,  $\alpha$ -tocopherol content; GTC,  $\gamma$ -tocopherol content; TTC, total tocopherol content; AGR,  $\alpha/\gamma$  ratio; GSL, glucosinolate; SPC, seed protein content; SOC, seed oil content.

**Table 4 | Tocopherol biosynthesis genes of *B. napus*, their genomic gene length, amplified gene region, total fragment length, and number of base pairs aligned after sequencing of the gene fragment of panel 1 accessions.**

Gene name	Genbank	Genomic sequence size (bp)	Amplified region	Fragment length (bp)	Sequence length (bp)			ORF coverage (%)	No. of <i>B. napus</i> accessions sequenced
					Coding region	Non-coding region	Total		
<i>BnaX.PDS1.a</i>	JN834026	1390	251–1332	1032	853	67	920	66.2	94
<i>BnaX.PDS1.b</i>	JN834015	1399	402–1199	798	598	0	598	42.7	90
<i>BnaA.PDS1.c</i>	JN834016	1418	1–461	461	304	0	304	72.8	96
			507–1327	821	649	80	729		
<i>BnaX.VTE1.a</i>	JN834017	2809	1707–2746	1040	579	281	860	30.6	92
<i>BnaX.VTE1.b</i>	JN834018	2866	1498–2649	1152	615	413	1028	35.9	86
<i>BnaX.VTE2.b</i>	JN834020	2352	1057–2086	1030	399	381	780	33.2	96
<i>BnaX.VTE3.a</i>	JN834021	1185	190–1045	856	612	141	753	63.5	96
<i>BnaX.VTE3.b</i>	JN834022	1181	48–617	570	435	43	478	40.5	96
<i>BnaX.VTE4.b</i>	JN834023	2062	1429–2008	580	280	214	494	24.0	95

17 bp. In total, 33 polymorphic sites formed 561 pairs of  $r^2$  calculations, of which 59% were observed to have significant LD ( $p < 0.05$ ). Plotting  $r^2$  values against physical distance

(bp) between linked SNP loci pairs indicated that LD decays from 0.45 to 0.25 when physical distance increased to 750 bp (Figure 3).

**Table 5 | Polymorphic sites within tocopherol candidate genes evaluated in 96 *B. napus* accessions (panel 1), their position in the gene and exon/intron-position, predicted amino acid change, and minor allele frequency.**

Position	Polymorphisms	Exon/Intron	Predicted amino acid change	Minor allele frequency (%) <sup>1</sup>
<b><i>BnaX.VTE2.b</i> (1057-2086)<sup>2</sup></b>				
1464 <sup>3</sup>	G/A <sup>4</sup>	exon 8	G → A	4.16
<b><i>BnaX.VTE3.a</i> (190-1045)</b>				
285	T/C	exon 1		16.67
342	C/T	exon 1		10.41
399	T/C	exon 1		9.60
657	G/T	exon 2		15.62
738	T/A	exon 2		15.62
741	T/C	exon 2		16.67
<b><i>BnaX.VTE3.b</i> (48-617)</b>				
326–329	4/0	exon 1	Frame shift	1.78
<b><i>BnaX.PDS1.a</i> (251-1332)</b>				
318	C/A	exon 1		1.06
336	G/T	exon 1		1.06
339	C/T	exon 1		1.06
340	0/9	exon 1	S → RTS	1.06
340	T/C	exon 1	S → A	1.06
342	T/C	exon 1		1.06
360	C/G	exon 1		1.06
456	C/A	exon 1		1.06
460	A/C	exon 1		1.06
687	C/G	exon 1		1.06
720	C/A	exon 1		1.06
800	A/G	exon 1	N → S	1.06
861	C/T	exon 1		1.06
909	C/T	exon 1		1.06
915	G/T	exon 1		1.06
927	G/A	exon 1		1.06
930	C/T	exon 1		1.06
1011	T/C	exon 1		1.06
1029	G/A	exon 1		1.06
1033	A/G	exon 1	I → V	1.06
<b><i>BnaA.PDS1.c</i> (1-461)</b>				
35	A/G	exon 1	Q → R	8.88
50–55	6/0	exon 1	DEA → A	23.40
60	T/A	exon 1		29.67
61	G/A	exon 1	A → T	5.49
174	T/C	exon 1		25.00
183	C/T	exon 1		4.44
207	G/A	exon 1		10.00
213	C/A	exon 1		41.77
297	G/C	exon 1		13.48
300	C/A	exon 1		19.05
<b><i>BnaA.PDS1.c</i> (507-1327)</b>				
543	C/T	exon 1		28.42
996	T/C	exon 1		13.04
1167	C/T	intron 1		14.29
1169	A/T	intron 1		14.29
1170	A/T	intron 1		14.29
1171	G/A	intron 1		14.29
1172	A/C	intron 1		14.29
1173	C/A	intron 1		14.29

Continued

Table 5 | Continued

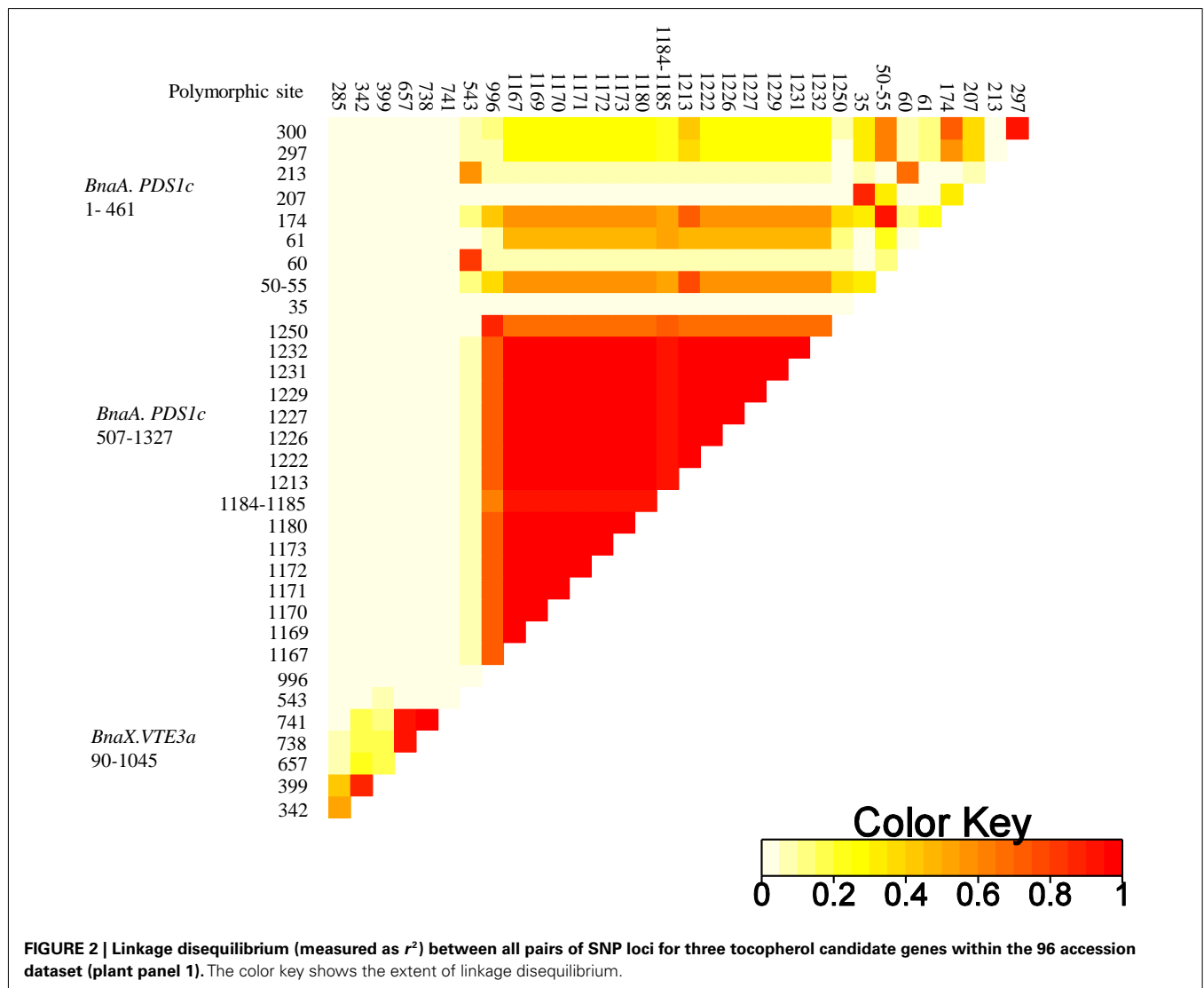
Position	Polymorphisms	Exon/Intron	Predicted amino acid change	Minor allele frequency (%) <sup>1</sup>
1180	A/T	intron 1		14.29
1184–85	2/0	intron 1		13.19
1213	T/A	intron 1		15.11
1222	G/A	intron 1		14.61
1226	C/G	intron 1		14.44
1227	T/A	intron 1		14.44
1229	G/C	intron 1		14.44
1231	1/0	intron 1		14.44
1232	T/C	intron 1		14.44
1250	G/T	exon 2		10.00

<sup>1</sup>Percentage of accessions of panel 1 with the minor allele.

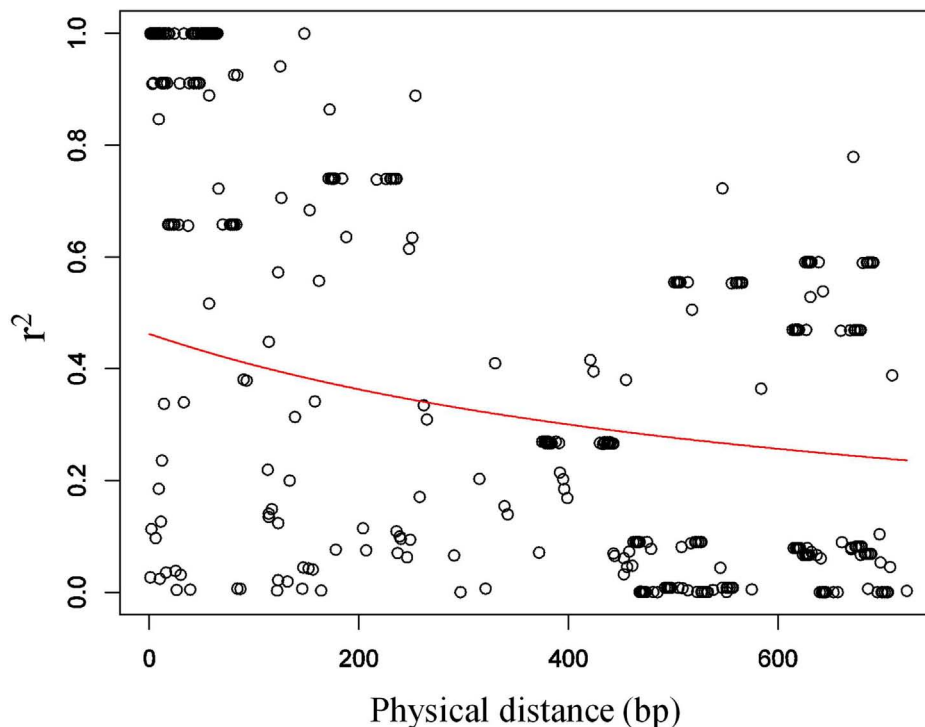
<sup>2</sup>Numbers in parentheses give the gene region amplified, relative to the ATG start codon.

