Theor Appl Genet (2012) 124:957–969
DOI 10.1007/s00122-011-1760-z

ORIGINAL PAPER

A mutation screening platform for rapeseed (*Brassica napus* L.) and the detection of sinapine biosynthesis mutants

Hans-Joachim Harloff · Susanne Lemcke · Juliane Mittasch · Andrej Frolov · Jian Guo Wu · Felix Dreyer · Gunhild Leckband · Christian Jung

Received: 21 July 2011 / Accepted: 21 November 2011 / Published online: 24 December 2011

Abstract We developed two mutant populations of oilseed rape (*Brassica napus* L.) using EMS (ethylmethanesulfonate) as a mutagen. The populations were derived from the spring type line YN01-429 and the winter type cultivar Express 617 encompassing 5,361 and 3,488 M_2 plants, respectively. A high-throughput screening protocol was established based on a two-dimensional 8× pooling strategy. Genes of the sinapine biosynthesis pathway were chosen for determining the mutation frequencies and for creating novel genetic variation for rapeseed breeding. The extraction meal of oilseed rape is a rich protein source containing about 40% protein. Its use as an animal feed or human food, however, is limited by anti-nutritive compounds like sinapine. The targeting induced local lesions in genomes (TILLING) strategy was applied to identify mutations of major genes of the sinapine biosynthesis pathway. We constructed locus-specific primers for several TILLING amplicons of two sinapine synthesis genes, *BnaX.SGT* and *BnaX.REF1*, covering 80–90% of the coding sequences. Screening of both populations revealed 229 and 341 mutations within the *BnaX.SGT* sequences (135 missense and 13 nonsense mutations) and the *BnaX.REF1* sequences (162 missense, 3 nonsense, 8 splice site mutations), respectively. These mutants provide a new resource for breeding low-sinapine oilseed rape. The frequencies of missense and nonsense mutations corresponded to the frequencies of the target codons. Mutation frequencies ranged from 1/12 to 1/22 kb for the Express 617 population and from 1/27 to 1/60 kb for the YN01-429 population. Our TILLING resource is publicly available. Due to the high mutation frequencies in combination with an 8× pooling strategy, mutants can be routinely identified in a cost-efficient manner. However, primers have to be carefully designed to amplify single sequences from the polyploid rapeseed genome.

Communicated by R. Visser.

Hans-Joachim Harloff and Susanne Lemcke have equally contributed.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1760-z) contains supplementary material, which is available to authorized users.

H.-J. Harloff · S. Lemcke · C. Jung (✉)
Plant Breeding Institute, Christian-Albrechts-University of Kiel, Olshausenstr. 40, 24098 Kiel, Germany
e-mail: c.jung@plantbreeding.uni-kiel.de

H.-J. Harloff
e-mail: h.harloff@plantbreeding.uni-kiel.de

S. Lemcke
e-mail: s.lemcke@plantbreeding.uni-kiel.de

J. Mittasch · A. Frolov
Department of Secondary Metabolism, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle (Saale), Germany
e-mail: juliane.mittasch@googlemail.com

A. Frolov
e-mail: andrej.frolov@bbz.uni-leipzig.de

Present Address:
J. G. Wu
College of Agriculture and Biotechnology, University of Zhejiang, Hangzhou, China
e-mail: jgwu@zju.edu.cn

F. Dreyer · G. Leckband
Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, Hohenlieth, Germany
e-mail: f.dreyer@npz.de

G. Leckband
e-mail: g.leckband@npz.de

Abbreviations

BAC	Bacterial artificial chromosome
CODDLE	Codons optimized to discover deleterious lesions
EMS	Ethylmethane sulfonate
IRD label	Infrared detection label
PCR	Polymerase chain reaction
REF1	Reduction of epidermal fluorescence1
SGT	UDP-glucose:sinapic acid glucosyltransferase
TILLING	Targeting induced local lesions in genomes

Introduction

Oilseed rape (*Brassica napus* L.) is the most important oil crop in temperate regions. It is grown for the production of biodiesel, animal feed, and vegetable oil for human consumption. The seed oil content ranges between 40 and 50% of dry matter, and the meal remaining after oil extraction contains about 40% protein (Hüsken et al. 2005). The seeds also accumulate secondary plant metabolites of the phenylpropanoid pathway, which contribute to the bitter taste, astringency, and dark color of rapeseed products (Shahidi and Naczk 1992) and limit the use of the protein-rich seed meal. Sinapoylcholine (sinapine) is the major phenolic compound of *B. napus* seeds. Therefore, its reduction below 2 mg/g seed is an important aim of oilseed breeding. In a study with 575 genotypically divergent winter rapeseed accessions (zum Felde et al. 2007), the sinapine content varied from 3.4 to 12.9 mg/g seeds, which indicates limited genetic variation within the rapeseed gene pool (Bell 1993; Hüsken et al. 2005; Shahidi and Naczk 1992; Wang et al. 2008; zum Felde et al. 2006, 2007). Hence, genetic modification or mutation induction targeting the sinapine metabolic pathway genes is the method of choice to breed rapeseed with low sinapine content.

The biosynthesis of sinapine in *Brassicaceae* is well known and starts via the phenylalanine/hydroxycinnamate pathway (Hüsken et al. 2005; Milkowski et al. 2004; Shahidi and Naczk 1992). The recent knowledge is summarized in the MetaCyc Pathway Database (http://biocyc.org/META/NEW-IMAGE?type=PATHWAY&object=PWY-5168). Important intermediate metabolites are 4-coumaroyl-CoA, caffeoyl-CoA, coniferaldehyde, and sinapaldehyde. Coniferaldehyde is converted in four steps via 5-hydroxy-coniferaldehyde, sinapic acid, and sinapoylglucose to sinapoylcholine (sinapine). Complete genomic and cDNA sequences have been published for *BnaX.CALD5H* (*BnF5H*, *BnFAH1*, coniferaldehyde 5-hydroxylase, EC 1.14.13.-) (Nair et al. 2000), *BnaX.SGT* (UGT84A9, UDP-glucose:sinapic acid glucosyltransferase, EC 2.4.1.120) (Mittasch et al. 2010), *BnaX.SCT* (sinapoylglucose:choline sinapoyltransferase, EC 2.3.1.91) (Weier et al. 2008), and *BnaX.REF1* (sinapaldehyde dehydrogenase/coniferaldehyde dehydrogenase, EC 1.2.1.68) (unpublished). The copy numbers varied between 2 and 4. Expression studies showed that only two out of four *BnaX.SGT* genes are expressed in developing seeds (Mittasch et al. 2010) as well as two *BnaX.REF1* genes that make them suitable targets for manipulating seed sinapine content.

To reduce sinapine production in *B. napus* seeds, gene expression was altered by genetic modification. Hüsken et al. (2005) described suppression of the *BnaX.SGT* gene by transformation with a dsRNAi construct that reduced the amount of sinapic acid esters in T_2 seeds by 61%, whereas sinapine content in homozygous T_2 seeds was reduced by 40% after transformation with an antisense construct of *BnaX.CALD5H* (Nair et al. 2000). A more recent study used a concomitant antisense suppression of *BnaX.CALD5H* and *BnaX.SCT* showing reductions of sinapine up to 90% (Bhinu et al. 2009). Unfortunately, genetically engineered rapeseed is not accepted in the EU and other countries due to legal restrictions. Therefore, the identification of mutants by targeting induced local lesions in genomes (TILLING) offers an alternative, particularly as sequence and copy number information of a number of genes of the sinapine biosynthesis pathway are available.

TILLING is a screening method for plant populations that have been treated with chemical mutagens like EMS (ethylmethanesulfonate) or MNU (*N*-nitroso-*N*-methylurea) causing point mutations. In the case of EMS, G residues are ethylated and result in G/C → A/T transitions. Irreversible point mutations accumulate in the population in a wide range of modified functions, most probably resulting in loss of gene function (Greene et al. 2003; McCallum et al. 2000). A TILLING resource consists of a population of several 1,000 M_2 plants and M_3 seeds derived from selfed M_2 plants. M_2 plant DNA is analyzed by sequence-specific PCR and CEL I endonuclease cleavage, preferably in pools of up to 8 DNA samples (Till et al. 2006). TILLING resources have been published for a wide variety of species including *Arabidopsis*, maize, rice, oat, and wheat (Chawade et al. 2010; Greene et al. 2003; Slade et al. 2005; Weil and Monde 2007).

