European Food Research and Technology (2018) 244:1473–1485
https://doi.org/10.1007/s00217-018-3062-z

ORIGINAL PAPER

Genetic and epigenetic characterization of the *cry1Ab* coding region and its 3′ flanking genomic region in MON810 maize using next-generation sequencing

Sina-Elisabeth Ben Ali[1,2] · Alexandra Schamann[1,2] · Stefanie Dobrovolny[1,3] · Alexander Indra[4] · Sarah Zanon Agapito-Tenfen[5] · Rupert Hochegger[1] · Alexander G. Haslberger[2] · Christian Brandes[1]

Received: 22 December 2017 / Revised: 6 March 2018 / Accepted: 10 March 2018 / Published online: 26 March 2018

Abstract
MON810 maize was first commercialized in 1997 and it is one of the most marketed genetically modified crops worldwide. Although MON810 maize has been studied extensively, its genetic stability and epigenetics have not been studied very well. We used next-generation sequencing to investigate the genetics and epigenetics of the *cry1Ab* coding region and its 3′ flanking genomic region in three different MON810 maize varieties. Genetic characterization of the *cry1Ab* coding region allowed us to identify and quantify several sequence variants. Samples from seeds containing a stacked MON810 event had more variants than MON810 single event varieties. Specifically, position 71 of the analyzed region varied in 15 of 600 samples tested and thus appears to be a mutational hotspot. In addition, position 71 varied at very different frequencies in the samples. Epigenetic analysis revealed a low degree of methylation, making it difficult to associate the coding region variants with methylation status. In conclusion, the variation in the coding region is either due to the increased age of the seeds from the tested maize varieties, which is known to increase the mutation rate, or due to the presence of a second (non-functional) *cry1Ab* fragment in the genome of the MON810 maize variety.

Keywords Genetic stability · GMO · MON810 · Amplicon sequencing · SNP · Bt maize · Methylation · Bisulfite sequencing

Introduction

A high percentage of economically important crops such as soybeans, maize, cotton, and canola are transgenic. Insect resistance is one of the most frequently marketed GM traits worldwide. In 2016, insect resistant maize hybrids were cultivated on an area of 53.7 million hectares, which corresponds to 88.6% of the farmland used for GM maize. The majority of GM maize varieties are stacked events, which contain mainly transgenes conferring herbicide resistance

✉ Christian Brandes
christian.brandes@ages.at

Sina-Elisabeth Ben Ali
sina-elisabeth.ben-ali@ages.at

Alexandra Schamann
alexandra.schamann@ages.at

Stefanie Dobrovolny
stefanie.dobrovolny@ages.at

Alexander Indra
alexander.indra@ages.at

Sarah Zanon Agapito-Tenfen
sarah.agapito@genok.no

Rupert Hochegger
rupert.hochegger@ages.at

Alexander G. Haslberger
alexander.haslberger@univie.ac.at

[1] Institute for Food Safety, Austrian Agency for Health and Food Safety, Spargelfeldstrasse 191, 1220 Vienna, Austria

[2] Department of Nutritional Sciences, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

[3] Department of Analytical Chemistry, University of Vienna, Währinger Straße 38, 1090 Vienna, Austria

[4] Institute for Medical Microbiology and Hygiene, Austrian Agency for Health and Food Safety, Währingerstrasse 25a, 1096 Vienna, Austria

[5] Genøk-Centre for Biosafety, PB 6418 Science Park, 9294 Tromsø, Norway

and insect resistance [1]. In this work, the term single event is used for GM plants containing a single transgene, and the term stacked event is used for plants produced via conventional cross-breeding of single events, which results in plants with more than one transgene. Next to conventional breeding, re-transformation, or co-transformation can be used for gene stacking, but these are not addressed in the current work. The recombinant *cry1Ab*-expressing insect resistant maize variety MON810 (trade name: Yieldgard®) was first commercialized in the US in 1997. It was developed by Monsanto to protect maize plants from feeding damage caused by the European corn borer (*Ostrinia nubilalis*), the southwestern corn borer (*Diatraea grandiosella*), and the pink borer (*Sesamia cretica*) [2].

MON810 maize cultivars contain a transgene cassette consisting of a 35S promoter derived from the cauliflower mosaic virus, an hsp70 intron from a maize heat shock protein, and a *cry1Ab* element derived from *Bacillus thuringiensis* (see Fig. 1) [3]. The MON810 event was produced by microprojectile bombardment of embryogenic maize tissue with a plasmid [4]. This method of transformation can lead to multi-copy events [5] with complex integration patterns if whole plasmids are used for bombardment (instead of low doses of cassette DNA) [6]. Rearranged fragments of the full-length transgene with varying copy numbers are frequently detected after bombardment [7]. Rearrangements may also be caused by regenerating transformed plants through tissue culture [5, 8, 9]. In fact, MON810 shows a complex integration pattern because of a truncation event that led to the complete loss of the NOS terminator at the 3′ end of the recombinant *cry1Ab* gene [10]. The MON810 transgene also interrupts exon 8 of the putative maize HECT E3 ligase gene 8 (on chromosome 5) and has an additional insertion [11] with 100% homology to the *Zea mays* mitochondrial genome within the first 44 bp of its 3′ flanking region. However, Monsanto [2] stated in their technical report that "Southern blot analysis of corn event MON810 demonstrated that a single functional copy of the *cry1Ab* coding sequence was integrated into the corn genome". Thus, the integration of non-functional fragments of the recombinant *cry1Ab* coding sequence or other parts of the transformation vector at additional sites in the genome has not been explicitly excluded. Nevertheless, other studies assessing the MON810 copy number by real-time PCR and digital PCR showed that only one insert is present in the genome of different MON810 maize varieties [12, 13]. Figure 1 shows the MON810 transgene cassette.