<sup>3</sup>Positions indicate the polymorphic site position in the gene relative to the ATG start codon. Positions of indels refer to the start of the insertion in the *B. napus* cv. *Tapidor* allele of the respective gene.

<sup>4</sup>Numbers or letters preceding the slash point the size of the insertion or nucleotide of the *Tapidor* allele.







**FIGURE 3 |** Plot of linkage disequilibrium measured as squared correlation of allele frequencies ( $r^2$ ) against physical distance (bp) between linked SNP loci pairs from tocopherol biosynthesis genes in the 96 accession dataset. The red line is the non-linear regression trend line of  $r^2$  vs. genetic distance (bp).

**Table 6 |** Genotyping results of panel 2.

Gene	SNP/indel position	No of accessions per allele					Minor allele frequency (%)	
		C	T	G	A	Ins		Del
<i>BnaA.PDS1.c</i>	543	24	99				19.5	
<i>BnaX.VTE3.a</i>	285	10	123				7.5	
<i>BnaX.VTE3.a</i>	342	126	5				3.8	
<i>BnaX.VTE3.b</i>	326–329					0	132	0
<i>BnaX.VTE2.b</i>	1464			2	130			1.5

Polymorphic sites within the genes *BnaA.PDS1.c*, *BnaX.VTE3.a*, *BnaX.VTE3.b*, and *BnaX.VTE2.b* and their occurrence in the 133 accessions panel determined with CAPS markers. Due to missing data accessions do not add up to the total number of 133 accessions for each gene.

ins, insertion; del, deletion.

Panel 2 was genotyped with allele-specific CAPS markers, which were based on polymorphisms of the candidate genes *BnaA.PDS1.c*, *BnaX.VTE3.a*, *BnaX.VTE3.b*, and *BnaX.VTE2.b* (Table A3 in Appendix). CAPS marker analysis enabled the determination of the nucleotide composition at the respective position. In panel 2 the minor allele frequencies of the analyzed SNPs ranged between 1.5 and 19.5% (Table 6). The deletion within *BnaX.VTE3.b* which had been detected in panel 1 was not polymorphic in panel 2 and was therefore excluded from further analysis.

#### POPULATION STRUCTURE

We analyzed the population structure of panel 1 with 31 SSR markers. Of these, seven markers turned out to be monomorphic or

gave ambiguous results and were excluded from further analysis. The remaining 24 SSR loci were polymorphic and resulted in 52 different alleles. The highest likelihood for a subpopulation was obtained with  $K = 4$  and  $\ln p(D) = -1986.6$  and a variance value of 264.4 using the software STRUCTURE (Table 7, Figure 4A). The population structure was also examined by PCA using the same data of the 24 SSR markers. The first and second principal component explained 11.5 and 7.8% of the variations, respectively (Figure 4B). No distinct subgroups were observed.

The kinship coefficient matrices between all accessions were calculated based on the data of the above mentioned 24 SSR markers. The highest kinship coefficient frequency (99.18%) was detected for values between 0 and 0.05, whereas 0.8% of the values

**Table 7 | Population substructure estimation of 96 *B. napus* accessions (panel 1) by evaluating 24 SSR markers using STRUCTURE.**

K	Ln p(D)	SD
1	-2226.2	22.9
2	-2048.9	83.5
3	-1997.6	152.6
4	-1986.6	264.4
5	-2082.8	493.3
6	-2317.0	1011.3
7	-2247.9	894.5
8	-2386.9	1183.7
9	-2424.4	1221.4
10	-2475.2	1246.7

The burn-in length period and the Markov Chain steps were both set to 100,000 iterations in a model allowing admixture and correlated allele frequencies. One to ten subpopulations were tested.

K, no. of populations, Ln p(D), estimated natural logarithm probability of the data; SD, standard deviation.

were above 0.05, indicating that most of the panel 1 accessions had a low level of relatedness (Figure 5). Population structure of panel 2 will be described elsewhere (Wang et al., in preparation).

### ASSOCIATION ANALYSIS

Association analyses were performed with all polymorphic sites (minor allele frequency > 5%) of the candidate genes and the trait data for each field trial of panel 1 by using two models (GLM + Q/ PK-mixed model).

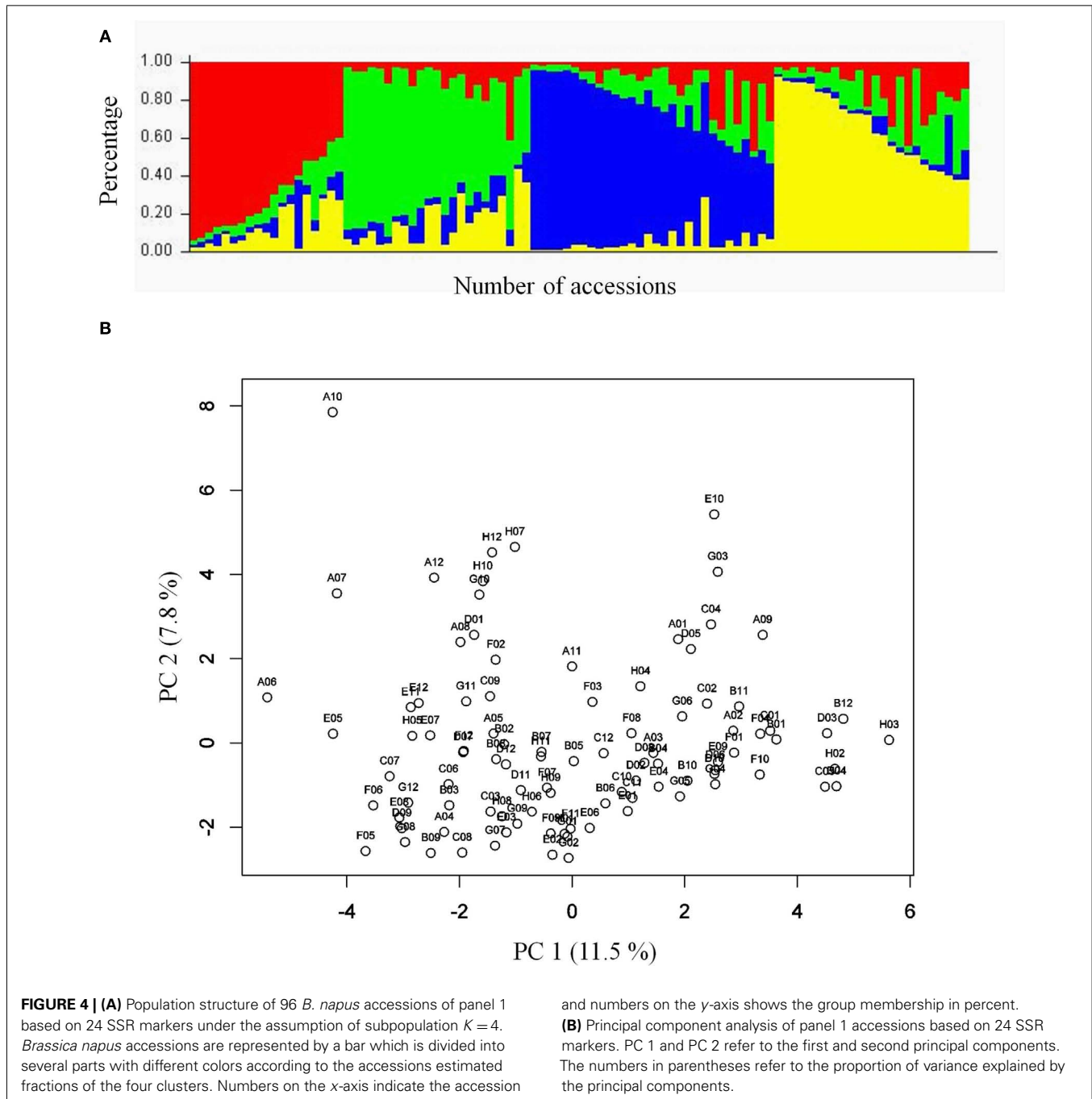
With GLM + Q in total, 53 significant associations ( $p < 0.05$ ) were found for 26 polymorphisms (Table 8). The second model applied during this study, the PK-mixed model, combines the kinship coefficient between individuals with population structure, estimated by PCA. Thus, some significant associations of the GLM + Q were excluded and the result was reduced to 26 significant associations of 12 polymorphisms (Table 8). Considering these results in total, seven polymorphisms within the candidate region of *BnaA.PDS1.c* (position 1–405) were significantly associated with ATC, one of these (SNP 61) was associated with the trait at both field trial locations. We found significant associations with AGR for four SNPs of which SNP 61 was associated with this trait at both field trial sites. Three SNPs (positions 35, 174, 207) were significantly associated with TTC as well as OTR. The indel on position 50–55 was found to be associated with TTC at the field trial site Giessen. Among the quality traits, only SOC was found to be associated with three SNPs in that gene fragment. In the second amplified region of *BnaA.PDS1.c* (positions 507–1327), SNP 543 was found to be significantly ( $p < 0.05$ ) associated with ATC, and TTC in both locations according to both models. The phenotypic variance ( $R^2$ ) explained by that single SNP was between 5.42 and 10.87% in GLM + Q and between 4.96 and 9.09% in PK-mixed model (Table 8). Within *BnaX.VTE3.a* (positions 190–1045), SNP 285 was found to be significantly associated with AGR (Holtsee and Giessen) as well as ATC (Giessen). This SNP explained 6.37–10.05% of the phenotypic variance of the

investigated traits. Another SNP detected at position 342 was associated with the trait GTC (Holtsee). By comparing both models, 29 of the 53 significant associations of the GLM + Q were consistent with the second model. Twenty-four associations of the GLM + Q model were not significant in the PK-mixed model. None of the models found associations to GSL content.

In panel 2 we detected in total five significant associations with three polymorphic sites. SNP 543 within *BnaA.PDS1.c* was significantly associated with AGR in 2009 ( $p = 0.033$ ). On average, all panel 2 accessions with the T-allele had -0.14 AGR (Table 9). Similarly, significant associations between SNP 285 within *BnaX.VTE3.a* and AGR ( $p = 0.014$ ) as well as ATC ( $p = 0.017$ ) in 2009 were found. The effect of the T-allele in panel 2 was on average  $-25.69 \text{ mg kg}^{-1}$   $\alpha$ -tocopherol. SNP 1464 within *BnaX.VTE2.b* was included in the calculations although the allele frequency was found to be 1.5% (Table 6). A significant association of SNP 1464 was found for GTC ( $p = 0.024$ ) and AGR ( $p = 0.035$ ) in 2009.