Brassica napus (AACC, $2n = 4x = 38$; 1,130 Mbp/C) is an amphidiploid species originating from several independent spontaneous hybridization events of *B. rapa* (AA, $2n = 2x = 20$) and *B. oleracea* (CC, $2n = 2x = 18$) (Allender and King 2010; U 1935), which both underwent segmental genome triplications during their evolution. On the average, three paralogous genes corresponding to single orthologues in *Arabidopsis* were found in the *Brassica* A or C genome (Parkin et al. 2005), which is in line with two to four homologous sequences found after BAC library

screening and Southern hybridization (Mittasch et al. 2010; Stephenson et al. 2010; Wang et al. 2008).

First results on TILLING in rapeseed were presented by Wang et al. (2008). Screening of 1,344 M_2 plants resulted in 19 mutants of the *BnFAE1* gene family, which were phenotypically verified by M_3 analysis. However, lack of a pooling strategy to facilitate the screening procedure limits the use of this TILLING resource. Moreover, non-specific primer combinations were used that did not differentiate between the respective *BnFAE1* gene loci.

The aim of this study was to establish a TILLING resource of spring and winter type *B. napus* for the rapid and cost-effective identification of mutants. Two major genes of the sinapine biosynthesis pathway have been chosen for TILLING. Locus-specific PCR and a two-dimensional pooling technique have been applied. A protocol for routine screening of 5,361 and 3,488 M_2 plants for the spring and winter rapeseed population is presented resulting in the identification of 570 *BnaX.SGT* and *BnaX.REF1* mutants. Homozygous M_3 plants were identified for future crossing experiments to combine different mutations in a single plant.

Materials and methods

Plant material and EMS mutagenesis

A Canadian yellow-seeded spring type inbred line YN01-429 (F_8) kindly provided by Prof. G. Rakow (AAFC Saskatoon, Canada) and a winter type inbred line Express 617 (F_{11}) derived from the German cultivar Express were used in this study. Seeds were soaked in tap water for 12 h prior to 12 h of EMS treatment. Winter type rapeseed plants with six leaves were vernalized at 4°C for 12 weeks.

DNA extraction and pooling strategy

Leaf samples from M_2 plants were harvested for DNA extraction. Leaves from spring type plants were sampled in 2-ml Eppendorf tubes, whereas the winter type leaves were sampled in 96-well plates. Genomic DNA was isolated from freeze-dried material (sample dry weights 20–50 mg spring and 10–20 mg winter type material) in a 96-microtiter plate format using a NucleoSpin® 96 Plant I Kit (Macherey and Nagel, Düren, Germany) and the TECAN Freedom Evo 200 Liquid Handling Robot (4 × 96 samples/day; TECAN GmbH, Crailsheim, Germany). DNA concentrations were measured in a Genios Microplate Reader (TECAN GmbH, Crailsheim, Germany) using Quant-It-Picogreen dsDNA Reagent (Invitrogen, Karlsruhe, Germany). Average DNA yields were 10.3 (YN01-429) and 5.8 (Express 617) µg DNA/sample, which is sufficient for screening 50 and 24 amplicons, respectively. As the same kit was used for DNA isolation, differences in DNA yield were due to different amounts of leaf material.

For normalization, DNA aliquots were diluted to a PCR-ready final concentration of 5 ng/µl and arranged in one-dimensional 4× pools for YN01-429 and two-dimensional (2D) 8× pools for Express 617. In the 2D pools, 4 microtiter plates with normalized DNA samples were combined in one pool plate with columns 1–6 containing the 8× column pool DNA and columns 7–12 containing the 8× row pool DNA. Due to this arrangement, every single sample is represented twice on a 96-lane LI-COR gel and can be directly identified by assignment of the lanes (Supplement S1). As leaf sampling and DNA extraction were performed according to family number with subsequent column-wise storage in the microtiter plate and 4×-pools were built up by combining equivalent positions of 4 plates, 4× pools never contained sibling M_2 samples, whereas in the 8× pools, the column pools contained samples from 8 families (8 × 1), whereas the row pools contained samples from only 2 families (4 × 2). In the case of 1D-4× pools, we sequenced the respective amplicons of all four plants for mutant identification, whereas in the 2D-8× pool, the mutant plant could be directly identified, and only one amplicon had to be sequenced, which greatly facilitated the mutant detection procedure.

Sequence information of the *BnaX.SGT* and *BnaX.REF1* genes

The coding sequence of *BnaX.SGT* (UGT84A9) is 1,494 bp in size organized in one exon, whereas both *BnaX.REF1* genes have 9 exons and 8 introns. Their known genomic sequences are 3,977 and 3,973 bp in size, respectively, with a coding sequence of 1,503 bp. Sequence data and copy numbers are based on BAC library screening, expression studies, and Southern hybridization. The GenBank accession numbers for *BnaA.SGT.a* (UGT84A9b), *BnaC.SGT.a* (UGT84A9a), *BnaX.REF.a,* and *BnaX.REF.b* are FM872285, FM872284, FN995990, and FN995991, respectively. In addition to BLAST analysis, the software CLC Main Workbench (CLC bio, Aarhus, Denmark) was also used for *in silico* sequence evaluation.

Primer design and PCR conditions

Locus-specific primers for the PCR amplification of the coding regions of these genes were designed with the program FastPCR© (Kalendar et al. 2009) and tested with unlabeled and 5′ labeled primers (IRD labels Dy-681 in the forward and Dy-781 in the reverse primers, Biomers, Ulm,

Germany) according to the protocol of Till et al. (2006). For PCR, we were using a DYAD thermal cycler (MJ Research Inc., Waltham, MA, USA).

Heteroduplex analysis, fragment detection, and calculation of mutation frequencies

The CEL I enzyme was extracted from celery as described by Till et al. (2006). Heteroduplex formation of the PCR product and digestion with CEL I were performed according to the same reference. Prior to loading to the gel, 2 µl of the digestion product was mixed with 2 µl of formamide loading dye and denatured for 3 min at 95°C. Aliquots of 0.3–0.5 µl were applied to a 6.5% polyacrylamide gel (KB^{Plus} Gel Matrix, LI-COR®, Bad Homburg, Germany) and separated on a LI-COR 4300 DNA Analyzer with double laser detection system for IR-labeled primers. The gel was run for 4:15 h at 1,500 V, 40 mA, and 40 W. The fragments were analyzed with the GelBuddy software (Zerr and Henikoff 2005). After the sample assignment of the fragments, mutations were identified by Sanger sequencing of the corresponding PCR products. Sequence analysis was performed using dye-terminator chemistry (Applied Biosystems, Foster City, CA, USA) on a 3730xL DNA Analyzer (Applied Biosystems). Mutation frequencies F [1/kb] were calculated using amplicon sizes corrected by 100 bp for LI-COR gel border effects according to the formula:

$$F[1/\mathrm{kb}] = 1 \Big/ \left(\frac{(\text{amplicon size}\,[\mathrm{bp}] - 100) \times (\text{number of } M_1 \text{ plants})}{(\text{number of mutations}) \times 1{,}000} \right)$$

Determination of sinapic acid metabolites

The M_3 plants were grown in the greenhouse under 16 h of light. Four different dates of sowing were chosen (=four experiments). M_4 seeds were harvested after bag isolation. Sinapine and sinapoylglucose were determined by HPLC as described in Milkowski et al. (2004). If not otherwise indicated, single M_4 seeds were analyzed. Sinapic acid ester equivalents (SAE) were determined in single seed extracts after alkaline hydrolysis for 3 h at 50°C in 5 N KOH (Wolfram et al. 2010).