Commercialization of a GM crop requires a regulatory approval that is usually contingent on a positive safety assessment (positive means that the GM crop is considered as safe) [14]. Molecular characterization is a crucial part of safety assessments and usually relies on the results of Southern blot analysis and Sanger sequencing. It is used to determine the copy number and function of the insert, transgene stability and integrity, the DNA sequence of the insert and the flanking genomic regions, as well as the general characteristics of the introduced trait and possible interactions among stacked transgenes [15, 16]. However, while genetic stability of the transgenic locus has to be verified in single events (over five generations), it is sufficient for stacked events to establish that the integrity of authorized single events has been preserved [15].

We know from earlier investigations of MON810 maize that the coding region of the MON810 *cry1Ab* gene has a

Fig. 1 Illustration of the genetic elements of the MON810 transgene cassette and the analyzed regions. The integrated transgene cassette consists of the constitutive 35S promoter from the cauliflower mosaic virus, the hsp70 intron, and a truncated *cry1Ab* coding sequence. The 180-bp target region (Amplicon-3m810fw + rev; marked with arrows), which was selected for genetic characterization, ranges from the *cry1Ab* element to the mitochondrial (Mt) DNA in the 3′ flanking region (3′FR). The frequently detected variant at position 71 lies within this 180-bp target region. The 345-bp target region (Meth F fw + rev) was used for methylation analysis

variant (a heterozygous C → T transition at position 71) in a few samples of the stacked MON810 variety 4421VT3, but investigations of a MON810 single event did not reveal any variants [17, 18]. The influence of gene stacking on the genetic stability of transgenic inserts remains unclear. Therefore, we have analyzed three other MON810 commercial varieties to improve the safety assessment of MON810 and to investigate the impact of transgene stacking in recently detected variants. In this study, we focused on the molecular characterization of the *cry1Ab* coding region by NGS. We employed amplicon sequencing for the detection and quantification of very low-frequency variants, as well as to more accurately characterize specific genomic regions [19]. In addition, we used conventional PCR to determine the zygosity of the MON810 insert and to gain a better understanding of its genetic structure. We also investigated the methylation patterns in the *cry1Ab* coding region. 5-methylcytosine resulting from the methylation of cytosine is an important epigenetic mechanism and very common in eukaryotes [20]. Therefore, we hypothesized that this epigenetic regulatory mechanism could be linked to the observed genetic instability in that region. We used bisulfite sequencing, which is considered to be the "gold standard" method for DNA methylation studies, to assess the link between DNA methylation and the frequent occurrence of C → T transitions in the *cry1Ab* coding region [21].

Materials and methods

Biological materials

Three maize varieties were analyzed: hybrid 631RR2/Bt, hybrid DKB 350YG, and a third MON810 hybrid (name not specified). Hybrid 631RR2/Bt (Croplan Genetics®) is a stacked event (single cross hybrid) with the unique identifier MON-ØØ6Ø3–6 x MON-ØØ81Ø-6. It contains two transgenes, a herbicide resistance transgene (event NK603, trade name: Roundup Ready), and an insect resistance transgene (event MON810, tradename: Yieldgard). NK603 enables resistance against the herbicide Roundup. MON810 produces a toxin (*Cry1Ab*, derived from Bacillus thuringiensis) against lepidopteran insects. Each transgene of the stacked event is located on a different chromosome. Hybrid 631RR2/Bt was obtained from the US market (Indiana, 2005) and the hybrid seeds were used for our analyses. DKB 350YG and the unspecified MON810 hybrid are single events, each containing the single transgene MON810 with the unique identifier MON-ØØ81Ø-6. DKB 350YG (Dekalb Brasil®) was obtained from the Brazilian market in 2013. The unspecified MON810 variety was obtained from Germany in 2008. Genomic DNA was extracted from 200 individual maize grains from the unspecified MON810 variety, the 631RR2/Bt variety, and the DKB 350YG variety, respectively.

Extraction of genomic DNA

Maize grains were individually homogenized with a household garlic crusher. Then, 150 mg of the homogenized grains was incubated (o/n) with a mixture of 820 µL of TNE buffer [10 mM Tris (pH 8), 150 mM NaCl, 2 mM EDTA (pH 8), 1% SDS], 150 µL of 5 M guanidine HCl, and 30 µL of proteinase K (600 µg/mL) at 60 °C under constant shaking. Next, the mixture was centrifuged for 5 min at 16,100×*g*. Then, 300 µL of chloroform (99%) was added to 600 µL of the supernatant and the mixture was vortexed for 20 s. Then, the mixture was centrifuged at 16,100×*g* for 8 min to separate the phases and 500 µL of the aqueous phase was transferred into a new tube. Two microliters of RNAse (8 µg/mL) was added to the aqueous phase and the mixture was incubated at 60 °C for 30 min under constant shaking. The extracted DNA was purified using the Wizard DNA Clean-up System (Promega). After a pre-elution with 20 µL of ddH_2O, the DNA was eluted with 10 mM Tris buffer (pH 7.4, 70 °C) for 10 min. The eluted DNA was immediately stored at − 20 °C.