### DISCUSSION

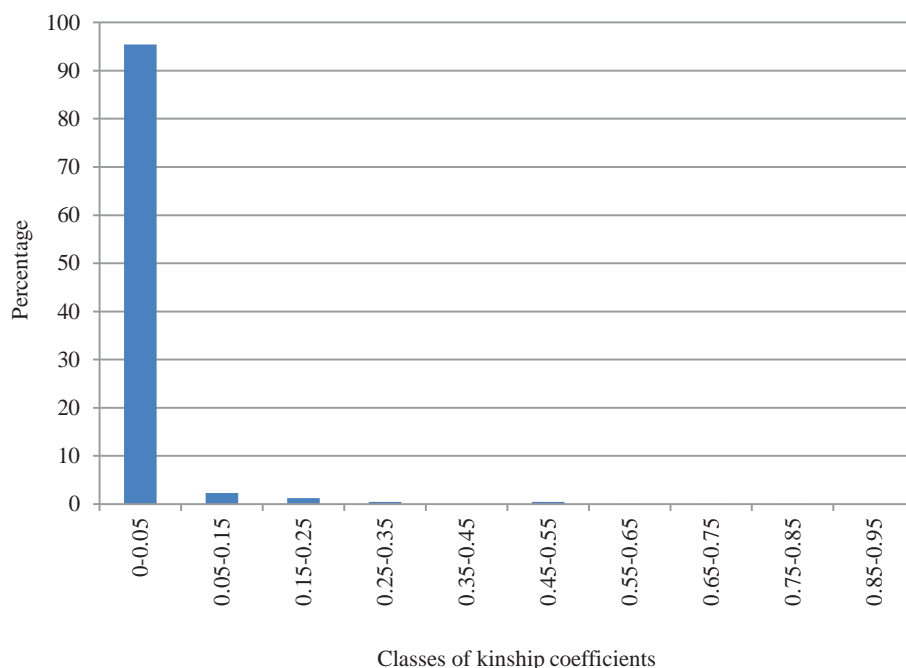
Enhancing the content and composition of tocopherol is one important step to further improve oil quality of rapeseed. In the present study we have demonstrated for the first time an association between tocopherol traits and allelic variations at various candidate gene loci. These polymorphisms represent promising candidates for the development of molecular markers for marker-assisted breeding of rapeseed varieties with enhanced tocopherol qualities. Originally, association studies were developed to dissect the genetics of human diseases but rapidly they have also become an important method in plant genetics to identify alleles and loci responsible for phenotypic trait variation. Association studies can be classified into genome-wide and candidate gene approaches (Zhu et al., 2008). In rapeseed, the genome-wide approach was applied in numerous studies but in none of them the marker density was sufficient for genome-wide association mapping (Hasan et al., 2008; Honsdorf et al., 2010; Zou et al., 2010; Jestin et al., 2011; Rezaeizad et al., 2011). Until now, only one candidate gene-based study has been carried out, investigating the association of *BnaA.FRL.a* haplotypes with flowering time (Wang et al., 2011). The genetic architecture of the tocopherol biosynthetic pathway has been almost completely unraveled in the model species *A. thaliana* (Mène-Saffrané and DellaPenna, 2010), a close relative of *B. napus*. These findings provided the incentive to study tocopherol biosynthesis genes in rapeseed as already performed for oil crops as soybean and sunflower (Li et al., 2010; Dwiyanti et al., 2011; Haddadi et al., 2011) as well as non-oil crops as tomato and maize (Wong et al., 2003; Chander et al., 2008; Almeida et al., 2011). In a separate study, we have identified all genes and several orthologs from the biosynthesis pathway of *B. napus* according to their high sequence homologies to *A. thaliana* genes (Wang et al., in preparation). Further on, we have studied their expression and function and mapped them to *B. napus* linkage groups (Endrigkeit, 2007; Wang et al., in preparation). These data together with partial coincidence between map positions and the positions of major QTL for tocopherol content and composition provided convincing evidence for their role as functional genes for tocopherol biosynthesis in oilseed rape.



There is substantial phenotypic variation for tocopherol traits in *B. napus*. The two diversity sets used in our study (panels 1 and 2) displayed different ranges of variation for tocopherol content and composition. The variation in panel 1 consisting mainly of winter types was comparable to the results obtained by Goffman and Becker (2002), who found a maximum of  $367 \text{ mg kg}^{-1}$  among 87 winter type rapeseed accessions. As expected, genetic variation for tocopherol content ( $197.54\text{--}460.07 \text{ mg kg}^{-1}$ ) and composition ( $0.33\text{--}2.14 \alpha/\gamma$  ratio) was much higher in panel 2, which is possibly due to its different composition (98 winter type and 35 spring

type accessions). This higher genetic variation explains the high heritability estimates in both panels ( $h^2 = 0.62\text{--}0.94$ ) compared to Marwede et al. (2004) ( $h^2 = 0.23\text{--}0.50$ ), who analyzed three doubled haploid populations with a lower genetic variation for tocopherol traits.

Rapeseed is an allopolyploid species, thus it was not surprising to find several homologous sequences for each *A. thaliana* tocopherol gene (Endrigkeit, 2007; Wang et al., in preparation). In total, we analyzed 13 candidate genes for polymorphism screening in panel 1 and used between 5 and 28 primer pairs to amplify parts



**FIGURE 5 |** The distribution of kinship relative coefficients between 96 *B. napus* accessions of panel 1.

of the genes (data not shown). To circumvent the known problems in direct gene sequencing of allopolyploid species, where several orthologous and paralogous gene copies often result in insufficient sequence quality for SNP detection, we used only those primer pairs producing a single PCR fragment and yielding high-quality sequence trace files. This approach was already applied successfully in several previous studies (Ganal et al., 2009; Westermeier et al., 2009; Durstewitz et al., 2010). We found large differences in the density of polymorphisms within the analyzed tocopherol genes and the allele frequencies of these polymorphisms in panel 1. In two genes (*BnaA.PDS1.c*, *BnaX.VTE3.a*) many nucleotide variations were identified, while in the amplified fragments of seven candidate genes no or only rare polymorphisms (frequency < 5%) were detected (Table A5 in Appendix). One possible reason for these findings may be the short and intensive breeding history of rapeseed that has led to a reduced allelic diversity in conventional winter oilseed *B. napus* material (Becker et al., 1995; Hasan et al., 2006; Bus et al., 2011). The discovery of rapeseed varieties with low erucic acid and low GSL content represents major achievements in the rapeseed breeding history but also constituted genetic bottlenecks. Today's spring and winter rapeseed is derived from a limited number of genetic resources, thus most of them share the same genetic background (Friedt and Snowdon, 2009). Panel 1 almost exclusively consisted of winter rapeseed accessions, mainly from Europe; therefore, we decided to use a second panel which encompasses also spring type accessions. A further possible explanation for the low SNP frequency in panel 1 may be the short size of the amplified fragments and the high stringency conditions chosen for obtaining high-quality sequences. Future studies will have to clarify whether sequence variations detected here or any other variations beyond

the amplified regions are the reasons for the observed phenotypic variations.

The SNP density of *BnaA.PDS1.c* (1 SNP/41 bp) was compared with earlier studies in rapeseed. Similar SNP densities were found in EST derived amplicons from 16 rapeseed cultivars (1 SNP/42 bp; Durstewitz et al., 2010) and in the *BnaA.FRI.a* gene after genotyping 95 rapeseed accessions (1 SNP/66 bp; Wang et al., 2011). A considerably lower rate (1 SNP/247 bp) was reported by Westermeier et al. (2009), who surveyed 18 genomic candidate sequences across six rapeseed genotypes. When considering the SNP distribution in the transcriptome of the two parents of the Tapidor and Ningyou7 DH population, which has been frequently used in genetic studies of rapeseed, an overall polymorphism rate between 1 SNP/1.2 kb and 1 SNP/2.1 kb was found (Trick et al., 2009). SNP densities were also calculated for other oil crops such as soybean (1 SNP/273 bp; Zhu et al., 2003), sunflower (1 SNP/69 bp; Fusari et al., 2008), and olive (1 SNP/156 bp; Reale et al., 2006). Taken together, SNP density varies much between species. Thus, comparisons between species should be ideally restricted to orthologous genes (Krutovsky and Neale, 2005). Consequently, we performed an *in silico* comparison of the SNP density of known *A. thaliana* tocopherol loci (Table A4 in Appendix) by using POLYMORPH (Fitz, J. et al., personal communication). We observed SNP density values ranging from 1/30 to 1/624 bp (Table A4 in Appendix), thus demonstrating a broad spectrum of genetic variation within tocopherol biosynthesis genes. Similar to our findings, *A. thaliana* *VTE1* and *VTE2* are less polymorphic than the other genes. In *B. napus*, we found a similar SNP density in *BnaX.VTE3.a* as compared to the *A. thaliana* *VTE3* gene (1/170 bp) but not for *PDS1* (1/415 bp). Interestingly, SNP densities for *A. thaliana* actin genes (*ACT2*, *ACT8*) were 1/194 and 1/203 bp and two randomly

**Table 8 | Polymorphisms of three tocopherol candidate regions of the 96 *B. napus* accession panel significantly associated ( $p < 0.05$ ) with tocopherol and seed quality traits and their percentage on phenotypic variation evaluated by GLM + Q and PK-mixed model.**

Gene	Region	SNP/indel position	Trait	Environment	GLM + Q		PK-mixed model			
					p-Value*	R <sup>2</sup> (%)	p-Value*	R <sup>2</sup> (%)		
<i>BnaA.PDS1.c</i>	1–461	35	OTR	Giessen	0.001	11.41	<0.001	14.72		
			TTC	Giessen	0.011	6.92	0.003	9.75		
			SOC	Giessen	0.030	5.51	0.028	5.64		
			ATC	Giessen	n.s.	n.a.	0.033	5.32		
			TTC	Holtsee	n.s.	n.a.	0.043	4.45		
		50–55	ATC	Giessen	0.040	4.53	0.023	5.72		
			TTC	Giessen	n.s.	n.a.	0.039	4.76		
			60	ATC	Giessen	0.039	4.74	0.031	5.19	
			61	AGR	Giessen	0.003	10.17	0.002	11.21	
				AGR	Holtsee	0.005	9.12	0.008	8.23	
	SOC	Giessen		0.006	8.77	0.019	5.99			
	ATC	Giessen		0.010	7.42	0.005	8.79			
	ATC	Holtsee		0.011	7.36	0.012	7.00			
	SOC	Holtsee		0.029	5.50	n.s.	n.a.			
	GTC	Giessen		0.037	4.89	n.s.	n.a.			
	174	ATC	Giessen	0.009	7.59	0.006	8.46			
		AGR	Giessen	0.009	7.50	0.014	6.85			
		ATC	Holtsee	0.042	4.75	n.s.	n.a.			
		TTC	Giessen	0.048	4.17	0.025	5.73			
		OTR	Giessen	n.s.	n.a.	0.042	4.53			
		207	OTR	Giessen	<0.001	14.43	<0.001	16.93		
			TTC	Giessen	0.005	8.34	0.002	10.48		
			SOC	Giessen	0.008	8.17	0.009	7.83		
			ATC	Giessen	0.013	6.98	0.012	7.26		
		213	SOC	Giessen	0.027	6.58	0.020	7.27		
	SOC		Holtsee	n.s.	n.a.	0.040	5.76			
	297		AGR	Giessen	0.018	6.50	0.046	4.72		
	297	ATC	Giessen	0.026	5.77	0.045	4.74			
		300	AGR	Giessen	0.007	11.98	0.006	9.35		
	AGR		Holtsee	0.024	6.42	n.s.	n.a.			
	ATC		Giessen	n.s.	n.a.	0.008	8.68			
	<i>BnaA.PDS1.c</i>	507–1327	543	ATC	Giessen	0.002	10.87	0.003	9.09	
				ATC	Holtsee	0.004	10.01	0.011	7.26	
				TTC	Giessen	0.006	8.04	0.011	7.11	
				TTC	Holtsee	0.022	6.17	0.030	4.96	
				OTR	Giessen	0.029	5.42	0.025	5.48	
				AGR	Giessen	0.035	5.23	n.s.	n.a.	
				AGR	Holtsee	0.043	4.91	n.s.	n.a.	
				1167	SOC	Giessen	0.043	4.89	n.s.	n.a.
				1169	SOC	Giessen	0.043	4.89	n.s.	n.a.
1170				SOC	Giessen	0.043	4.89	n.s.	n.a.	
1171		SOC	Giessen	0.043	4.89	n.s.	n.a.			
1172		SOC	Giessen	0.043	4.89	n.s.	n.a.			
1173		SOC	Giessen	0.043	4.89	n.s.	n.a.			
1184–85		SOC	Giessen	0.043	4.72	n.s.	n.a.			
1180		SOC	Giessen	0.043	4.89	n.s.	n.a.			
1213		SOC	Giessen	0.002	13.71	n.s.	n.a.			
		OTR	Giessen	0.016	9.12	n.s.	n.a.			