Results

EMS mutagenesis and development of TILLING populations

Two different rapeseed lines were employed in this experiment. First we produced an EMS mutant population with the spring type rapeseed line YN01-429. We used different concentrations of 0.5, 0.8, 1.0, and 1.2% EMS, which were in the range of previously published data (Röbbelen 1967; Thurling and Depittayanan 1992) (Table 1). Survival rates in the M_1 generation dropped from 80% (0.5% EMS) to 50% (1.2% EMS). Therefore, higher EMS concentrations were avoided and concentrations between 0.5 and 1.2% were chosen for further studies.

According to different EMS treatments, the YN01-429 population was subdivided into two subpopulations. The first subpopulation consisted of 500 vigorous YN01-429-derived M_1 plants derived from 2,400 EMS-treated seeds (0.5 and 1.0% EMS). They were self-pollinated by bag isolation, and M_2 seeds were harvested. Of each M_2 family, 4 plants were grown to avoid loss of mutations due to the chimeric character of the M_1 plants. After bag isolation, M_3 seeds were harvested from 1,724 M_2 plants grown in the greenhouse. The second subpopulation derived from 0.8 to 1.2% EMS treatments consisted of 2,833 vigorous M_1 plants. Bag isolation resulted in seeds from 2,833 M_2 families. Three plants of each M_2 family were grown in the field, and M_3 seeds of 3,629 plants were harvested without bag isolation. Leaf material was taken, and DNA was isolated from all 5,361 plants of the YN01-429 spring rapeseed M_2 population. DNA samples of all treatments were later jointly investigated by TILLING.

For the production of a winter type rapeseed TILLING population, the Express 617 inbred line was treated with 1% EMS, resulting in an M_1 lethality rate of ~40%. A total of 2,103 viable M_1 plants were obtained, and seeds were harvested from 1,902 M_1 plants. A total of 7,608 M_2 plants (4 plants/family) were grown in the greenhouse, and M_3 seeds were harvested from 6,775 M_2 plants. DNA was isolated from 3,488 M_2 plants representing 940 M_1 plants. Leaf samples of another 3,732 M_2 plants were freeze-dried and stored for further use (Table 1).

Primer design and CEL I digest

TILLING in *B. napus* is hampered by the fact that many sequences exist as paralogues and orthologues with high sequence similarity among each other. Therefore, we designed locus-specific primer combinations that cover gene regions coding for functional domains of the polypeptide. The amplicon size should be <1,500 bp because of technical limitations due to resolution and background of the LI-COR gels (Till et al. 2006). Primers were carefully tested prior to TILLING: first, a so-called PCR crash test was carried out with primer pairs and single primers from paralogous and orthologous loci. Suitable primer combinations should not give any non-specific amplicons visible as additional bands or smear after gel electrophoresis. Second, the obtained PCR products were sequenced to confirm locus specificity, and third, the PCR was repeated

Table 1 Features of the rapeseed TILLING platform

Population	YN01-429		Express 617
Number of seeds	2,400	5,000	3,120
EMS concentration	0.5, 1%	0.8, 1.2%	1%
Number of M_1 plants	2,000	3,980	2,103
Number of M_1 plants with seed set	500	2,833	1,902
Number of seeds per M_2 family	4	3[a]	4
Number of M_2 plants	2,000	3,860	7,608
Number of selfed M_2 plants	2,000	0	7,608
Open pollinated M_2 plants	0	3,860	0
Leaf samples, DNA extraction	1,905	3,456	3,488[c]
M_1 plants represented	500[b]	1,500[b]	940
DNA samples in M_2 population	5,361		3,488

[a] 8,499 seeds were sown in the field, only 3,860 M_2 plants survived

[b] Estimated values of represented M_1 plants due to combination of different subsets

[c] Only half of the leaf material was extracted, leaf material of 3,732 additional M_2 plants is available (originating from another 950 M_1 plants)

with IRD-labeled primers. Figure 1 shows the genomic structure of the target genes and the location of the amplicons. All primers were 20–30 nucleotides long with melting temperatures of 60–65°C to avoid interaction between the two IRD labels during the PCR. It was known that four *BnaX.SGT* loci are present in the *B. napus* genome, with only two of them (*BnaA.SGT.a* and *BnaC.SGT.a*) being expressed in ripening seeds (Mittasch et al. 2010). Therefore, we designed three locus-specific primer combinations for *BnaA.SGT.a* and *BnaC.SGT.a* that gave rise to amplicons in a range between 1,270 and 1,420 bp covering between 85 and 95% of the coding sequence (Supplement S2).

Two *REF1* homologues (*BnaX.REF1.a* and *BnaX.REF1.b*) had been discovered in the rapeseed genome. For each locus, primers were designed for two amplicons with sizes between 943 and 1,361 bp. Taken together, they cover 84% of the coding sequence including 7 out of 9 exons (Supplement S2).

Apart from primer design, an optimized CEL I digest in combination with a refined pooling strategy is critical for successful mutant detection. We performed a number of heteroduplex digestion experiments with varying amounts of CEL I enzyme to determine the optimal signal-to-noise ratio after LI-COR gel electrophoresis (data not shown). An existing SNP within *BnaC.SGT.b* between YN01-429 and Express 617 served as a positive control, as two fragments became visible after CEL I digestion of the mixed amplicons.

Detection and characterization of EMS mutations

We used a pooling strategy to screen the rapeseed EMS population, 4× for the spring type and 2D-8× for the winter type population (see "Materials and methods"). Our protocol enabled the detection of TILLING fragments, even in the case of a low signal-to-noise ratio. Fifty to 80% of the polymorphic fragments identified after gel electrophoresis indicated real point mutations as verified by Sanger sequencing. The position of the SNP matched the position of the CEL I cleavage with a precision of 5–20 bp. In the 8× pools, no effect of increased effective concentrations of mutant alleles in the 2D row pools containing sibling samples was observed (see "Materials and methods").

TILLING of four sinapine genes resulted in a total of 683 mutations, which were later verified by Sanger sequencing (Table 2). As in some cases the same mutation was found in more than one plant of the same M_2 family, we corrected the number of heritable germ-line mutations by counting those mutations only once in each family. This resulted in a total of 570 different transitions by subtracting 113 mutations within one and the same family. A complete list of all mutants is provided in Supplement S3. The mutation frequencies (formula described in "Materials and methods") varied between the different sequences investigated. We found the lowest mutation frequency within the spring rapeseed population (*BnaC.SGT.a,* 1/60 kb) and the highest frequency within the winter rapeseed population (*BnaX.REF1.b*, 1/12 kb). Average mutation frequencies within the *BnaX.SGT* and *BnaX.REF1* genes were calculated as 1/31 and 1/12 kb, respectively. Differences were also found between both populations with lower frequencies in the spring population.

In the following, we grouped the mutation events into distinct classes to address the question whether EMS mutations are randomly distributed across the different genes. It has been reported that only a limited number of