DNA quality check

Agarose gel electrophoresis, as well as spectrophotometry, were used to assess DNA quality. DNA samples and Quantitas DNA Marker (Biozym) were loaded onto a 1% w/v agarose gel containing ethidium bromide and were then run for 23 min (140 V, 400 mA, 100 W). Afterwards, the gel was viewed under UV light using a Chemi XRS Gel documentation system (Bio-Rad). Samples with clear and sharp bands in the upper region of the gel were considered acceptable. DNA from these samples was then measured with a Nanophotometer (Implen). DNA sample purity was assessed using the 260/280 absorbance ratio. Samples with a ratio between 1.75 and 1.90 were used for additional analyses.

DNA quantitation

DNA concentrations were measured using a Qubit® 2.0 Fluorometer (Life Technologies) with a Qubit® dsDNA BR Assay Kit (Life Technologies). The instrument uses fluorescent dyes that bind to specific target molecules to determine the concentration of nucleic acids. Therefore, quantitation by the Qubit® 2.0 Fluorometer method is more sensitive than the widespread UV absorbance method [22].

PCR-based zygosity testing

MON810 zygosity was determined by adapting a PCR-based testing method developed by Liu and Chen [23]. Three primers (HECTExfwd, CRYfwd, and HECTupRev3, see Table 3) [11] were used to identify the degree of MON810 zygosity in the maize samples. The primer pair HECTExfwd/HECTupRev3 was used to detect the wild type (no MON810 transgene). CRYfwd and HECTupRev3 were used to identify the 3′-border region of MON810 in genomic and transgenic DNA (in a hemizygous or homozygous state). Maize samples that are homozygous for MON810 will produce only the MON810-specific CRYfwd/HECTupRev3 fragment and not the wild-type fragment. Samples without MON810 will produce only the wild-type HECTExfwd/HECTupRev3 fragment. Samples that are hemizygous for MON810 will produce both the wild type and MON810 fragments.

The PCR was performed using genomic DNA as a template. Each 20 µL reaction contained 4 µL of 5X Green GoTaq® Reaction Buffer (Promega), 250 nM of each forward primer, 500 nM of the reverse primer, 250 µM of each deoxynucleoside triphosphate (dNTP), 1.5 U of GoTaq® G2 DNA Polymerase (Promega), and 60 ng of DNA. A Mastercycler® (Eppendorf) was used for this analysis. The temperature profile is shown in Table 1.

The commercial hybrid NK603 AG8025 was used as a negative control (producing only the wild-type band) and the commercial hybrid PR33P67 (Pioneer Hi-Bred), which contains the MON810 trait in a hemizygous state, was used as a positive control (producing two bands, one for MON810 and one for the wild type).

SNP genotyping using high-resolution melting analysis

The first step of high-resolution melting (HRM) analysis uses real-time PCR to amplify the target sequence [24]. Each 16 µL reaction used the Type-it HRM PCR Kit (Qiagen), 8 µL master mix, 700 nM primers, and 1.6 µL of diluted DNA (40 ng/µL) template. HRM analysis was performed on a Rotor-Gene Q instrument (Qiagen) with a temperature ramping rate of 0.2 °C per 4 s. The initial and final temperatures were 80 and 92 °C. The thermocycling profile used is shown in Table 2. The melting curves and the temperature-shifted curves were normalized to enable sample-to-sample comparisons. Modified curves and HRM scores were obtained using the Rotor-Gene Q series software (version 2.0.2.4, Qiagen). After normalization, the normalized and temperature-shifted melting curves were used to derive the final curve. One sample (per run) was selected as a reference genotype (HRM score of 100%) for the difference plot. The software compared the melting curves between the samples and the reference and generated confidence intervals relative to the selected genotype.

To identify samples with sequence variations in the 180 bp target sequence (Fig. 1), 600 genomic samples containing the MON810 insert were screened by HRM analysis. Amplicons with sequence variations have different melting curve shapes and, therefore, lower HRM scores, which enables the identification of DNA samples containing these variations. Each variety was tested separately and a maximum of 70 genomic DNA samples was analyzed per run (one replicate per run). Every run was repeated on a different day. The mean of the HRM scores was calculated, and in each run, the seven samples with the lowest HRM scores and the reference sample (with an HRM score of 100%) were selected for amplicon sequencing. In total, 24 samples of each variety were sequenced. We have described the use of HRM analysis for variant detection several times [17, 25, 26].

Amplicon sequencing

Amplicon paired-end sequencing of the 180 bp target region (Fig. 1) was performed to characterize samples with low HRM scores in detail. Deep sequencing of amplicons (PCR products) enables the detection and quantification of very low-frequency variants (even frequencies of 1% are detectable), as well as accurate characterization of specific genomic regions [19].

The 16S Metagenomic Sequencing Library Preparation protocol [27] and the protocol of Dobrovolny et al. (2018, under review) were modified to sequence amplicons on an Illumina MiSeq sequencer. According to the 16S

Table 1 Thermal cycling profile of the PCR zygosity experiment

Cycle step	Temperature (°C)	Time	Number of cycles
Initial denaturation	94	10 min	
Denaturation	94	30 s	30
Annealing	54	60 s	
Extension	72	60 s	
Final extension	72	8 min	

Table 2 Thermal cycling profile of the HRM analysis experiment

Cycle step	Temperature (°C)	Time	Number of cycles
Initial denaturation	95	5 min	
Denaturation	95	10 s	40
Annealing	55	30 s	
Extension	72	20 s	
HRM	80–92	+0.2 °C every 4 s	