(Continued)

Table 8 | Continued

Gene	Region	SNP/indel position	Trait	Environment	GLM + Q		PK-mixed model	
					p-Value*	R <sup>2</sup> (%)	p-Value*	R <sup>2</sup> (%)
<i>BnaX.VTE3.a</i>	190–1045	285	SPC	Giessen	0.033	7.65	n.s.	n.a.
			ATC	Giessen	0.045	7.10	n.s.	n.a.
			SOC	Giessen	0.049	4.66	n.s.	n.a.
			SOC	Giessen	0.049	4.66	n.s.	n.a.
			SOC	Giessen	0.049	4.66	n.s.	n.a.
			SOC	Giessen	0.040	4.72	n.s.	n.a.
			SOC	Giessen	0.049	4.66	n.s.	n.a.
			AGR	Giessen	0.003	9.50	0.002	10.05
			ATC	Giessen	0.004	8.86	0.002	9.76
			AGR	Holtsee	0.011	7.14	0.016	6.37
			GTC	Holtsee	n.s.	n.a.	0.028	5.35
			GTC	Holtsee	0.049	4.28	n.s.	n.a.

The two field trial locations of panel 1 were analyzed separately due to significant genotype × environment effects.

R<sup>2</sup>, variance explained; ATC, tocopherol content; GTC, γ-tocopherol content; TTC, total tocopherol content; AGR, α/γ ratio; GSL, glucosinolate; SPC, seed protein content; SOC, seed oil content; n.s., not significant; n.a., not available.

\*p-values after Bonferroni multiple test correction.

Table 9 | Association between SNPs and tocopherol traits and allele mean differences of panel 2 accessions.

Gene	SNP position	Trait	Environment <sup>1</sup>	p-Value	Mean difference				R <sup>2</sup> (%)
					C	T	G	A	
<i>BnaA.PDS1.c</i>	543	AGR	2009	0.033	0	−0.14			3.49
<i>BnaX.VTE3.a</i>	285	ATC	2009	0.017	0	−25.69			5.70
		AGR	2009	0.014	0	−0.24			5.66
<i>BnaX.VTE2.b</i>	1464	GTC	2009	0.024			44.29	0	4.51
		AGR	2009	0.035			−0.44	0	4.04

R<sup>2</sup>, variance explained; ATC, α-tocopherol content (mg kg<sup>−1</sup>); GTC, γ-tocopherol content (mg kg<sup>−1</sup>); TTC, total tocopherol content (mg kg<sup>−1</sup>); AGR, α/γ ratio.

<sup>1</sup>For panel 2 we termed the two field trial seasons as environment 2008 or 2009 Figures.

chosen loci (*RPP1*, *LOV1*) involved in defense response exhibited SNP densities of 1/28 and 1/209 bp, respectively. In conclusion, the SNP densities of orthologous genes mainly depend on the genes function and on the plant material chosen for the analysis.

The decay rate of LD with distance is an important parameter determining the resolution of an association study. In our study LD decayed within a physical distance of only 750 bp. The rapid LD decline observed here confirms the power of a candidate gene-based association approach as had been demonstrated for a rapeseed *FRI* homolog tested for association with flowering time (Wang et al., 2011). Comparable to our results, LD decayed rapidly within 2 kb. Other association studies in rapeseed were based on mapped markers, where LD was found to extend over 2 cM in different cultivars (Ecke et al., 2010; Bus et al., 2011) and over 5 cM in parental lines (Würschum et al., 2012). A whole chromosome association approach limited to chromosome A09 revealed an LD extent of 1 cM, when  $r^2$  dropped to 0.2 (Wang et al., in preparation). Altogether, these

data demonstrate that in rapeseed the degree of LD strongly depends on the plant material and the genomic region analyzed as previously examined for other crops (Rafalski, 2002; Jung et al., 2004; Stich et al., 2006; Gore et al., 2009; Ecke et al., 2010).

A further key aspect in conduction of an association approach is the careful consideration of population structure to avoid confounding effects. We applied two different methods (analysis by STRUCTURE software and PCA) with differing results for panel 1. STRUCTURE suggests four subpopulations, indicating a sufficient number of markers for subpopulation calculation, whereas PCA contrasted these findings. No individual subgroups were separated by the principal components 1 and 2, which is probably due to the large proportion of winter rapeseed accessions in panel 1. In comparison, Ecke et al. (2010) could not find population substructure after genotyping 85 winter rapeseed varieties with 89 markers. This reflected the low genetic diversity of winter type rapeseed (Hasan et al., 2006; Bus et al., 2011). We got supporting evidence that the number of background markers were sufficient

for population structure and kinship estimation by calculating the  $p$ -values for the association of the SSRs with the phenotypic traits. The corresponding QQ-plot showed a uniform distribution of the  $p$ -values for most of the SSR loci (Figure A3 in Appendix). This indicates that population structure and kinship are adequately modeled by the markers. In our second panel, three subgroups were detected with STRUCTURE which proved to be in accordance with the results from the PCA (Wang et al., in preparation). The data of panel 2 indicate that winter and spring cultivars of *B. napus* represent genetically distinct groups and corroborate results from former studies (Diers and Osborn, 1994; Hasan et al., 2006).

Because the integration of population structure is considered to be an important factor for straight analysis in association models (Flint-Garcia et al., 2005; Myles et al., 2009; Hall et al., 2010) we decided to integrate two different models in our analysis of which the GLM + Q model was first applied. In order to eliminate spurious associations as a result of relatedness between individuals we also performed the PK-mixed model, which included both population structure (by PCA) and kinship (Yu et al., 2006; Stich et al., 2008; Stich and Melchinger, 2009). The number of associations was corrected when applying the PK-mixed model: we identified an extra of seven significant associations but also enabled us to reduce associations, indicating that kinship can cause confounding effects on associations in panel 1. We used panel 2, which has a higher variation in tocopherol content and composition, to verify results obtained in the first experiment. The panel 2 accessions were genotyped with CAPS markers which were based on SNPs with significant associations in panel 1. First analyses indicate a validation of the previously detected associations of panel 1. Although association analysis of this panel was based on PK-mixed model, not all associations could be confirmed. This might be explainable by the fact that population structure and kinship relationships of this panel is known to contribute strongly to the phenotype variation for all traits (Wang et al., in preparation). Therefore, future association studies with panel 2 will provide further evidence on whether nucleotide variations in the tocopherol candidate genes detected in this study can explain phenotypic variation.

Several SNPs were found to be associated with more than one trait and of those, some were consistently associated with a trait at both field trial sites, indicating an independent environmental effect of these allelic variations. Two of them, the functional SNPs 35 and 61, were located within exon 1 of *BnaA.PDS1.c* and represent non-synonymous substitutions, whereas there was no evidence of SNP 285 and SNP 543 being functional (both are synonymous) or linked to the functional SNPs (Table 5, Figure 2). The associations of the synonymous SNPs to the traits indicate that they are linked to other adjacent causative functional polymorphisms in LD distance. Moreover, silent SNPs can be involved in regulatory functions like alteration of mRNA splicing, stability, and structure and therefore, can affect the structure, function, and expression level of proteins (Chamary and Hurst, 2009; Hunt et al., 2009). However, due to the selection of candidate genes based on their properties like high homology to *A. thaliana* genes, mapping position, or function in

tocopherol biosynthesis, it was more likely to identify associations with tocopherol traits. We identified several polymorphisms in *BnaA.PDS1.c* and *BnaA.VTE3.a* correlated with tocopherol phenotypes and other seed quality traits. Indeed, both genes encode for enzymes required for tocopherol synthesis, e.g., *PDS1* being responsible for an enzyme which catalyzes the formation of homogentisate, an essential substrate for the formation of the aromatic head group of the tocopherol forms. To obtain DMPBQ, a precursor of  $\gamma$ -tocopherol, the enzyme MPBQ/MSBQ methyltransferase, encoded by the gene *VTE3*, is needed. In consistency with our results, *BnaA.PDS1.c* was mapped on chromosome A10 next to a QTL for tocopherol composition in the Tapidor  $\times$  Ningyou7 population (Wang et al., in preparation) whereas *BnaA.VTE3.a* was mapped on chromosome A07 in the Mansholt  $\times$  Samurai population in close proximity to a QTL for  $\alpha$ -tocopherol (Endrigkeit, 2007).

In addition to the tocopherol traits, we analyzed seed quality traits in panel 1 and could also find allelic variants that are associated with SOC or OTR. These results may reflect the fat-solubility property of tocopherols and also, the role of tocopherols in the oxidative stability of oil (Jung and Min, 1990; Isbell et al., 1999; Kamal-Eldin, 2006). Interestingly, the principal constituents and the corresponding pathways needed for the biosynthesis of tocopherol and seed oil are not inter-linked (Somerville et al., 2000; DellaPenna and Pogson, 2006). Nonetheless, the relation between fatty acids, natural components of triglycerides, and phospholipids, with tocopherols was demonstrated by some studies (Hasan and Erbas, 2004; Rani et al., 2007; Richards et al., 2008). Moreover, the mapping position of *BnaA.PDS1.c* in the TN-DH population is located within a QTL region for SOC (Qiu et al., 2006) which gives further evidence for our results.

The association between SNPs and tocopherol content/composition could be also of interest for rapeseed breeders. Tocopherols are essential components of human diet and animal feed and hold important functions in plants, such as the protection of lipids and membranes, response to abiotic stress, and oil stability. Hence, they represent an important target for rapeseed breeding. Rapeseed oil contains high amounts of tocopherol and is therefore an important dietary source. So far, the breeding of rapeseed with higher tocopherol is hampered by the ineffective phenotypic selection procedure by HPLC, a destructive, laborious, and costly method. This study provides rapeseed breeders with molecular markers as a tool for the selection of germplasm with higher tocopherol content and quality.