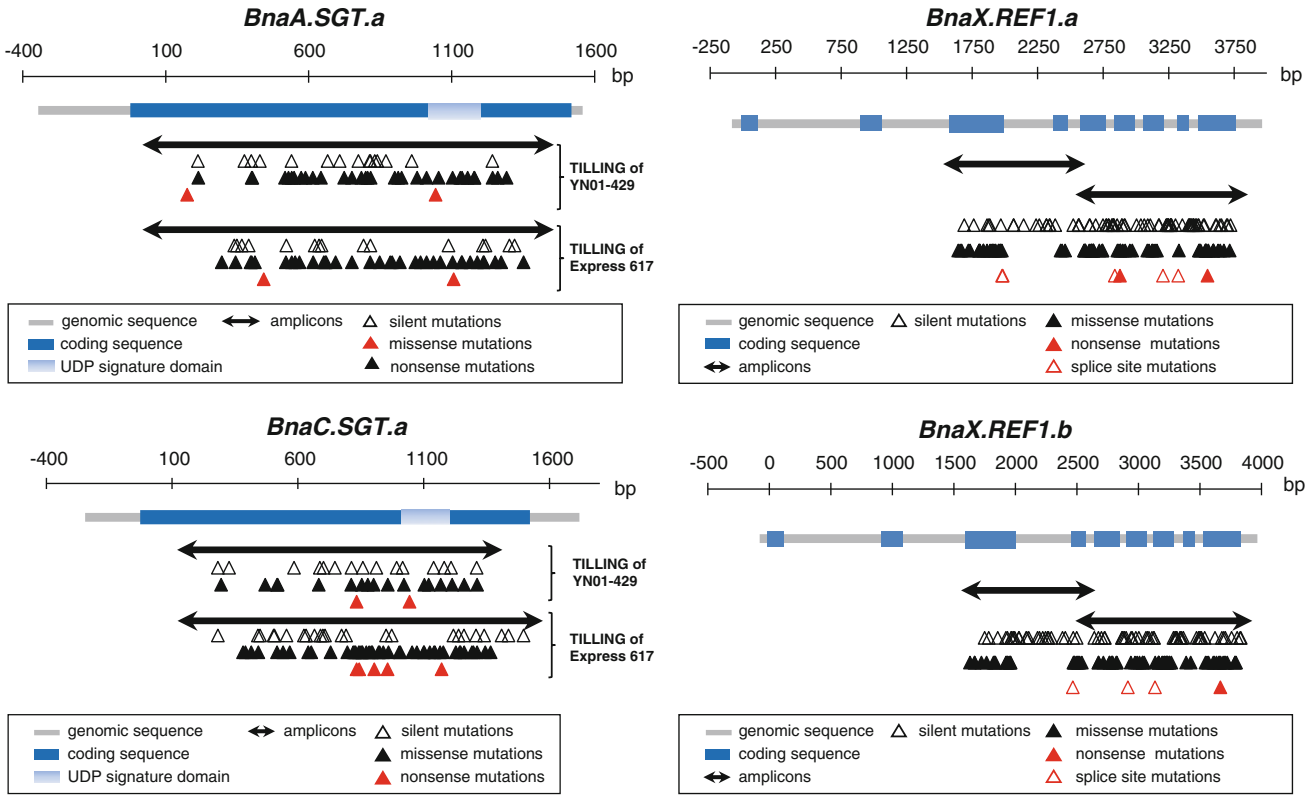

Fig. 1 Gene structure of *BnaX.SGT* and *BnaX.REF1* genes and mutations detected by TILLING

Table 2 TILLING with sinapine biosynthesis gene sequences

Gene	ORF (bp)	% covered by TILLING	LI-COR fragments detected	Mutations verified by sequencing	Mutations verified by M_3 analysis	M_3 phenotypes	Mutations in M_1	Mutation frequency [1/kb][a]	Mutations/ 1,000 M_1/ 1,000 bp
YN01-429									
BnaA.SGT.a	1,494	94	95	61	14	2	55	1/27[b]	20
BnaC.SGT.a	1,494	85	82	39	13	3	39	1/60[c]	15
Express 617									
BnaA.SGT.a	1,494	94	107	60	20	1	56	1/22	42
BnaC.SGT.a	1,494	95	127	90	6	1	79	1/16	59
BnaX.REF1.a	1,503	84	250	205	4	0	164	1/12	77
BnaX.REF1.b	1,503	84	291	228	3	0	177	1/12	83

For *BnaX.REF1,* values of both amplicons were added. Mutation frequencies were calculated as number of mutations/M_1 plant that was determined by analyzing the M_2 families

[a] Amplicon correction for LI-COR gel border effects by 100 bp

[b] Screening of 1,140 M_1

[c] Screening of 2,000 M_1

base triplets can be changed by EMS treatment, with guanine being the predominant target of ethylation (Stephenson et al. 2010). Thus, the frequency of mutagenic events should be correlated either with the frequency of G on both strands or with the frequency of G/C (G + C) on the coding strand. Accordingly, 99.3% of our mutations were G/C → A/T transitions, while only 0.7% (4 among all 570 mutations) were non-G/C → A/T transitions (Supplement S3). We classified the mutations in relation to the number of G/C residues (i.e., the maximum number of EMS targets) excluding the 4 non-G/C → A/T transitions. The frequency of mutated G/C residues ranged between 6% (*BnaX.SGT*) and 20% (*BnaX.REF1*). The distribution between the classes becomes more distinct if the

probability of a mutation is compared with its observed frequency (Table 3a, b). Probabilities and frequencies are expressed as percentages of total G/C and total mutations. In most cases, the observed values differed from the expected ones by <3%. They showed the same distribution pattern over different classes, suggesting that up to the M_2 generation, selection processes favouring for instance silent mutations or mutations within the introns do not play a major role. Even within a comparatively small and conservative domain like the UDPGT signature domain (proposed UDP-glucose binding site, Hughes and Hughes 1994) of *BnaX.SGT,* the mutation rates were as expected. Moreover, no apparent strand selectivity could be found, with the ratio of G → A to C → T transitions in most cases almost equaling the G/C ratios in the codon strand.

To visualize the distribution of single mutation events over the whole gene, the number of events in the amplicons was added up in intervals of 100 bp and plotted against the sum of the G/C residues (Fig. 2). Graphs clearly demonstrate that the number of mutation events in a given region of the gene follows the G/C content. In the case of *BnaX.REF1,* the regions of higher G/C content are exons. A slight deviation from this rule occurs at the beginning and the end of the respective amplicon also confirming previous findings (Stephenson et al. 2010) and justifying the correction of the amplicon for the calculation of mutation frequencies. A correlation between mutation events and G/C content is further demonstrated by a finer segmentation of 10 bp and by plotting the number of mutations in each segment against the product of the number of G/C residues in the same segments and its frequency, suggesting a proportionality to both G/C concentration and frequency. The calculations for Express 617 in Fig. 3 show a high probability for such a correlation. For the *BnaX.SGT* genes in YN01-429, the R^2 values ranged between 0.9156 and 0.9396 (data not shown).These findings correspond to the distribution of mutations in classes as described above and support the idea of a random distribution of EMS mutations over the whole genome depending on the G/C content.

We further looked for multiple mutation events (double, triple mutations) among all 8,849 M_2 plants investigated. We found 18 nucleotide positions within the *BnaX.SGT* gene (8%) and 40 nucleotide positions within the *BnaX.REF1* gene (12%) with double and triple mutations. In total, 52 double and 6 triple mutations were found in both genes (8% within *BnaX.REF1* exons, see Supplement S3; visualized on the protein level in Supplement S4). We further analyzed the randomness of multiple mutations at the same nucleotide position giving a probability between 1/144 and 1/3,600 kb [$(1/12)^2$ and $(1/60)^2$] for double and between 1/1,728 and 1/216,000 kb [$(1/12)^3$ and $(1/60)^3$] for triple mutations, resulting in an expected number of 38 double and 3 triple mutations. Since these values are close to the observed ones, we speculate that most of the multiple mutation events can be explained by random mutation. However, "pseudo mutations" due to cross-pollination or DNA probe admixture cannot be completely ruled out. The latter one could account for 6 double and 3 triple "mutations" where the samples came from one and the same pool plate. Moreover, double/triple mutations were found more often in G/C-rich regions with higher transition probabilities (see Supplement S5). Therefore, the occurrence of double/triple mutations at the same nucleotide position should be a function of local concentrations of G/C residues as well as their frequency within the amplicon (see Supplement S6). The corresponding correlation was high for *BnaX.REF1* ($R^2 = 0.6926$), whereas the correlation for *BnaX.SGT* was low ($R^2 = 0.3245$), probably due to the small number of double mutations.

We further calculated the frequency of multiple mutations within one gene (>1 mutation/kb/M_2 plant) with regard to the total number of mutations (see Supplement S3). It ranged between 3–4% for *BnaC.SGT.a* (no double/triple mutations for *BnaA.SGT.a*) and 2–5% for *BnaX.REF*. We calculated the average number of mutations per single plant by multiplying mutation frequencies by genome size (2,258 Mbp/2C; Arumuganathan and Earle 1991) and corrected for an estimated average G/C content in *B. napus* of 35.7% [as a mean of 36.0% in *B. oleracea* (Town et al. 2006) and 35.4% in *B. rapa* (Trick et al. 2009)]. As a result, the number of mutations/plant in the YN01-429 and in the Express 617 EMS population was 40,000 and 130,000, respectively.