Table 3 Primers used in this work

Primer	Sequence
HECTExfwd	5′-TCAATCATCAAAGCATCATCG-3′ [11]
CRYfwd	5′-TCTTCACGTCCAGCAATCAG-3′ [11]
HECTupRev3	5′-TTTGGGAAGGAAAAGGTATC-3′ [11]
Amplicon-3m810fw	5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCAAGCACGAGACCGTCAA-3′ modified after [18]
Amplicon-3m810rev	5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCTCGCAAGCAAATTCGGAA-3′ modified after [18]
Nextera XT Index 1 Primers	[28]
Nextera XT Index 2 Primers	[28]
Meth F fw	5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGYGAYGATGTGTTYAAGGAGAA-3′
Meth F rev	5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAAARCCAAACTCARATRAATCAAAA-3′
AmpliconATP1-1 fw	5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGAAYGAGATTYAAGYTGGGGAAATGGT-3′ modified after [30]
AmpliconATP1-1 rev	5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCCTCTTCCATCAATARRTACTCCCA-3′ modified after [30]

Table 4 Thermal cycling profile of the amplicon PCR and the index PCR

Cycle step	Temperature (°C)	Time	Number of cycles
Initial denaturation	95	15 min	
Denaturation	95	30 s	40 for amplicon PCR 8 for index PCR
Annealing	62 for amplicon PCR 55 for index PCR	30 s	
Extension	72	30 s	
Final extension	72	10 min	

Metagenomic Sequencing Library Preparation protocol [27], Illumina overhang adapter sequences were added to the MON810 coding region primer sequences [18] to generate amplicon primers (Amplicon-3m810fw and Amplicon-3m810rev). All primers used in this work are listed in Table 3. Amplicon PCR was performed in a thermal cycler (Eppendorf) to amplify the region of interest. Then, 2.5 µL of genomic DNA (5 ng/µL), 5 µL of each primer (1 µM) (Amplicon-3m810fw and rev, see Table 3), and 12.5 µL of HotStarTaq Master Mix (Qiagen) were used to set up the reaction. The temperature profile of the amplicon PCR is shown in Table 4.

Subsequently, amplicon PCR products were purified using AMPure XP beads to eliminate free primers and primer dimers. Index PCR was used to attach dual indices and sequencing adapters to the purified amplicons. Each index PCR reaction contained 5 µL of purified PCR amplicon product, 5 µL of each primer [Nextera XT Index Primers 1 and 2 (Illumina)] [28], 25 µL of HotStarTaq Master Mix (Qiagen), and 10 µL of ddH_2O. The thermal cycler profile used for index PCR is shown in Table 4. A second clean-up step was performed on the PCR products and the purified products were quantified using the Qubit® 2.0 Fluorometer (Life Technologies). Depending on the size of the DNA amplicon, products were diluted with 10 mM Tris (pH 8.5) to a 4 nM concentration. Five microliters of each diluted sample was pooled together, and 5 µL of this pool was combined with 5 µL 0.2 N NaOH and incubated for 5 min at room temperature to denature the DNA. The denatured DNA was diluted with 990 µL of pre-chilled HT1 buffer and the resulting 20 pM denatured library was stored at − 20 °C until sequencing. For amplicon sequencing, 2 µL of the 20 pM denatured library was spiked into a whole bacterial genome sequencing run (300 bp paired-end sequencing) using the MiSeq Reagent Kit v3 (Illumina). Sequencing data were evaluated using the CLC Genomics Workbench version 10.1.1 (Qiagen).

Bisulfite sequencing

For bisulfite sequencing, 24 selected DNA samples (see section "Methylation status in the coding region of the MON810 *cry1Ab gene*") were bisulfite-converted and purified using the EpiTect® Bisulfite Kit (Qiagen) according to the manufacturer's protocol. Next, amplicon sequencing was performed as described above (section "Amplicon sequencing"). Only the amplicon PCR primers, reagents, and PCR temperature profiles were different. For amplicon PCR with bisulfite-converted samples, a specific primer pair targeting the coding region of the MON810 *cry1Ab* gene was designed using Kismeth [29] and Illumina overhang adapter sequences [27] were added to the primer sequences. The resulting primers are depicted in Table 3 (Meth F fw and Meth F rev). Each 25 µL amplicon PCR reaction mixture contained 50 ng of bisulfite-converted DNA, 5 µL of 5X Colorless GoTaq® Buffer (Promega), 0.625 µL of each primer (10 µM) (Meth F fw and Meth F rev, see Table 3), 250 µM of each dNTP, and 1.875 U of GoTaq® Hot Start

Table 5 Thermal cycling profile of the amplicon PCR and the index PCR with bisulfite-converted samples

Cycle step	Temperature (°C)	Time	Number of cycles
Initial denaturation	94	10 min	
Denaturation	94	30 s	40 for amplicon PCR
Annealing	50	30 s	8 for index PCR
Extension	72	45 s	
Final extension	72	10 min	

Table 6 Thermal cycling profile of the Amplicon PCR and the Index PCR for the endogenous gene control

Cycle step	Temperature (°C)	Time	Number of cycles
Initial denaturation	95	15 min	
Denaturation	95	30 s	35 for amplicon PCR
Annealing	54	30 s	8 for index PCR
Extension	72	45 s	
Final extension	72	10 min	

DNA Polymerase (Promega). Amplicon PCR products were purified as described in the "Amplicon sequencing" section. The purified amplicons were used as templates for index PCR. Each 50 µL index PCR reaction contained 5 µL of purified amplicon PCR product, 5 µL of each primer [Nextera XT Index Primers 1 and 2 (Illumina)], 10 µL of 5X Colorless GoTaq® Buffer (Promega), 3.75 U of GoTaq® Hot Start DNA Polymerase (Promega), and 250 µM of each dNTP. Thermal cycling temperature profiles are shown in Table 5. Afterwards, the index PCR samples were treated as described in the "Amplicon sequencing" section. CG, CHG, and CHH methylation patterns were analyzed and displayed using the CLC Genomics Workbench version 10.1.1 (Qiagen).