## ACKNOWLEDGMENTS

The research was supported by the Deutsche Forschungsgemeinschaft (DFG, JU205/14-1) and the Stiftung Schleswig-Holsteinische Landschaft. We thank Jens Hermann from the Institute of Botany and Monika Zuba from NPZ Lembke Company for their excellent technical assistance in HPLC and NIRS analytics and Monika Bruisch for her support in the greenhouse. We are grateful to the Institute for Clinical Molecular Biology, University Kiel, Germany for performing the Sanger-based sequencing.

## REFERENCES

- Almeida, J., Quadrana, L., Asis, R., Setta, N., de Godoy, F., Bermúdez, L., Otaiza, S. N., Corrêa da Silva, J. V., Fernie, A. R., Carrari, F., and Rossi, M. (2011). Genetic dissection of vitamin E biosynthesis in tomato. *J. Exp. Bot.* 62, 3781–3798.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- Amar, S., Ecke, W., Becker, H. C., and Moellers, C. (2008). QTL for phytoosterol and sinapate ester content in *Brassica napus* L. collocate with the two erucic acid genes. *Theor. Appl. Genet.* 116, 1051–1061.
- Becker, H. C., Engqvist, G. M., and Karlsson, B. (1995). Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. *Theor. Appl. Genet.* 91, 62–67.
- Bergmüller, E., Porfirova, S., and Dörmann, P. (2003). Characterization of an *Arabidopsis* mutant deficient in  $\gamma$ -tocopherolmethyltransferase. *Plant Mol. Biol.* 52, 1181–1190.
- Bernardo, R. (1993). Estimation of coefficient of coancestry using molecular markers in maize. *Theor. Appl. Genet.* 85, 1055–1062.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635.
- Bus, A., Körber, N., Snowdon, R., and Stich, B. (2011). Patterns of molecular variation in a species-wide germplasm set of *Brassica napus*. *Theor. Appl. Genet.* 123, 1413–1423.
- Chamary, J. V., and Hurst, L. D. (2009). The price of silent mutations. *Sci. Am.* 300, 46–53.
- Chander, S., Guo, Y. Q., Yang, X. H., Yan, J. B., Zhang, Y. R., Song, T. M., and Li, J. S. (2008). Genetic dissection of tocopherol content and composition in maize grain using quantitative trait loci analysis and the candidate gene approach. *Mol. Breed.* 22, 353–365.
- Chen, G., Geng, J., Rahman, M., Liu, X., Tu, J., Fu, T., Li, G., McVetty, P. B. E., and Tahir, M. (2010). Identification of QTL for oil content, seed yield, and flowering time in oilseed rape (*Brassica napus*). *Euphytica* 175, 161–174.
- Cheng, X. M., Xu, J. S., Xia, S., Gu, J. X., Yang, Y., Fu, J., Qian, X. J., Zhang, S. C., Wu, J. S., and Liu, K. (2009). Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*. *Theor. Appl. Genet.* 118, 1121–1131.
- Collakova, E., and DellaPenna, D. (2003). Homogentisate phytyltransferase activity is limiting for tocopherol biosynthesis in *Arabidopsis*. *Plant Physiol.* 131, 632–642.
- Dähnhardt, D., Falk, J., Appel, J., van der Kooij, T. A., Schulz-Friedrich, R., and Krupinska, K. (2002). The hydroxyphenylpyruvate dioxygenase from *Synechocystis* sp. PCC 6803 is not required for plastoquinone biosynthesis. *FEBS Lett.* 523, 177–181.
- DellaPenna, D., and Last, R. L. (2006). Progress in the dissection and manipulation of plant vitamin E biosynthesis. *Physiol. Plant* 126, 356–368.
- DellaPenna, D., and Pogson, B. (2006). Vitamin synthesis in plants: tocopherols and carotenoids. *Annu. Rev. Plant Biol.* 57, 711–738.
- Diers, B. W., and Osborn, T. C. (1994). Genetic diversity of oilseed *Brassica napus* germplasm based on restriction fragment length polymorphisms. *Theor. Appl. Genet.* 88, 662–668.
- Durstewitz, G., Polley, A., Plieske, J., Luerssen, H., Graner, E. M., Wieseke, R., and Ganai, M. W. (2010). SNP discovery by amplicon sequencing and multiplex SNP genotyping in the allopolyploid species *Brassica napus*. *Genome* 53, 948–956.
- Dwiyanti, M., Yamada, T., Sato, M., Abe, J., and Kitamura, K. (2011). Genetic variation of gamma-tocopherol methyltransferase gene contributes to elevated alpha-tocopherol content in soybean seeds. *BMC Plant Biol.* 11, 152. doi:10.1186/1471-2229-11-152
- Ecke, W., Clemens, R., Honsdorf, N., and Becker, H. C. (2010). Extent and structure of linkage disequilibrium in canola quality winter rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 120, 921–931.
- Endrigkeit, J. (2007). *Identifikation und Charakterisierung von Genen der Tocopherol-Biosynthese aus Raps (Brassica napus L.)*. Dissertation, Christian-Albrechts-University, Kiel.
- Endrigkeit, J., Wang, X. X., Cai, D. G., Zhang, C. Y., Long, Y., Meng, J. L., and Jung, C. (2009). Genetic mapping, cloning, and functional characterization of the BnaX.VTE4 gene encoding a gamma-tocopherol methyltransferase from oilseed rape. *Theor. Appl. Genet.* 119, 567–575.
- Falk, J., Andersen, G., Kernebeck, B., and Krupinska, K. (2003). Constitutive overexpression of barley 4-hydroxyphenylpyruvate dioxygenase in tobacco results in elevation of the vitamin E content in seeds but not in leaves. *FEBS Lett.* 540, 35–40.
- Flint-Garcia, S. A., ThUILlet, A. C., Yu, J. M., Pressoir, G., Romero, S. M., Mitchell, S. E., Doebley, J., Kresovich, S., Goodman, M. M., and Buckler, E. S. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44, 1054–1064.
- Friedt, W., and Snowdon, R. J. (2009). “Oilseed rape,” in *Handbook of Plant Breeding Vol. 4: Oil Crops*, eds J. Vollmann and I. Rajcan (New York: Springer), 91–126.
- Fusari, C. M., Lia, V. V., Hopp, H. E., Heinz, R. A., and Paniego, N. B. (2008). Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. *BMC Plant Biol.* 8, 7. doi:10.1186/1471-2229-8-7
- Ganal, M. W., Altmann, T., and Roder, M. S. (2009). SNP identification in crop plants. *Curr. Opin. Plant Biol.* 12, 211–217.
- Goffman, F. D., and Becker, H. C. (1999). “Inheritance of tocopherol contents in seeds of rapeseed (*Brassica napus* L.),” in *Proceedings of the 10th Rapeseed Congress*, Canberra.
- Goffman, F. D., and Becker, H. C. (2002). Genetic variation of tocopherol content in a germplasm collection of *Brassica napus* L. *Euphytica* 125, 189–196.
- Gore, M. A., Chia, J.-M., Elshire, R. J., Sun, Q., Ersoz, E. S., Hurwitz, B. L., Peiffer, J. A., McMullen, M. D., Grills, G. S., Ross-Ibarra, J., Ware, D. H., and Buckler, E. S. (2009). A first-generation haplotype map of maize. *Science* 326, 1115–1117.
- Goujon, M., McWilliam, H., Li, W. Z., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* 38, W695–W699.
- Grusak, M. A., and DellaPenna, D. (1999). Improving the nutrient composition of plants to enhance human nutrition and health. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50, 133–161.
- Gupta, P., Rustgi, S., and Kulwal, P. (2005). Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol. Biol.* 57, 461–485.
- Haddadi, P., Ebrahimi, A., Langlade, N., Yazdi-samadi, B., Berger, M., Calmon, A., Naghavi, M., Vincourt, P., and Sarrafi, A. (2011). Genetic dissection of tocopherol and phytoosterol in recombinant inbred lines of sunflower through quantitative trait locus analysis and the candidate gene approach. *Mol. Breed.* 29, 1–13.
- Hall, D., Tegström, C., and Ingvarsson, P. K. (2010). Using association mapping to dissect the genetic basis of complex traits in plants. *Brief. Funct. Genomics* 9, 157–165.
- Hamama, A., and Bhardwaj, H. (2011). Characterization of total and individual sterols in canola sprouts. *J. Am. Oil Chem. Soc.* 88, 361–366.
- Hasan, B., and Erbas, S. (2004). Influence of seed development and seed position on oil, fatty acids and total tocopherol contents in sunflower (*Helianthus annuus* L.). *Turk. J. Agr. Forest.* 29, 179–186.
- Hasan, M., Friedt, W., Pons-Kühnemann, J., Freitag, N., Link, K., and Snowdon, R. J. (2008). Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *napus*). *Theor. Appl. Genet.* 116, 1035–1049.
- Hasan, M., Seyis, F., Badani, A., Pons-Kühnemann, J., Friedt, W., Lühs, W., and Snowdon, R. (2006). Analysis of genetic diversity in the *Brassica napus* L. gene pool using SSR markers. *Genet. Resour. Crop Evol.* 53, 793–802.
- Heuertz, M., De Paoli, E., Kallman, T., Larsson, H., Jurman, I., Morgante, M., Lascoux, M., and Gyllenstrand, N. (2006). Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. *Genetics* 174, 2095–2105.
- Honsdorf, N., Becker, H. C., and Ecke, W. (2010). Association mapping for phenological, morphological, and quality traits in canola quality winter rapeseed (*Brassica napus* L.). *Genome* 53, 899–907.
- Hunt, R., Sauna, Z. E., Ambudkar, S. V., Gottesman, M. M., and Kimchi-Sarfaty, C. (2009). Silent (synonymous) SNPs: should we care about them? *Methods Mol. Biol.* 578, 23–39.
- Isbell, T. A., Abbott, T. P., and Carlson, K. D. (1999). Oxidative stability index of vegetable oils in binary mixtures with meadowfoam oil. *Ind. Crops Prod.* 9, 115–123.
- Jestin, C., Lodé, M., Vallée, P., Domin, C., Falentin, C., Horvais, R., Coedel, S., Manzanares-Dauleux, M., and Delourme, R. (2011). Association mapping of quantitative resistance