Missense, nonsense, and codon usage mutations within sinapine genes

We found a number of putative loss-of-function mutations that cause amino acid changes (missense), stop codons within coding regions (nonsense), or splice site mutations at intron borders. We detected 16 stop codon mutants (2.8%) and 8 splice site mutations (2.3%) with G → A exchanges at the 5′ and 3′ ends of the introns. Those mutations should result in non-functional enzymes that are expected to have an impact on sinapine content. We also looked for codon usage mutations, whose effects are more difficult to predict. For evaluation of putative low codon usage mutants, we used a Japanese codon usage database (http://www.kazusa.or.jp/codon/), which is based on 625 CDS and 231,574 codons. According to this database, only two EMS mutations would have a major impact on codon usage reduction, GTG → GTA (reduction from 19.9 down to 7.4%) and TTG → TTA (reduction from 18.2 down to 8.7%). Among the mutations from this study, we found 10 transitions giving rise to the above-mentioned triplet changes (Table 3).

Table 3 Predicted properties of detected mutations; (a) total numbers, (b) frequencies

Gene/screening	Parameter	Total	Silent	Low codon usage	Missense	Nonsense	UDPG_total	UDPG_missense	UDPG_non-sense
(a)									
YN01-429 screening									
BnaA.SGT.a	G/C residues	672	233	24	398	41	62	41	9
	Mutations	55	18	1	34	3	7	5	2
BnaC.SGT.a	G/C residues	614	218	23	360	36	61	42	9
	Mutations	39	17	1	19	3	5	3	1
Express 617 screening									
BnaA.SGT.a	G/C residues	672	233	24	398	41	62	41	9
	Mutations	55	18	3	35	2	7	5	1
BnaC.SGT.a	G/C residues	633	224	24	369	40	61	42	9
	Mutations	79	28	1	46	5	9	8	1

		Total	Intron	Splice site	Exon	Silent	Low codon usage	Missense	Nonsense
BnaX.REF1.a	G/C residues	873	300	16	573	175	18	378	20
	Mutations	164	58	5	106	30	2	74	2
BnaX.REF1.b	G/C residues	877	300	16	577	176	19	380	21
	Mutations	176	55	3	121	32	2	88	1

Gene/screening	%	Silent	Low codon usage	Missense	Nonsense	UDPG_total	UDPG_missense	UDPG_non-sense
(b)								
YN01-429 screening								
BnaA.SGT.a	Expected	35	4	59	6	9	6	1
	Observed	33	2	62	5	13	9	4
BnaC.SGT.a	Expected	36	4	59	6	10	7	1
	Observed	44	3	49	8	13	8	3
Express 617 screening								
BnaA.SGT.a	Expected	35	4	59	6	9	6	1
	Observed	33	5	64	4	13	9	2
BnaC.SGT.a	Expected	35	4	58	6	10	7	1
	Observed	35	1	58	6	11	10	1

		Intron	Splice site	Exon	Silent	Low codon usage	Missense	Nonsense
BnaX.REF1.a	Expected	34	2	66	20	2	43	2
	Observed	35	3	65	18	1	45	1
BnaX.REF1.b	Expected	34	2	66	20	2	43	2
	Observed	31	2	69	18	1	50	1

In (a) the numbers of mutations in M_1 grouped in classes and the distribution of G/C residues in each class are shown. In (b) probabilities for the occurrence of mutations in each class are calculated on the basis of G/C distribution and compared with the frequencies of mutations in each class. UDPG = UDPGT signature domain

A protein databases search did not reveal larger stretches of conserved sequences sensitive to mutations within the *BnaX.REF1* genes. In contrast, the ~44 amino acid UDPGT signature domain encoded by the *BnaX.SGT* gene is highly conserved within the glucosyltransferase family. Therefore, mutations within this sequence should have a large impact on enzyme stability and function. We have identified 21 mutants carrying mutations within this region, which will be further investigated.

We harvested M_3 seeds from all M_2 mutant plants. In a first step, we selected 14 *BnaA.SGT.a* and 13 *BnaC.SGT.a* mutants from the spring rapeseed population and 4 *BnaA.SGT.a* and 5 *BnaC.SGT.a* mutants as well as 4 *BnaX.REF1.a* and 3 *BnaX.REF1.b* mutants from the winter

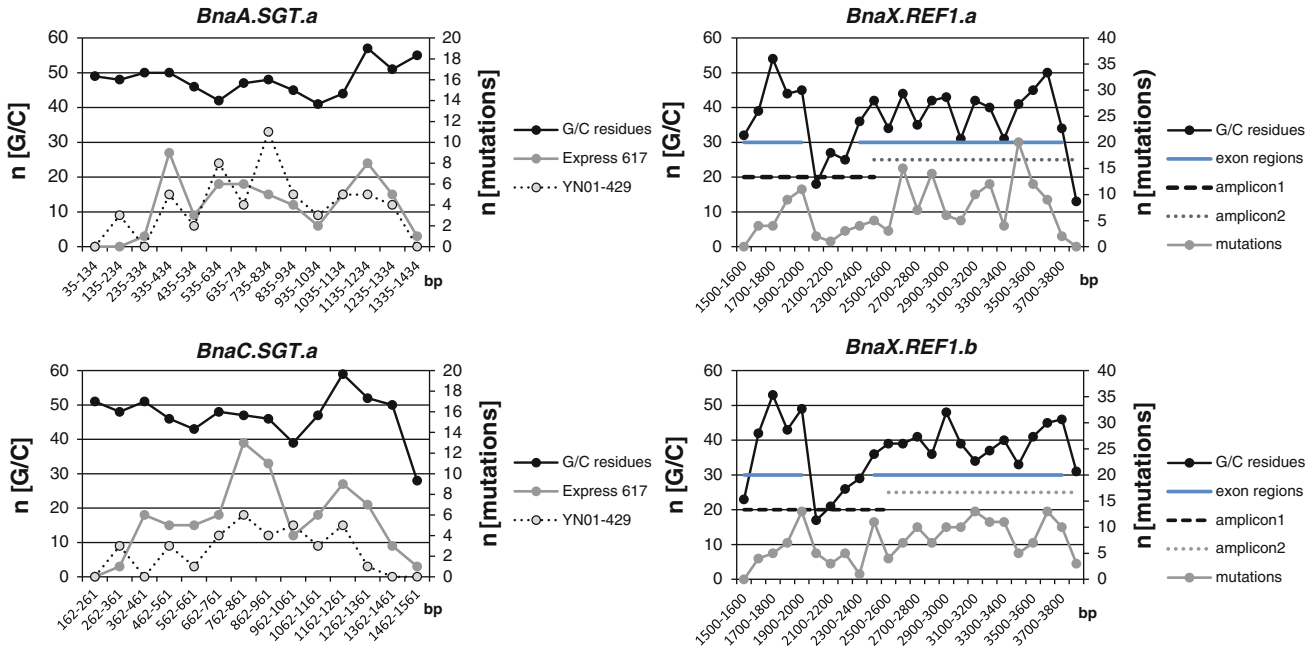

Fig. 2 Frequency distribution of EMS mutations within the *BnaX.SGT* and *BnaX.REF1* genes. The numbers (*n*) of G/C residues or mutation events were added up in 100-bp segments over the whole amplicon sequence

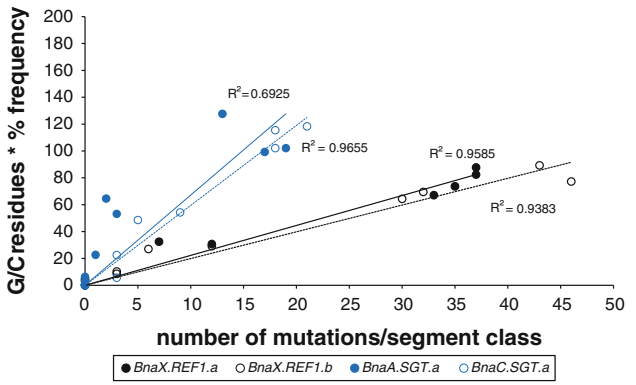

Fig. 3 Mutation events as a function of G/C distribution and frequency. The number of mutations in 10-bp segments of the amplicons is plotted against the product of the number of G/C residues in the segment and its frequency in the amplicon in %. Data are shown for the Express 617 population

rapeseed population with promising base-pair transitions as described above. We aim to select homozygous M_4 offspring for crossing and phenotyping experiments. Homozygous plants were found for all stop codon mutations clearly demonstrating that loss of one gene alone did not seem to have a deleterious or even lethal effect, because those plants showed a normal growth habit.