Bisulfite conversion frequency

To determine the bisulfite conversion rate, a universal endogenous control gene developed by Wang et al. [30] was used. Illumina overhang adapter sequences were added to the sequences specific to the endogenous ATP1-1 control gene. The resulting primers are shown in Table 3 (AmpliconATP1-1 fw and rev). Amplicon PCR and index PCR were set up, as described in "Amplicon sequencing" section, and differed only by the specific primers used. The thermal cycler temperature profiles for amplicon PCR and index PCR are shown in Table 6. After index PCR, control samples were treated as described in "Amplicon sequencing" section. The bisulfite conversion frequency was calculated using the CLC Genomics Workbench version 10.1.1 (Qiagen).

Results

Genetic characterization of the MON810 *cry1Ab* coding region

First, we screened all samples by HRM analysis to detect mutations in the 180 bp target region. The mean of the HRM scores of two HRM runs was calculated and the samples with the lowest mean HRM scores as well as the reference samples were chosen for amplicon sequencing (a total of 72 samples, 24 of each variety). The mean HRM scores of these samples are presented in Table 7. Next, the 3′ flanking region of MON810 was analyzed for sequence variations by amplicon sequencing. Insert Table 7.

The 72 sequenced samples (Table 7) generated between 5308 and 26,990 sequencing reads. All sequenced samples passed a quality filter (trimming) to eliminate sequencing errors. Thereby, between 0 and 289 reads were removed for each sample. Reads were mapped to the reference sequence and the CLC Variant Detector was used for SNP calling. The minimum variant frequency was set to 1% to prevent false-positive variants due to sequencing signal noise. Table 7 gives an overview in which of selected samples variations were detected. Seventeen of the 72 sequenced samples had a variant in the coding region of MON810 *cry1Ab* with a minimum variant frequency of 1% (see Fig. 2). Fifteen of the 17 samples with a variant harbored the previously described variant (T instead of C) at position 71 of the analyzed region. However, the frequency of the variant at position 71 ranged from 1.2 to 53.2%. 631RR2/Bt samples 63, 68, 74, and 207 had the highest frequencies of variant 71 (31.7, 53.2, 52.2, and 26.8%). Low-frequency (<2%) transitions were detected at positions 39 (C → T), 50 (G → A), 53 (C → T), 71 (C → T), 79 (C → T), 83 (C → T), and 91 (G → A). Of the 17 samples with a variant, 14 were derived from the stacked MON810 event (631RR2/Bt), two were from the unspecified (single event) MON810 variety, and one was from the (single event) hybrid DKB 350YG. An overview of the results from the SNP analysis is depicted in Table 8.

Taken together, noticeably, more samples from the stacked MON810 event (631RR2/Bt) had a variant in the *cry1Ab* coding region and at higher frequencies compared to samples from the single event varieties (DKB 350YG and the unspecified MON810 variety). The variant frequencies of samples from sequenced single events were <2%, whereas the samples from the stacked event variety had variant frequencies up to 53.2% (see Table 8). Furthermore, the heterozygous variant at position 71 (C → T) was the most common variant in all of the tested samples and had

Table 7 Results of HRM analysis and amplicon sequencing

631RR2/Bt samples	Mean HRM score (%)	SNP detection*	DKB 350YG samples	Mean HRM score (%)	SNP detection*	Unspecified MON810 samples	Mean HRM score (%)	SNP detection*
82	4.07	✓	242	52.74	✗	22	23.51	✗
130	28.93	✓	136	56.24	✗	282	38.62	✗
68	51.05	✓	66	56.73	✗	296	49.56	✗
74	51.61	✓	185	66.18	✗	46	50.11	✗
117	54.28	✓	187	66.33	✗	41	53.12	✗
121	60.35	✓	190	69.01	✗	40	54.53	✗
35	64.80	✗	216	70.03	✗	34	55.82	✓
129	64.82	✓	191	71.34	✗	44	55.98	✗
207	64.92	✓	181	72.61	✗	5	57.31	✓
81	67.02	✓	194	74.73	✗	323	60.10	✗
72	68.68	✓	48	75.71	✓	234	61.94	✗
1	69.35	✗	163	75.83	✗	116	62.61	✗
120	69.82	✓	184	76.12	✗	206	63.07	✗
41	70.13	✗	186	76.39	✗	187	63.12	✗
63	70.49	✓	138	76.59	✗	269	66.82	✗
205	79.58	✗	141	76.70	✗	126	68.58	✗
182	80.22	✗	219	78.23	✗	319	68.79	✗
206	81.48	✗	155	78.29	✗	435	77.24	✗
158	83.35	✗	126	79.10	✗	444	77.57	✗
181	83.77	✓	189	80.50	✗	446	82.82	✗
201	84.12	✓	54	89.00	✗	402	83.26	✗
39	100.00	✗	123	100.00	✗	14	100.00	✗
108	100.00	✗	253	100.00	✗	276	100.00	✗
173	100.00	✗	225	100.00	✗	292	100.00	✗

The mean values of the obtained HRM scores are shown. Low HRM scores indicate mutations