- for *Leptosphaeria maculans* in oilseed rape (*Brassica napus* L.). *Mol. Breed.* 27, 271–287.
- Jung, M., Ching, A., Bhatramakki, D., Dolan, M., Tingey, S., Morgante, M., and Rafalski, A. (2004). Linkage disequilibrium and sequence diversity in a 500-kbp region around the *adh1* locus in elite maize germplasm. *Theor. Appl. Genet.* 109, 681–689.
- Jung, M. Y., and Min, D. B. (1990). Effects of  $\alpha$ -,  $\gamma$ -, and  $\delta$ -tocopherols on oxidative stability of soybean oil. *J. Food Sci.* 55, 1464–1465.
- Kamal-Eldin, A. (2006). Effect of fatty acids and tocopherols on the oxidative stability of vegetable oils. *Eur. J. Lipid Sci. Technol.* 108, 1051–1061.
- Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., and Eskin, E. (2008). Efficient control of population structure in model organism association mapping. *Genetics* 178, 1709–1723.
- Krutovskiy, K. V., and Neale, D. B. (2005). Nucleotide diversity and linkage disequilibrium in cold-hardiness- and wood quality-related candidate genes in Douglas Fir. *Genetics* 171, 2029–2041.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and clustal X version 2.0. *Bioinformatics* 23, 2947–2948.
- Li, H. Y., Liu, H. C., Han, Y. P., Wu, X. X., Teng, W. L., Liu, G. F., and Li, W. B. (2010). Identification of QTL underlying vitamin E contents in soybean seed among multiple environments. *Theor. Appl. Genet.* 120, 1405–1413.
- Long, Y., Shi, J., Qiu, D., Li, R., Zhang, C., Wang, J., Hou, J., Zhao, J., Shi, L., Park, B. S., Choi, S. R., Lim, Y. P., and Meng, J. (2007). Flowering time quantitative trait loci analysis of oilseed Brassica in multiple environments and genomewide alignment with *Arabidopsis*. *Genetics* 177, 2433–2444.
- Magee, L. (1990).  $R^2$  measures based on Wald and likelihood ratio joint significance tests. *Am. Stat.* 44, 250–253.
- Marwede, V., Gul, M. K., Becker, H. C., and Ecke, W. (2005). Mapping of QTL controlling tocopherol content in winter oilseed rape. *Plant Breed.* 124, 20–26.
- Marwede, V., Schierholt, A., Christian, M., and Becker, H. C. (2004). Genotype  $\times$  environment interactions and heritability of tocopherol contents in canola. *Crop Sci.* 44, 728–731.
- Mei, D. S., Li, Y. C., Wang, H. Z., Hu, Q., Li, Y. D., and Xu, Y. S. (2009). QTL analysis on plant height and flowering time in *Brassica napus*. *Plant Breed.* 128, 458–465.
- Mène-Saffrané, L., and DellaPenna, D. (2010). Biosynthesis, regulation and functions of tocopherols in plants. *Plant Physiol. Biochem.* 48, 301–309.
- Munné-Bosch, S., and Alegre, L. (2002). The function of tocopherols and tocotrienols in plants. *CRC Crit. Rev. Plant Sci.* 21, 31–57.
- Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., and Buckler, E. S. (2009). Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21, 2194–2202.
- Norris, S. R., Shen, X., and DellaPenna, D. (1998). Complementation of the *Arabidopsis* *pds1* mutation with the gene encoding p-hydroxyphenylpyruvate dioxygenase. *Plant Physiol.* 117, 1317–1323.
- Öhrvall, M., Vessby, B., and Sundlöf, G. (1996). Gamma, but not alpha, tocopherol levels in serum are reduced in coronary heart disease patients. *J. Intern. Med.* 239, 111–117.
- Pocock, S. J., Geller, N. L., and Tsiatis, A. A. (1987). The analysis of multiple endpoints in clinic trials. *Biometrics* 43, 487–498.
- Pongracz, G., Weiser, H., and Matziger, D. (1995). Tocopherole-Antioxidantien der Natur. *Eur. J. Lipid Sci. Technol.* 3, 90–104.
- Porfirova, S., Bergmüller, E., Tropf, S., Lemke, R., and Dörmann, P. (2002). Isolation of an *Arabidopsis* mutant lacking vitamin E and identification of a cyclase essential for all tocopherol biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12495–12500.
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multi-locus genotype data. *Genetics* 155, 945–959.
- Qiu, D., Morgan, C., Shi, J., Long, Y., Liu, J., Li, R., Zhuang, X., Wang, Y., Tan, X., Dietrich, E., Weihmann, T., Everett, C., Vanstraelen, S., Beckett, P., Fraser, F., Trick, M., Barnes, S., Wilmer, J., Schmidt, R., Li, J., Li, D., Meng, J., and Bancroft, I. (2006). A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. *Theor. Appl. Genet.* 114, 67–80.
- Radoev, M., Becker, H. C., and Ecke, W. (2008). Genetic analysis of heterosis for yield and yield components in rapeseed (*Brassica napus* L.) by QTL mapping. *Genetics* 179, 1547–1558.
- Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* 5, 94–100.
- Rafalski, J. A. (2010). Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13, 174–180.
- Rani, A., Kumar, V., Verma, S., Shakya, A., and Chauhan, G. (2007). Tocopherol content and profile of soybean: genotypic variability and correlation studies. *J. Am. Oil Chem. Soc.* 84, 377–383.
- Reale, S., Doveri, S., Diaz, A., Angiolillo, A., Lucentini, L., Pilla, F., Martin, A., Donini, P., and Lee, D. (2006). SNP-based markers for discriminating olive (*Olea europaea* L.) cultivars. *Genome* 49, 1193–1205.
- Rezaeizad, A., Wittkop, B., Snowdon, R., Hasan, M., Mohammadi, V., Zali, A., and Friedt, W. (2011). Identification of QTLs for phenolic compounds in oilseed rape (*Brassica napus* L.) by association mapping using SSR markers. *Euphytica* 177, 335–342.
- Richards, A., Wijesundera, C., and Salisbury, P. (2008). Genotype and growing environment effects on the tocopherols and fatty acids of *Brassica napus* and *B. juncea*. *J. Am. Oil Chem. Soc.* 85, 159–168.
- Rücker, B., and Röbbelen, G. (1996). Impact of low linolenic acid content on seed yield of winter oilseed rape (*Brassica napus* L.). *Plant Breed.* 115, 226–230.
- Schierholt, A., and Becker, H. C. (2001). Environmental variability and heritability of high oleic acid content in winter oilseed rape. *Plant Breed.* 120, 63–66.
- Schledz, M., Seidler, A., Beyer, P., and Neuhaus, G. (2001). A novel phytyltransferase from *Synechocystis* sp. PCC 6803 involved in tocopherol biosynthesis. *FEBS Lett.* 499, 15–20.
- Schneider, C. (2005). Chemistry and biology of vitamin E. *Mol. Nutr. Food Res.* 49, 7–30.
- Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18, 233–234.
- Schuelke, M., Mayatepek, E., Inter, M., Becker, M., Pfeiffer, E., Speer, A., Hübner, C., and Finckh, B. (1999). Treatment of ataxia in isolated vitamin E deficiency caused by [alpha]-tocopherol transfer protein deficiency. *J. Pediatr.* 134, 240–244.
- Sheehy, P. J. A., Bramley, P. M., Elmadfa, I., Kafatos, A., Kelly, F. J., Manios, Y., Roxborough, H. E., Schuch, W., and Wagner, K. H. (2000). Vitamin E. *J. Sci. Food Agric.* 80, 913–938.
- Shewmaker, C. K., Sheehy, J. A., Daley, M., Colburn, S., and Ke, D. Y. (1999). Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. *Plant J.* 20, 401–412.
- Smooker, A. M., Wells, R., Morgan, C., Beaudoin, F., Cho, K., Fraser, F., and Bancroft, I. (2011). The identification and mapping of candidate genes and QTL involved in the fatty acid desaturation pathway in *Brassica napus*. *Theor. Appl. Genet.* 122, 1075–1090.
- Snowdon, R., Lühs, W., and Friedt, W. (2006). “Oilseed rape,” in *Genome Mapping and Molecular Breeding, Vol. 2, Oilseeds*, ed. C. Kole (Heidelberg: Springer).
- Soll, J., Kemmerling, M., and Schultz, G. (1980). Tocopherol and plastoquinone synthesis in spinach chloroplasts subfractions. *Arch. Biochem. Biophys.* 204, 544–550.
- Somerville, C., Browse, J., Jaworski, J. G., and Ohlrogge, J. B. (2000). “Lipids,” in *Biochemistry and Molecular Biology of Plants*, eds B. B. Buchanan, W. Gruissem, and R. L. Jones (Rockville: American Society of Plant Physiologists), 456–527.
- Stich, B., Maurer, H. P., Melchinger, A. E., Frisch, M., Heckenberger, M., van der Voort, J. R., Peleman, J., Sorensen, A. P., and Reif, J. C. (2006). Comparison of linkage disequilibrium in elite European maize inbred lines using AFLP and SSR markers. *Mol. Breed.* 17, 217–226.
- Stich, B., and Melchinger, A. E. (2009). Comparison of mixed-model approaches for association mapping in rapeseed, potato, sugar beet, maize, and *Arabidopsis*. *BMC Genomics* 10, 94. doi:10.1186/1471-2164-10-94
- Stich, B., Mohring, J., Piepho, H. P., Heckenberger, M., Buckler, E. S., and Melchinger, A. E. (2008). Comparison of mixed-model approaches for association mapping. *Genetics* 178, 1745–1754.
- Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen, D., and Buckler, E. S. (2001). Dwarf8 polymorphisms associate with variation in flowering time. *Nat. Genet.* 28, 286–289.
- Trick, M., Long, Y., Meng, J., and Bancroft, I. (2009). Single nucleotide polymorphism (SNP) discovery in the polyploid *Brassica napus* using Solexa transcriptome sequencing. *Plant Biotechnol. J.* 7, 334–346.
- Valentin, H. E., Lincoln, K., Moshiri, E., Jensen, P. K., Qi, Q., Venkatesh, T. V., Karunanandaa, B., Baszisz, S. R., Norris, S. R., Savidge, B.,