We did first experiments with M_4 seeds to analyze the contents of sinapine, sinapoylglucose, and sinapic acid equivalents by HPLC. For these measurements, we chose three spring type rapeseed M_3 families that were homozygous and non-segregating for the mutation (090052, 090974, and 090977) and four segregating families (spring types 090001, 090053, winter types 101612 and 101650), indicating that the parents were heterozygous. All three genotypic classes could be identified within M_3 families 090001, 101612, and 101650, while homozygous individuals were missing from family 090053. No significant reductions could be observed, which is probably due to the presence of several gene copies (Supplement S7). While in the case of the low codon usage and missense mutations, a change in protein function could hardly be predicted, and the two nonsense mutations (101612 and 101650) clearly demonstrate that the expected complete loss of function of only one of the two seed-expressed genes was not sufficient to produce a measurable effect.

As we expect drastic reductions of sinapine contents after downregulation of both seed-specific *BnaX.SGT* and *BnaX.REF1* genes, crossings of homozygous stop codon and splice site mutants in the *BnaX.SGT* and *BnaX.REF1* genes from the spring and winter rapeseed population are currently on the way in order to combine two loss-of-function mutations in one plant.

Discussion

The purpose of our study was to establish a mutant screening protocol for rapeseed by selecting mutations within two major genes of the sinapine biosynthesis pathway. TILLING procedures have been established for a

number of plants. The main features of TILLING platforms are the number of M_2 families represented by their DNA samples and the availability of $M_{2:3}$ seeds. Their efficiency relies mainly on (1) the mutation frequency, (2) the number of M_2 plants jointly tested in an experiment (pooling strategy), and (3) the costs for DNA extraction, enzyme reactions, and fragment analysis. Here, we will address these questions comparing our results with previously published TILLING protocols. The spring and winter type rapeseed TILLING platforms presented here are open for scientists to screen their sequences in our institute.

The M_1 mutation frequency is a critical parameter for TILLING. It depends on the species and the target tissue, the mutagen, the developmental stage of the mutagenic treatment, and the mutagen concentration. Typically, mutation frequencies are measured in the M_2 generation that is derived from selfed M_1 plants. Further generations can be produced by single seed descent with an increased number of families or, to avoid loss of mutations, small-sized M_2 families are grown (Suzuki et al. 2008; Rigola et al. 2009; Stephenson et al. 2010). We decided to grow and analyze four plants/M_2 family with the exception of the 2nd YN01-429 subpopulation, where due to space limitations in the field, 3 plants/M_2 family were grown. Among our M_2 families, we found ratios of heterozygous/homozygous (mutant) plants, as verified by sequencing, between 2.1 and 2.7 (with the exception of 3.88 in *BnaC.SGT.a*, YN01-429, see below), quite close to the expected 2:1 segregation ratio (disregarding the chimeric nature of the M_1s). Likewise, Greene et al. (2003) reported matching M_2 segregation ratios for *Arabidopsis*, whereas Stephenson et al. (2010) found deviating segregation ratios for *B. rapa*.

Often, data from different mutagenesis projects cannot be compared with each other because M_1 mutation frequencies have been calculated in different ways. In the following, we will present transformed data for better comparison (formula under "Materials and methods"). In our winter rapeseed population, we measured an average mutation frequency of 1/15 kb, which is higher as reported for most *Brassicaceae* like *Arabidopsis* (1/170 kb, Greene et al. 2003), *B. napus* (1/130 and 1/42 kb, Wang et al. 2008), *B. oleracea* (1/447 kb, Himelblau et al. 2009), or *B. rapa* (1/30 kb, Stephenson et al. 2010). Our mutation frequency is comparable to hexaploid (*Triticum aestivum*, 1/24 kb) or tetraploid (*T. turgidum* subsp. *durum*, 1/40 kb) wheat (Slade et al. 2005) or oat (1/20 and 1/40 kb, Chawade et al. 2010) suggesting that polyploids can tolerate a higher mutation load due to gene redundancy. This is also a reason to use the comparatively high EMS concentration of 1% for mutagenesis resulting in 130,000 mutations/plant in our Express population. The number of mutations was substantially lower in the YN01-429 population, however, with different EMS concentrations (0.5–1.2% EMS; 40,000 mutations/plant). Likewise, much lower mutation frequencies have been reported for EMS-treated populations of *B. rapa* R-o-18 ($2C = 2n$) (0.3/0.4% EMS; 20,000 mutations/plant, Stephenson et al. 2010) and *B. napus* Ningyou7 (0.6% EMS; 29,000 mutations/plant, Wang et al. 2008).

Three *Brassica* TILLING platforms have been published so far. They differ substantially from our TILLING platform with regard to size and screening efficiency. Mutant populations of the diploid species *B. rapa* and *B. oleracea* were screened by a standard 5× or 8× pooling strategy (Himelblau et al. 2009; Stephenson et al. 2010). To avoid the selection of locus-specific primer combinations, Wang et al. (2008) screened single DNA samples of a *B. napus* M_2 population subtracting natural SNPs for mutant detection. In contrast, we applied a 2D-8× pooling strategy in combination with locus-specific primers. This protocol is much more efficient for gel-based mutant detection, as it drastically reduces the scoring of false-positive fragments due to background and *Taq* polymerase error rate, and it enables mutation detection of all orthologous or paralogous sequences of a polyploid genome.

Despite these improvements, the methodical and technical demands of TILLING remain to be a challenge. In our experience, each sequence performed differently, and the differences in mutation frequencies often seem to result from methodical problems like DNA quality, biases in PCR amplification, and limited reproducibility of LI-COR gels (missing lanes, varying background) rather than from different biological background. Although random effects cannot be ruled out due to the limited number of sequences investigated, this could explain varying results after screening the *BnaA.SGT.a* gene (Express) and the *BnaC.SGT.a* gene (YN01-429), because the mutation frequencies were lower as those found for the corresponding orthologues (Table 2). The greatest differences were obvious for the silent/missense mutation frequencies (Table 3b), the heterozygote/homozygote (mutant) ratio (see "Discussion", paragraph 2), and the G → A/C → T ratio (2.93 vs. 1.13 expected).

The method of choice for accurate measurement of mutation frequencies throughout the whole genome seems to be re-sequencing of EMS mutants and wild type. Here, we were able to show by indirect means, that is, correlation of mutation events to G/C residues, that there is strong evidence for a random distribution of mutations within genes. We did not find any evidence for individual hotspots for EMS mutations in the *B. napus* genome. Since the number of sequences investigated here is considerably small, we were not able to evaluate local compositional biases defined as the proportions of observed to expected mutations in the neighborhood of mutated G residues as described by Greene et al. (2003).