*By amplicon sequencing

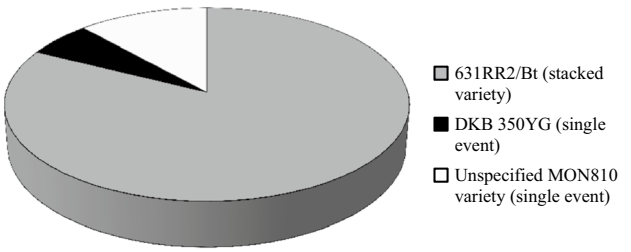

Fig. 2 Distribution of samples with a sequence variant in the MON810 *cry1Ab* coding region. The stacked variety 631RR2/Bt had the highest variant frequency

the highest variant frequencies (between 1.2 and 53.2%). The cytosine-to-thymine transition at position 71 does not lead to an amino acid change (silent mutation) [17] and the low-frequency variants at positions 50, 53, and 83 are also silent mutations. In contrast, the low-frequency variants at positions 39, 79, and 91 result in amino acid changes. The C → T transition at position 39 results in the translation of a phenylalanine (F) instead of a leucine (L), the C → T transition at position 79 leads to a valine (V) instead of an alanine (A), and the G → A transition at position 91 results in an arginine (R) instead of a histidine (H). It was striking that all variants were either C → T or G → A transitions. C → T transitions are very frequent at 5-methylcytosine sites. This is because deamination of 5-methylcytosine leads to thymine, which is not recognized by the enzyme uracil-DNA glycosylase and, therefore, not repaired [31]. This prompted us to investigate the methylation status of this region.

Methylation status of the MON810 *cry1Ab* coding region

Bisulfite sequencing was used to investigate the methylation status of the MON810 *cry1Ab* coding region in detail. The bisulfite-sequenced region was 345 bp, which included the 180 bp region used for SNP characterization (Fig. 1). Position 71 of the 180 bp fragment appears to be a mutational hotspot, which prompted our detailed

Table 8 Overview of detected variants in the three tested varieties

	SNP 39 (C → T)	SNP 50 (G → A)	SNP 53 (C → T)	SNP 71 (C → T)	SNP 79 (C → T)	SNP 83 (C → T)	SNP 91 (G → A)
631RR2/Bt samples							
63	–	–	–	31.71	–	–	–
68	–	–	–	53.15	–	–	–
72	–	–	–	6.15	–	–	–
74	–	–	–	52.24	–	–	–
81	–	–	–	4.27	–	–	–
82	–	–	1.42	1.16	–	–	–
117	–	–	–	2.61	–	–	–
120	–	–	–	1.28	–	–	–
121	1.11	1.24	–	1.68	–	–	–
129	–	–	–	1.74	–	–	–
130	–	–	–	12.68	1.26	–	–
181	–	–	–	–	–	1.17	–
201	–	–	–	1.78	–	–	–
207	–	–	–	26.80	1.25	–	–
DKB 350YG samples							
48	–	–	–	–	–	–	1.46
Unspecified MON810 samples							
5	–	–	–	1.60	–	–	–
34	–	–	–	1.87	–	–	–

The upper column shows the position and type of variant. Frequencies of the different variants are depicted in %

investigation of its methylation status. In total, eight samples of each variety were bisulfite-sequenced. For variety 631RR2/Bt with stacked transgenes, four samples with a high frequency of variants at position 71 and four samples without any nucleotide variations were selected for bisulfite sequencing. Only samples without variants were chosen for methylation analysis from the DKB 350YG and unspecified MON810 varieties. The analyzed samples generated between 4612 and 49,062 reads. A maximum of two reads was removed via quality trimming for each sample. Bisulfite reads were mapped to the reference sequence. Methylation levels were quantified using the Bisulfite Sequencing tool of the CLC Genomics Workbench (version 10.1.1). The analyzed region (excluding primer regions) included 25 cytosines in CpG motifs, 22 cytosines in CHG (where H is an A or T) motifs and 71 cytosines in CHH (where H is any base other than G) motifs. Overall, the average methylation level of the 24 bisulfite-sequenced samples was low (0.83% overall methylation, with 0.86% CpG, 0.78% CHG, and 0.86% CHH). The bisulfite conversion frequency averaged 97.9%. Figure 3 shows the methylation levels for each individual variety.

To determine if DNA methylation differed significantly, we compared the methylation levels in the different motifs (CpG, CHG, and CHH)

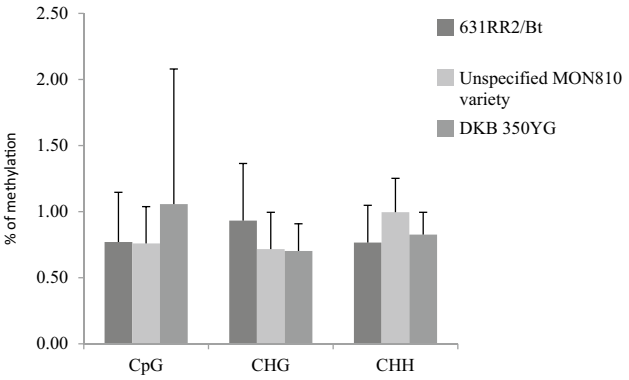

Fig. 3 Cytosine methylation levels in the *cry1Ab* coding region and the 3′ flanking region among the three tested MON810 maize varieties. The levels of cytosine methylation were measured in the CG, CHG, and CHH motifs. Vertical bars show the positive standard deviation

1. among the three different varieties (631RR2/Bt, DKB 350YG and the unspecified MON810 variety);
2. between 631RR2/Bt samples with a high-frequency variant and 631RR2/Bt samples without a variant.

In addition, we compared the methylation levels exclusively at position 71 (a CHH motif) between 631RR2/Bt

samples with a high-frequency variant and 631RR2/Bt samples without a variant.