- Gruys, K. J., and Last, R. L. (2006). The *Arabidopsis* vitamin E pathway gene5-1 mutant reveals a critical role for phytol kinase in seed tocopherol biosynthesis. *Plant Cell* 18, 212–224.
- Van Eenennaam, A., Lincoln, K., Durrett, T., Valentin, H., Shewmaker, C., Thorne, G., Jiang, J., Baszis, S., Levering, C., Aasen, E., Hao, M., Stein, J., Norris, S., and Last, R. (2003). Engineering vitamin E content: from *Arabidopsis* mutant to soy oil. *Plant Cell* 15, 3007–3019.
- Wang, N., Qian, W., Suppanz, I., Wei, L., Mao, B., Long, Y., Meng, J., Müller, A. E., and Jung, C. (2011). Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the FRIGIDA homologue BnaA.FRI.a. *J. Exp. Bot.* doi: 10.1093/jxb/err249
- Wei, S., Yu, B., Gruber, M. Y., Khachatourians, G. G., Hegedus, D. D., and Hannoufa, A. (2010). Enhanced seed carotenoid levels and branching in transgenic *Brassica napus* expressing the *Arabidopsis* miR156b gene. *J. Agric. Food Chem.* 58, 9572–9578.
- Westermeier, P., Wenzel, G., and Mohler, V. (2009). Development and evaluation of single-nucleotide polymorphism markers in allotetraploid rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 119, 1301–1311.
- Wittkop, B., Snowdon, R. J., and Friedt, W. (2009). Status and perspectives of breeding for enhanced yield and quality of oilseed crops for Europe. *Euphytica* 170, 131–140.
- Witzum, J. L. (1993). Role of oxidised low density lipoprotein in atherosclerosis. *Br. Heart J.* 69, S12–S18.
- Wong, J. C., Lambert, R. J., Tadmor, Y., and Rocheford, T. R. (2003). QTL associated with accumulation of tocopherols in maize. *Crop Sci.* 43, 2257–2266.
- Würschum, T., Liu, W., Maurer, H., Abel, S., and Reif, J. (2012). Dissecting the genetic architecture of agronomic traits in multiple segregating populations in rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 124, 153–161.
- Yin, X., Yi, B., Chen, W., Zhang, W., Tu, J., Fernando, W., and Fu, T. (2010). Mapping of QTLs detected in a *Brassica napus* DH population for resistance to *Sclerotinia sclerotiorum* in multiple environments. *Euphytica* 173, 25–35.
- Yu, B., Lydiate, D., Young, L., Schäfer, U., and Hannoufa, A. (2008a). Enhancing the carotenoid content of *Brassica napus* seeds by downregulating lycopene epsilon cyclase. *Transgenic Res.* 17, 573–585.
- Yu, J. M., Holland, J. B., McMullen, M. D., and Buckler, E. S. (2008b). Genetic design and statistical power of nested association mapping in maize. *Genetics* 178, 539–551.
- Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., and Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208.
- Zhang, H., Shi, C., Wu, J., Ren, Y., Li, C., Zhang, D., and Zhang, Y. (2004). Analysis of genetic effects and heritabilities for linoleic and  $\alpha$ -linolenic acid content of *Brassica napus* L. across Chinese environments. *Eur. J. Lipid Sci. Technol.* 106, 518–523.
- Zhang, L., Yang, G., Liu, P., Hong, D., Li, S., and He, Q. (2011). Genetic and correlation analysis of silique-trait in *Brassica napus* L. by quantitative trait locus mapping. *Theor. Appl. Genet.* 122, 21–31.
- Zhao, J., Dimov, Z., Becker, H. C., Ecke, W., and Moellers, C. (2008). Mapping QTL controlling fatty acid composition in a doubled haploid rapeseed population segregating for oil content. *Mol. Breed.* 21, 115–125.
- Zhu, C., Gore, M., Buckler, E. S., and Yu, J. (2008). Status and prospects of association mapping in plants. *Plant Genome* 1, 5–20.
- Zhu, Y. L., Song, Q. J., Hyten, D. L., Van Tassel, C. P., Matukumalli, L. K., Grimm, D. R., Hyatt, S. M., Fickus, E. W., Young, N. D., and Cregan, P. B. (2003). Single-nucleotide polymorphisms in soybean. *Genetics* 163, 1123–1134.
- Zou, J., Jiang, C. C., Cao, Z. Y., Li, R. Y., Long, Y., Chen, S., and Meng, J. L. (2010). Association mapping of seed oil content in *Brassica napus* and comparison with quantitative trait loci identified from linkage mapping. *Genome* 53, 908–916.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 13 March 2012; accepted: 30 May 2012; published online: 26 June 2012.

Citation: Fritsche S, Wang X, Li J, Stich B, Kopisch-Obuch FJ, Endrigkeit J, Leckband G, Dreyer F, Friedt W, Meng J and Jung C (2012) A candidate gene-based association study of tocopherol content and composition in rapeseed (*Brassica napus*). *Front. Plant Sci.* 3:129. doi: 10.3389/fpls.2012.00129

This article was submitted to *Frontiers in Plant Genetics and Genomics*, a specialty of *Frontiers in Plant Science*.

Copyright © 2012 Fritsche, Wang, Li, Stich, Kopisch-Obuch, Endrigkeit, Leckband, Dreyer, Friedt, Meng and Jung. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.

## APPENDIX

**Table A1 | *Brassica napus* accessions which were used for field trials at Giessen and Holtsee (Germany) 2007/08.**

Accession name	Geographical origin	Type	Accession name	Geographical origin	Type
Abukuma natane	Japan	Winter	Major	France	Winter
Akela	Germany	Winter	Mansholts Hamb. Raps	Germany	Winter
Anja	Germany	Winter	Marasaki natane	Japan	Winter
Aphid resistant rape	New Zealand	Winter	Markus	France	Winter
Askari	Germany	Winter	Matador	Sweden	Winter
Baltia	Soviet Union	Winter	Mestnij	Soviet Union	Winter
Belinda	Germany	Winter	Moana, Moana Rape	New Zealand	Winter
Bienvenue	France	Winter	Mytnickij	Ukraine	Winter
Binera	Germany	Winter	Nemertschanskij 1	Soviet Union	Winter
Bladkool	Netherlands	Winter	Norde	Sweden	Winter
Brauner Schnitkohl	Germany	Winter	Norin	unknown	Winter
Brink	Sweden	Winter	Nunsdale	United Kingdom	Winter
Ceres	Germany	Winter	Oleander	Germany	Winter
Chuoshenshu	Korea	Winter	Ölquell, Gülzower	Germany	Winter
Coriander	Germany	Winter	Palu	Italy	Winter
Darmor	France	Winter	Panter	Sweden	Winter
Diamant	Germany	Winter	Parapluie	France	Winter
Dippes	Germany	Winter	POH 285, Bolko	Poland	Winter
Doral	Germany	Winter	Primor, P-R	France	Winter
Dwarf Essex	United Kingdom	Winter	Quedlinburger Platzfester	Germany	Winter
Edita	Germany	Winter	Quinta	Germany	Winter
Elsoms giant	United Kingdom	Winter	Rafal, R 40	France	Winter
Emerald	Germany	Winter	Ramses	France	Winter
English giant	United Kingdom	Winter	Rapol	Germany	Winter
Erra	Germany	Winter	Regal	Sweden	Winter
Fertoedi	Hungary	Winter	Samo	Sweden	Winter
Fora	Sweden	Winter	Sarepta	France	Winter
Gogatsuna	Japan	Spring	Siberische Boerenkool	unknown	Winter
Goldgelber Zarter Butter	Germany	Winter	Silona	Sweden	Winter
Groene Groninger Snijmoes	Netherlands	Winter	Skrzeszowicki	Poland	Winter
Gross-Luesewitzer Spaets	unknown	Winter	Skziverskij	Soviet Union	Winter
Hektor	Sweden	Winter	Slovenska Krajova	Czech Republic	Winter
Hokkai 3-go	Japan	Winter	Sobotkowski	Poland	Winter
Janetzki Schlesischer	Austria	Winter	Sonnengold	Germany	Winter
Janpol	Poland	Winter	Start	Poland	Winter
Jet Neuf	France	Winter	Taisetsu	Japan	Winter
Jupiter	Sweden	Winter	Trebicska	Czech Republic	Winter
Krapphauser	unknown	Winter	Victor	Sweden	Winter
Kromerska	Czech Republic	Winter	Vinnickij 15/59	Soviet Union	Winter
Librador	Germany	Winter	Winfred	Germany	Winter
Libritta	Germany	Winter	Express	Germany	Winter
Liglory (Gelbsamige)	Germany	Winter	Tapidor <sup>1</sup>	France	Winter
Limburgse Bladkool	Netherlands	Winter	Ningyou <sup>1</sup>	Viet-Nam	Spring
Lingot, R-26	France	Winter	Ramon <sup>2</sup>	Netherlands	Winter
Liporta	Germany	Winter	Ridana <sup>2</sup>	Germany	Winter
Lirafit, WRF 22	Germany	Winter	Wolynski <sup>2</sup>	Soviet Union	Winter
Liragrün	Germany	Winter			
Lirakotta	Germany	Winter			
Madora	Germany	Winter			
MAH 1, Jantar	Poland	Winter			

*They are a part of the core collection of the EU project RESGEN CT99 109-112.*

<sup>1</sup>Accession grown only in Holtsee.

<sup>2</sup>Accession grown only in Giessen.

**Table A2 | Primer used for the amplification of tocopherol genes.**

Gene	Genbank nr	Primer sequence (5' → 3')
<i>BnaX.PDS1.a</i>	JN834026	F: ACTCGGAATGCGATTCTCCGCT R: GTAAACCTTCCCTTCCTCATCC
<i>BnaX.PDS1.b</i>	JN834015	F: CTCAGCATCTAATCAACGTAGCT R: CCCTCTATCGTCCTAAACGAC
<i>BnaA.PDS1.c</i>	JN834016	F1: AACTCTATGGGGCACGAAAA R1: TCTGCGTCTTCTACTTCAAC F2: CCCTCTATCGTCCTAAACGAC R2: CATCCCTCCACTCTGGTAAACTC
<i>BnaX.VTE1.a</i>	JN834017	F: CTGAAGAGACCGTTTGAGTAGC R: TTTTCTGTATTTGAGCCTCAT
<i>BnaX.VTE1.b</i>	JN834018	F: AACCAGAAGTTAGTTCATGGC R: CCTCAGTGGACTAGCTAAGC
<i>BnaX.VTE2.b</i>	JN834020	F: CATCTTCACTAGTGCTTAGATC R: GCCTTAATCACTCAAACGAGG
<i>BnaX.VTE3.a</i>	JN834021	F: GGTTTCATCCAGCACAGAAAA R: CCTCTCCTTTGGTCCAAGCTA
<i>BnaX.VTE3.b</i>	JN834022	F: GGTTTCCCGCTTCCAATCT R: CTAACCCAATCACACTCTGA
<i>BnaX.VTE4.b</i>	JN834023	F: CATTGAGTCTTCGTTGTGCAAT R: CCCTCAATCATCAATGG

**Table A3 | Polymorphisms within tocopherol candidate genes and enzymes, which were used as allele-specific markers for genotyping the 133 *B. napus* accessions of panel 2.**

Gene	Polymorphism	Position in gene	Enzyme
<i>BnaA.PDS1.c</i>	C/T <sup>1</sup>	543	<i>BseDI</i>
<i>BnaX.VTE3.a</i>	T/C	285	<i>Hin1II</i>
<i>BnaX.VTE3.a</i>	C/T	342	<i>SaI</i>
<i>BnaX.VTE3.b</i>	CCGG/–	326–329	<i>Kpn2I</i>
<i>BnaX.VTE2.b</i>	G/A	1464	<i>BbvI</i>

<sup>1</sup>Letter preceding the slash points the nucleotides of the *Tapidor* allele.