Non-G/C → AT transitions after EMS mutagenesis are very rare in *Brassicaceae* (Himelblau et al. 2009; Stephenson et al. 2010), *Arabidopsis*, maize, and wheat. In our study, two A → T, one T → A, and one G → T transition were found. In contrast, much higher frequencies (up to 30%) have been reported for *Drosophila*, barley, and rice (Till et al. 2007). For a better understanding of these findings, we will elucidate the basic chemistry of the in vitro DNA ethylation experiment as described by Sega (1984). Besides the ethylation of the phosphate, major modifications were found in guanine (N-7, O^6, N-3), adenine (N-3, N-1, N-7), and cytosine (N-3). Mispairing as a prerequisite for a pyrimidine/purine base exchange was only reported for the O^6 ethylation of guanine and might occur in the N-3 ethylation of cytosine, both leading to G/C → A/T transitions. In the case of the other modifications, subsequent depurination or depyrimidation (for the trace amounts of thymine ethylation) followed by arbitrary base exchange was favoured. Two possible explanations could be given for the varying amounts of non-G/C → A/T transitions in EMS populations of different species. One is the heterogeneity of lines caused by cross-pollination or seed impurities, a possibility that was not ruled out in rice (Till et al. 2007); the other postulating repair mechanisms of varying efficiency coping with the expected proportion of depurination during EMS mutagenesis. Greene et al. (2003) offered spontaneous mutation rates in a range of 10^{-7}–10^{-8} bp/generation as a third alternative for the explanation of rare sequence variants. The latter might give an explanation to our findings as we detected only one non-G/C → A/T transition in 0.2–0.4 × 10^7 bp as calculated on the basis of one transition per gene in 3,488 M_2 plants, that is, a total screening of 0.5–0.8 × 10^7 bp, divided by two generations.

When selecting an appropriate primer combination with homology to a given target gene, two different aspects have to be taken into account. One is the search for G/C-rich regions with at least two or three sequence variations between the loci to be able to design specific primers, and the other is to cover either a large region of the gene in order not to lose any functional mutations or at least cover functional domains like catalytic centers or substrate-binding domains. The probability of EMS-mediated transitions to stop codons or missense mutations can be estimated by analyzing the coding sequence. A mere selection by the expected chemical changes due to an amino acid exchange within the possible non-synonymous mutations is, however, precarious, as even minor changes might affect the functionality of the folded protein and should not be excluded a priori, for instance when using the widespread prediction software CODDLE (http://www.proweb.org/coddle). Stop codon transitions have highest relevance as loss-of-function mutations. Likewise, 3′ and 5′ splice site mutations with G → A transitions can drastically alter the function of the resulting protein (Grisvard et al. 1998; Yuan et al. 2009), whereas other intronic splice motives like yUrAy and the polypyrimidine tract (Brown et al. 2002; Simpson et al. 2002) are no targets for the EMS reagent.

The impact of pooling size on mutation detection is a critical point. Two factors may increase the error rate when using pooled DNA for TILLING. First, large DNA pools could reduce the chance to find a mutation due to a dilution problem. Second, more complicated fragment patterns could pose a problem to analyzing LI-COR gels. Some authors refrained from using 8× pooling in order not to overlook any mutation (Gady et al. 2009). The copy number of the mutant molecule within an 8× pool can be calculated for a heterozygous mutant allele by 10*1,000/2.34/8/2 (10 ng of template DNA for PCR and 2.34 pg/2C *B. napus* genome size, Arumuganathan and Earle 1991). The resulting ∼260 mutant sequence copies versus ∼4,000 non-mutant sequence copies should be sufficient for detecting mutations during the TILLING procedure as applied in this study. Hence, we did not observe reduced information content nor did we see a problem with background bands when comparing selected 4× and 8× pools in the spring rapeseed population. Fragments detected in 4× pools were detected as well in the 8× pools (data not shown). This is in good accordance with 8× pooling experiments in rice (Till et al. 2007).

Some recent publications describe alternative techniques for detecting mutations in large populations of tomato (*Solanum lycopersicum*) like conformation-sensitive capillary electrophoresis (CSCE), high-resolution DNA melting analysis (HRM) (Gady et al. 2009), and next generation sequencing (NGS) (Rigola et al. 2009; Tsai et al. 2011). All these technical alternatives have in common that they avoid the laborious CEL I digestion and LI-COR gel electrophoresis, and offer a higher and faster sample throughput. However, with one exception (Tsai et al. 2011), the sequences to be analyzed were much smaller (<400–600 bp). That requires the development of a two to three times higher number of locus-specific TILLING amplicons to attain the same gene coverage in a candidate gene. Gady et al. (2009) found an average mutation frequency of 1/737 kb after screening an M_2 population (1% EMS) of *S. lycopersicum* by CSCE and HRM. However, they detected a high percentage of false positives, which required much proof reading and re-screening activity. In conclusion, they regarded the "classical" LI-COR method to be more sensitive. This was in line with the NGS technique using the GS FLX 454 (Rigola et al. 2009) where a lower mutation rate (1/431 kb) was found as compared to the classical approach (1/322 kb, Minoia et al. 2010). Another technical improvement has been described recently by Tsai et al. (2011) who used

TILLING amplicons <1,500 bp for DNA library construction followed by Illumina sequencing-based mutation detection. Their results were in line with "classical" screening methods. However, this study was suffering from a very small population size of only 768 plants tested, an underrepresentation of GC-rich regions after Illumina sequencing and a loss of rare heterozygous mutants due to statistical noise. The advantage of a faster screening procedure by NGS contrasts to an increased statistical and bioinformatics input to analyze the sequence reads. In addition, our TILLING method is still very cost-effective allowing the screening of 3,500 plant DNA samples (1,500 bp amplicon) for ~800€. Notwithstanding, further technical improvement is needed in the future to facilitate mutant detection in large populations.

The selection of low sinapine lines is of great interest in rapeseed breeding, because sinapine contents below 2 mg/g seed are desired. We have found 34 promising mutants in the *BnaX.SGT* genes with altered aminoacid sequences in the UDPGT signature domain and stop codon mutations. Seven $M_{2:3}$ families have been phenotypically analyzed, and no significantly lower sinapine and sinapic acid ester contents were detected. This is not unexpected as only one of two functional seed-expressed genes was mutated. Our data demonstrate that favorable rapeseed mutants can be found by TILLING, although the sinapine contents of the first mutants tested so far are not beyond the range of natural variation as described by zum Felde et al. (2007). The great potential of TILLING, however, lies in the combination of phenotypically neutral mutants to produce double or triple mutants with drastically altered phenotypes undetectable by simple phenotypic mutant screening. A crossing program has already been launched to combine loss-of-function mutations in one genotype. In the case of paralogous or orthologous mutations, inbred lines homozygous for both mutations will be selected that are expected to have much lower sinapine contents. Even the combination of two mutations within one gene is possible in rapeseed hybrids that result from a cross between two mutant lines. Moreover, mutants totally depleted in sinapine will be a valuable resource to study the physiological role of the sinapate esters in plants and to study their overall appearance, vigor, and their response to biotic stress.

Acknowledgments We thank Gislind Bräcker, Meike Pfeiler, Anja Henning, Ingrid Otschik for technical assistance and Monika Bruisch for excellent support in the greenhouse. This project was funded by the seed companies Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (Hohenlieth, Germany), Deutsche Saatveredelung AG (Lippstadt, Germany) and KWS SAAT AG (Einbeck, Germany), and the German Federal Ministry for Education, Research and Technology BMBF (GABI Future grant no. 0315052C). We especially thank Dr. G. Rakow (AAFC, Canada) for providing the spring rapeseed genotype YN01-429, Dr. Frank Breuer (KWS SAAT AG) and Dr. Zeljko Micic (Deutsche Saatveredelung AG) for Express 617 M_2 and M_3 seed propagation and M_2 leaf sampling. We thank the Zentrum für Molekulare Biowissenschaften (ZBM), University of Kiel for providing the facilities for DNA isolation and normalization and especially Prof. Dr. Axel Scheidig for CEL1 preparation. We thank Prof. Dr. Wolfgang Bilger and Jens Hermann from the Institute of Botany, University Kiel for performing additional HPLC analyses. We are also grateful to the Institute for Clinical Molecular Biology, University Kiel for considerable Sanger sequencing and especially Prof. Dr. Philip Rosenstiel for the critical reading of the manuscript.