There were no statistically significant differences among the three different varieties (ANOVA, $p \leq 0.05$). No significant differences were detected between the tested 631RR2/Bt samples either (ANOVA, $p \leq 0.05$). Interestingly, the methylation status of the analyzed region was not correlated with the presence or the absence of sequence variants. To gain a better understanding of the genetic structure of the MON810 locus, the insert zygosity was investigated in the three different varieties (631RR2/Bt, DKB 350YG, and the unspecified MON810 variety).

MON810 zygosity

MON810 hybrid maize grains can have different genetic structures: hemizygous (produced by a transgenic male or female parent) or homozygous at the transgenic locus [32]. If the genetic structure is homozygous at the transgene locus, it would theoretically be possible to detect two different alleles for every nucleotide position of MON810 (one maternal and one paternal). Since we discovered different alleles in our samples, MON810 zygosity was determined by conventional PCR as described in the "Materials and methods". Overall, 20 samples of each variety were investigated to identify the MON810 insert zygosity. All samples analyzed produced two bands on the agarose gel, the wild-type band, and the MON810-specific band, which means that they were hemizygous for the transgene locus. Therefore, we assumed that all of the analyzed F1 seeds from 631RR2/Bt, DKB350YG, and the unspecified MON810 variety were hemizygous for the MON810 insert.

Conclusions and discussion

Although there are already studies describing the molecular characterization of MON810 maize [10–12, 17, 18, 33–37], the complex integration pattern of the MON810 event [10, 11], recent SNP findings in commercial MON810 maize [17], and a fragile 45S rDNA phenotype in a stacked MON810 variety [36], as well as its frequent use worldwide [1], clearly indicate that MON810 maize varieties warrant further investigation. Currently, NGS is the method of choice for a detailed molecular characterization of transgenic events. NGS has many advantages over traditional Southern blot analysis; for example, NGS is much better at identifying incomplete and multiple integration events in complex regions [16, 38].

Deep sequencing of amplicons from the *cry1Ab* coding region and its flanking region verified the frequent presence of a heterozygous C → T transition at position 71, which was reported previously [17]. In contrast to our earlier study [17], NGS was used, and therefore, it was possible to quantify the frequencies of this variant. Additional maize varieties were also analyzed, allowing us to directly compare the variant frequencies between the different varieties and between single and stacked events. We observed more samples with coding region variants and much higher variant frequencies in the stacked MON810 event (631RR2/Bt) than in the single events (DKB 350YG and the unspecified MON810 variety). The occurrence of a T at position 71 was up to 53% in samples from variety 631RR2/Bt compared to < 2% in samples from the unspecified MON810 variety. Thus, the occurrence of sequence variants differs markedly between single and stacked events. Whether the stacking of transgenes alone is the reason for the increased mutation rate in the tested stacked event has to be clarified in further experiments. Possibly, other factors (e.g., age of the seeds) may contribute to genetic differences as well.

Interestingly, all identified variants were either C → T or G → A transitions. Furthermore, four of the seven identified variants depicted in Table 8 were silent mutations. These were located at the third position of the codon, which is known to be degenerated. Many amino acids are encoded by more than one codon. Such codons are called synonymous codons [39]. Although synonymous codons encode the same amino acid, codon usage is not random. Codon usage influences gene expression levels, protein folding, and protein cellular function [39, 40]. Thus, also silent mutations can have huge impact on proteins. The recombinant *cry1Ab* toxin of MON810 maize seems to be significantly stronger expressed compared to natural *cry1Ab* toxins due to modified codon usage prior to transgene insertion [41]. Protein folding and toxin activation may also be influenced in the recombinant *cry1Ab* toxin due to modified codon usage, but also because of major changes on the DNA level [42]. Synonymous codon usage patterns can vary markedly among genomes, but also within a genome. The evolutionary pressure on codon bias is explained by selection for translational efficiency, GC mutational bias, and by mRNA secondary-structure stability [40]. *Cry* genes are AT-rich compared to endogenous plant genes. This preference can lead to serious problems such as low expression if a *cry* gene becomes inserted into a plant organism. To enable stable and strong transgene expression, codon optimization is commonly performed prior to transgene insertion [43]. This was also performed for the *cry1Ab* gene; several A and T nucleotides were replaced by C and G nucleotides, which are preferred codon endings in *Zea maize* [41, 44]. In view of this, it was remarkable that the variants we have detected are only A and T variants, which does not reflect the distribution of nucleotides in the *Z. mays* background.

Another important result from our study is the determination of the zygosity of the MON810 transgene. We found that the transgenic locus was hemizygous in all three tested

varieties. This means that the analyzed maize grains inherited the transgene only from one parent (a transgenic male or female parent). We detected two alleles at single nucleotide positions of the MON810 transgene in individual maize grain samples. If the maize varieties were homozygous at the transgene locus, it would theoretically be possible to detect two different alleles at single nucleotide positions. Given the fact that the analyzed maize grains were hemizygous for the transgene locus, somatic mutations might explain the identification of two alleles at single positions in the MON810 transgene. Moreover, it has to be noted that every single seed of an ear is an independent event [45]. Thus, a mutation in pollen grain or ovule might also be the cause of the observed mutations. Moreover, in many plant species, genetic mutations and changes in chromosome structure in embryo meristems increase as the seeds age. The mutation rate can increase substantially in later developmental phases (seed senescence) [46]. Since we analyzed older maize grains, this may have influenced the sequence variations that we observed.