**Table A4 | SNP densities with in *A. thaliana* tocopherol genes by using POLYMORPH (Fitz, J. et al., personal communication).**

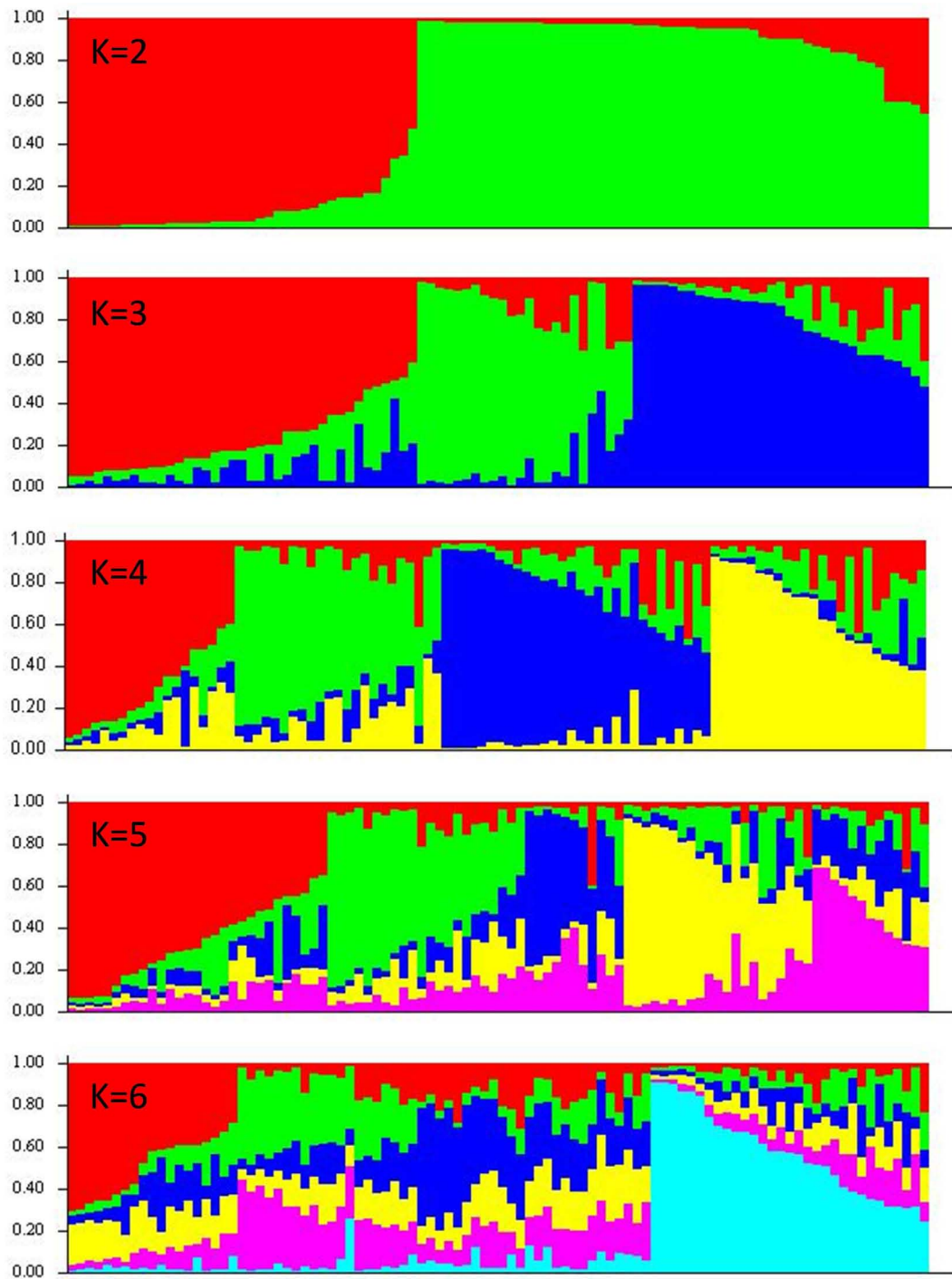
Gene	Gene ID	Genomic gene length (bp)	Chromosome	No. SNPs	SNP density
<i>VTE1</i>	AT4G32770	3069	4	10	1/307 bp
<i>VTE2</i>	AT2G18950	3122	2	5	1/624 bp
<i>VTE3</i>	AT3G63410	1530	3	9	1/170 bp
<i>VTE4</i>	AT1G64970	2025	1	68	1/30 bp
<i>VTE5</i>	AT5G04490	1910	5	23	1/83 bp
<i>PDS1</i>	AT1G06570	1661	1	4	1/415 bp
<i>ACT2</i>	AT3G18780	2333	3	12	1/194 bp
<i>ACT8</i>	AT1G49240	2441	1	12	1/203 bp
<i>LOV1</i>	AT1G10920	3348	1	16	1/209 bp
<i>RPP1</i>	AT3G44480	5795	3	207	1/28 bp

The database compares the genomic sequences of 80 *A. thaliana* accessions. For density calculation only SNPs with a frequency of >5% were considered and then related to the genomic length. SNP densities of random chosen actin genes (*ACT1*, *ACT8*) as well as loci, involved in defense response (*RPP1*, *LOV1*), were also calculated.

**Table A5 | Number of polymorphisms within the amplified gene regions of tocopherol candidate genes from *B. napus*, their properties and their frequency in panel 1 accessions.**

Gene	Base pairs sequenced (bp)	Sequence variations				Frequency (in %) <sup>1</sup>
		Total	Intron	Silent	Non-silent	
<i>BnaX.PDS1.a</i>	920	20	0	16	4	1.06
<i>BnaX.PDS1.b</i>	598	0	0	0	0	0
<i>BnaA.PDS1.c</i>	304	10	0	7	3	52.13
	729	18	15	37	0	29.16
<i>BnaX.VTE1.a</i>	860	0	0	0	0	0
<i>BnaX.VTE1.b</i>	1028	0	0	0	0	0
<i>BnaX.VTE2.b</i>	780	1	0	0	1	4.16
<i>BnaX.VTE3.a</i>	753	6	0	6	0	17.77
<i>BnaX.VTE3.b</i>	478	1	0	0	1	1.04
<i>BnaX.VTE4.b</i>	494	0	0	0	0	0

<sup>1</sup>Percentage of accessions different from reference genotype *Tapidor*.



**FIGURE A1 |** Presentation of the population structure of 96 *B. napus* accessions of panel 1 under the assumption of subpopulation  $K = 2-6$  which calculation based on 24 SSR markers.

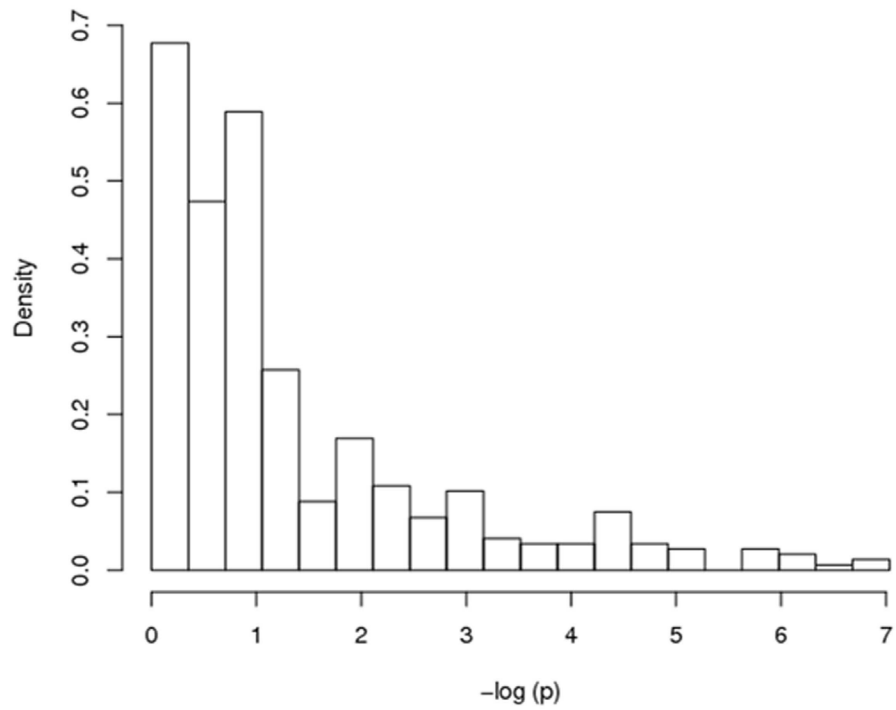
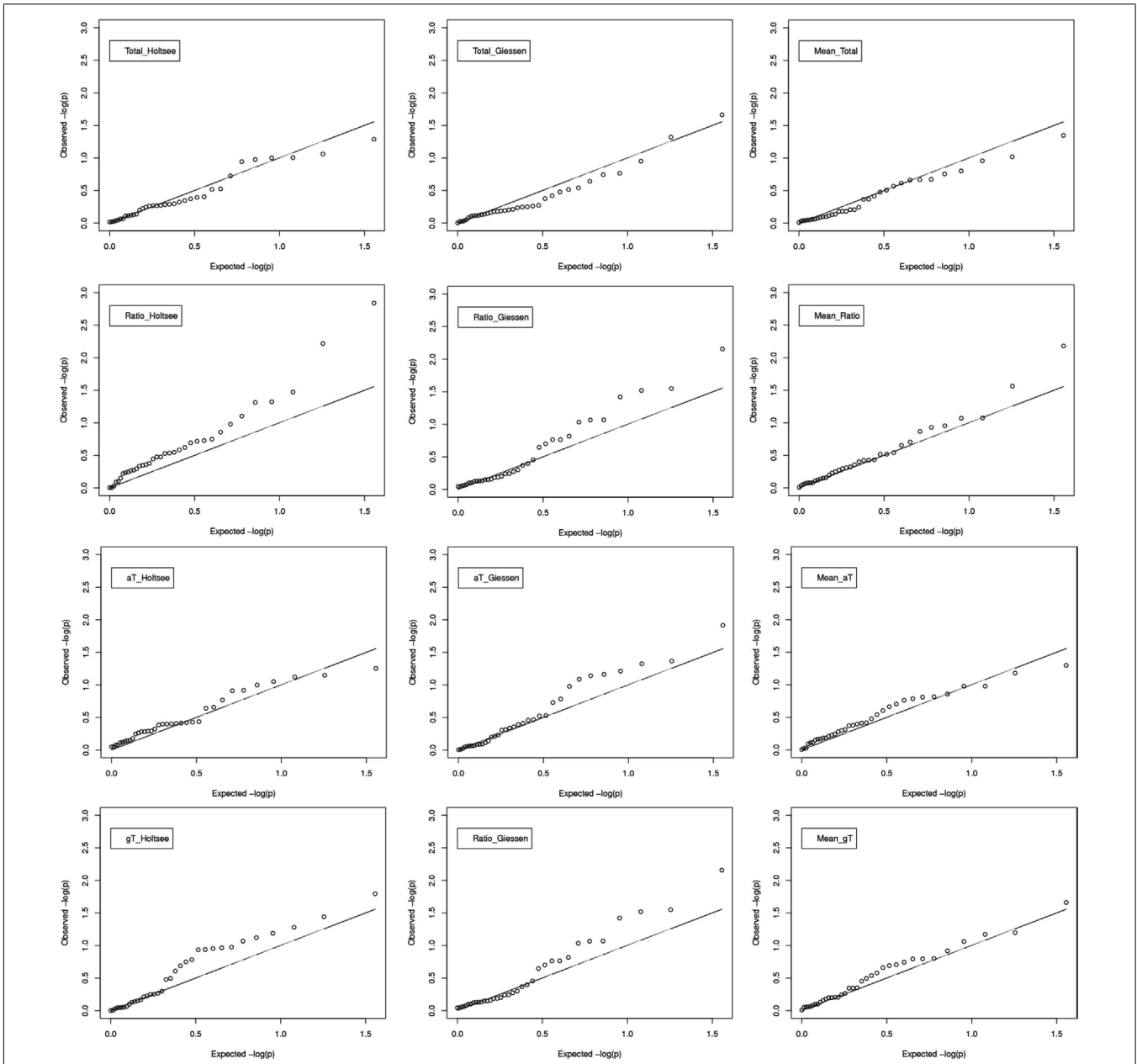


FIGURE A2 | Distribution of  $p$ -values of the tocopherol traits.



**FIGURE A3 |** QQ-plot shows the distribution of the  $p$ -values of the associations of the SSRs with the phenotypic traits.