References

Allender C, King G (2010) Origins of the amphiploid species *Brassica napus* L. investigated by chloroplast and nuclear molecular markers. BMC Plant Biol 10:54

Arumuganathan K, Earle E (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Report 9:208–218

Bell JM (1993) Factors affecting the nutritional value of canola meal: a review. Can J Anim Sci 73:679–697

Bhinu VS, Li R, Huang J, Kaminskyj S, Sharpe A, Hannoufa A (2009) Perturbation of lignin biosynthesis pathway in *Brassica napus* (canola) plants using RNAi. Can J Plant Sci 89:441–453

Brown JWS, Simpson CG, Thow G, Clark GP, Jennings SN, Medina-Escobar N, Haupt S, Chapman SC, Oparka KJ (2002) Splicing signals and factors in plant intron removal. Biochem Soc Trans 30:146–149

Chawade A, Sikora P, Brautigam M, Larsson M, Vivekanand V, Nakash MA, Chen TS, Olsson O (2010) Development and characterization of an oat TILLING-population and identification of mutations in lignin and beta-glucan biosynthesis genes. BMC Plant Biol 10:86

Gady ALF, Hermans FWK, Van de Wal MHBJ, van Loo EN, Visser RGF, Bachem CWB (2009) Implementation of two high through-put techniques in a novel application: detecting point mutations in large EMS mutated plant populations. Plant Methods 5:13

Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, Enns LC, Burtner C, Johnson JE, Odden AR, Comai L, Henikoff S (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. Genetics 164:731–740

Grisvard J, Keryer E, Takvorian A, Dever LV, Lea PJ, Vidal J (1998) A splice site mutation gives rise to a mutant of the C-4 plant *Amaranthus edulis* deficient in phosphoenolpyruvate carboxylase activity. Gene 213:31–35

Himelblau E, Gilchrist EJ, Buono K, Bizzell C, Mentzer L, Vogelzang R, Osborn T, Amasino RM, Parkin IAP, Haughn GW (2009) Forward and reverse genetics of rapid-cycling *Brassica oleracea*. Theor Appl Genet 118:953–961

Hughes J, Hughes MA (1994) Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. DNA Seq 5:41–49

Hüsken A, Baumert A, Strack D, Becker HC, Möllers C, Milkowski C (2005) Reduction of sinapate ester content in transgenic oilseed rape (*Brassica napus*) by dsRNAi-based suppression of *BnSGT1* gene expression. Mol Breed 16:127–138

Kalendar R, Lee D, Schulman AH (2009) FastPCR software for PCR primer and probe design and repeat search. Genes Genomes Genomics 3:1–14

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeted screening for induced mutations. Nat Biotechnol 18:455–457

Milkowski C, Baumert A, Schmidt D, Nehlin L, Strack D (2004) Molecular regulation of sinapate ester metabolism in *Brassica*

napus: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J 38:80–92

Minoia S, Petrozza A, D' Onofrio O, Piron F, Mosca G, Sozio G, Cellini F, Bendahmane A, Carriero F (2010) A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC Res Notes 3:69

Mittasch J, Mikolajewski S, Breuer F, Strack D, Milkowski C (2010) Genomic microstructure and differential expression of the genes encoding UDP-glucose:sinapate glucosyltransferase (UGT84A9) in oilseed rape (*Brassica napus*). Theor Appl Genet 120:1485–1500

Nair RB, Joy RWI, Kurylo E, Shi X, Schnaider J, Datla RSS, Keller WA, Selvaraj G (2000) Identification of a CYP84 family of cytochrome P450-dependent mono-oxygenase genes in *Brassica napus* and perturbation of their expression for engineering sinapine reduction in the seeds. Plant Physiol 123:1623–1634

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

Rigola D, van Oeveren J, Janssen A, Bonne A, Schneiders H, van der Poel HJA, van Orsouw NJ, Hogers RCJ, de Both MTJ, van Eijk MJT (2009) High-throughput detection of induced mutations and natural variation using keypoint (TM) technology. PLoS ONE 4:e4761

Röbbelen G (1967) Chemisch induzierte Mutationen. Fortschritte der Botanik 29:195–198

Sega GA (1984) A review of the genetic effects of ethyl methanesulfonate. Mutat Res 134:113–142

Shahidi F, Naczk M (1992) An overview of the phenolics of canola and rapeseed—chemical, sensory and nutritional significance. J Am Oil Chem Soc 69:917–924

Simpson CG, Thow G, Clark GP, Jennings SN, Watters JA, Brown JWS (2002) Mutational analysis of a plant branchpoint and polypyrimidine tract required for constitutive splicing of a mini-exon. RNA A Publ RNA Soc 8:47–56

Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN, Facciotti D (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. Nat Biotechnol 23:75–81

Stephenson P, Baker D, Girin T, Perez A, Amoah S, King GJ, Østergaard L (2010) A rich TILLING resource for studying gene function in *Brassica rapa*. BMC Plant Biol 10:62

Suzuki T, Eiguchi M, Kumamaru T, Satoh H, Matsusaka H, Moriguchi K, Nagato Y, Kurata N (2008) MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. Mol Genet Genomics 279:213–223

Thurling N, Depittayanan V (1992) EMS induction of early flowering mutants in spring rape (*Brassica napus*). Plant Breed 108:177–184

Till BJ, Zerr T, Comai L, Henikoff S (2006) A protocol for TILLING and Ecotilling in plants and animals. Nat Protoc 1:2465–2477

Till BJ, Cooper J, Tai TH, Colowit P, Greene EA, Henikoff S, Comai L (2007) Discovery of chemically induced mutations in rice by TILLING. BMC Plant Biol 7:19

Town CD, Cheung F, Maiti R, Crabtree J, Haas BJ, Wortman JR, Hine EE, Althoff R, Arbogast TS, Tallon LJ, Vigouroux M, Trick M, Bancroft I (2006) Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell 18:1348–1359

Trick M, Kwon SJ, Choi SR, Fraser F, Soumpourou E, Drou N, Wang Z, Lee SY, Yang TJ, Mun JH, Paterson AH, Town CD, Pires JC, Pyo LY, Park BS, Bancroft I (2009) Complexity of genome evolution by segmental rearrangement in *Brassica rapa* revealed by sequence-level analysis. BMC Genomics 10:539

Tsai H, Howell T, Nitcher R, Missirian V, Watson B, Ngo KJ, Lieberman M, Fass J, Uauy C, Tran RK, Khan AA, Filkov V, Tai TH, Dubcovsky J, Comai L (2011) Discovery of rare mutations in populations: TILLING by sequencing. Plant Physiol 156: 1257–1268

U N (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *Brassica napus* and peculiar mode of fertilization. Jpn J Bot 7:389–452

Wang N, Wang Y, Tian F, King GJ, Zhang C, Long Y, Shi L, Meng J (2008) A functional genomics resource for *Brassica napus*: development of an EMS mutagenized population and discovery of FAE1 point mutations by TILLING. New Phytol 180:751–765

Weier D, Mittasch J, Strack D, Milkowski C (2008) The genes BnSCT1 and BnSCT2 from *Brassica napus* encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression. Planta 227:375–385

Weil CF, Monde RA (2007) Getting the point—mutations in maize. Crop Sci 47:S60–S67

Wolfram K, Schmidt J, Wray V, Milkowski C, Schliemann W, Strack D (2010) Profiling of phenylpropanoids in transgenic low-sinapine oilseed rape (*Brassica napus*). Phytochemistry 71:1076–1084

Yuan YX, Wu J, Sun RF, Zhang XW, Xu DH, Bonnema G, Wang XW (2009) A naturally occurring splicing site mutation in the *Brassica rapa* FLC1 gene is associated with variation in flowering time. J Exp Bot 60:1299–1308

Zerr T, Henikoff S (2005) Automated band mapping in electrophoretic gel images using background information. Nucleic Acids Res 33:2806–2812

zum Felde T, Becker HC, Möllers C (2006) Genotype × environment interactions, heritability and trait correlations of sinapate ester content in winter rapeseed (*Brassica napus* L.). Crop Sci 46:2195–2199

zum Felde T, Baumert A, Strack D, Becker HC, Möllers C (2007) Genetic variation for sinapate ester content in winter rapeseed (*Brassica napus* L.) and development of NIRS calibration equations. Plant Breed 126:291–296