The identification of two alleles at single positions of the MON810 transgene may also be explained by the existence of a second (possibly only partial) copy of the recombinant *cry1Ab* gene, which contains a thymine instead of a cytosine at position 71. A BLAST search of the primer sequences (Amplicon-3m810fw and rev) used for SNP characterization revealed that the reverse primer can also bind to maize mitochondrial DNA. This is because the last base of the transgene and the first 44 bp of the 3′ flanking region of MON810 are 100% homologous to maize mitochondrial DNA (see Fig. 1). This additional, most-likely unintended, insertion of mitochondrial DNA has become a functional part of the *cry1Ab* coding region since it contains the stop codon for the *cry1Ab* gene. However, the forward primer can only bind to the *cry1Ab* coding region, so it can be assumed that this primer pair is specific to only one locus in the MON810 varieties' genomes (providing that no other copies are present). Monsanto states that only one functional copy of the *cry1Ab* gene is present in the maize genome and other studies confirmed the presence of only one MON810 insert in the MON810 maize genome [12, 13]. Nevertheless, since particle bombardment frequently results in variable copy numbers of rearranged transgene fragments [7], it is possible that small non-functional fragments of the *cry1Ab* coding sequence coupled to the mitochondrial DNA sequence are also contained in the MON810 maize genome and that these fragment(s) have been overlooked until now. Mandatory transgene characterization might be complicated by these sequence variations. Consequently, transgene presence and frequency may be underestimated [47].

Most remarkably, position 71 of the *cry1Ab* sequence seems to be a mutational hotspot in at least three MON810 maize varieties (631RR2/Bt, the unspecified MON810 variety, and variety 4421VT3 [17]). Mutational hotspots can arise through the pro-mutagenic activities of cytosine modifications such as methylation, deamination, and halogenation [20]. DNA methylation of cytosine in symmetrical CG motifs is the most common methylation pattern in plants. However, in plants cytosine can also be methylated at symmetrical CHG (H is A or T) or asymmetrical CHH (H is A, C or T) sites [30]. Moreover, C → T transitions frequently occur at 5-methylcytosine sites because deamination of 5-methylcytosine leads to thymine, which is not recognized by the enzyme uracil-DNA glycosylase and, therefore, not repaired [31]. Since all detected variants were either C → T or the complementary G → A transitions, we analyzed the methylation status of the *cry1Ab* coding region and its 3′ flanking genomic region by bisulfite sequencing. However, the methylation status was very low (0.83% overall methylation, of which 0.86% were CpG, 0.78% were CHG, and 0.86% were CHH) and no significant differences were detected among the varieties or between samples with or without variants. Therefore, we conclude that the presence of variants in the *cry1Ab* locus is not influenced by methylation status.

The methylation status of stacked versus single events (MON-89Ø34-3 and MON-ØØ6Ø3-6) was previously investigated [48] and differences in cytosine methylation levels were found in the FMV promoter and *cry2Ab2* transgene in four Bt-expressing hybrid varieties. Comparing single and stacked events in the same genetic background also revealed differences in the 35S promoter sequence and the results of transgene transcript accumulation levels showed differences in both *cry1A.105* and *cry2Ab2* transgenes among the four Bt-expressing hybrid varieties. However, only the methylation level of the *cry2Ab2* transgene differed between single and stacked events. Such differences are important for the ongoing discussion about whether stacked events should be regulated as conventional hybrids or as new GM plants [49]. If they are regulated as conventional hybrids, it is sufficient to refer to the validation parameters for each individual single event. In contrast, if they are regulated as new GM plants, the risk assessment has to be repeated with plants materials from stacked events.

Spontaneous mutations can occur in all living organisms during DNA replication, homologous or intrachromosomal recombination, mitosis and meiosis [9, 18]. Furthermore, hemizygous DNA architectures show considerably more ectopic recombination events compared to homozygous DNA architectures, which can lead to severe chromosomal rearrangements [50]. In addition, transposable elements and repetitive DNA sequences have the ability to alter genomes, e.g., create or reverse mutations. Transposable elements can also increase gene copy number or even lead to the creation of new genes through exon shuffling [51, 52]. Moreover, DNA repeats can cause instabilities via homologous

recombination. Mobile elements and repetitive DNA strongly influence genome diversity [53]. Due to plasticity, plant genomes have the ability to reorganize themselves at the DNA level, but also at the chromatin level [54]. It needs to be clarified, if transgene stacking increases genetic reorganization, which would mean an increased mutation rate compared to the natural mutation rate.

The results of our study support the statement by Waminal et al. [36]: "In the context of genetic engineering, it may be useful to note that transformation itself is known to be mutagenic, and during plant transformation, each transformation event receives a unique transgene integration pattern, and thus, a unique accompanying genomic rearrangement. Thus, it is logical to treat each event individually and uniquely". We also share the opinion that every event should be examined and evaluated separately. Since in GM risk assessment, toxicological tests are usually performed with *cry* proteins purified from laboratory strains of bacteria engineered to express *cry* protein instead of using *cry* proteins extracted from the assessed GM plant [42], toxicological relevant issues may be overlooked. This is especially relevant when evaluating stacked events, which may contain genetic differences in the transgene compared to single events. In view of this, we recommend a separate assessment for stacked events, which includes DNA-sequencing data of the transgenes.

Given our results, we have to clarify, if transgene stacking has an influence on the genetic stability, which would justify a separate evaluation of single and stacked events in the authorization of GM plants. Furthermore, molecular mechanisms responsible for the differences between single and stacked events have to be identified. Nevertheless, transgene integration processes can have unpredictable plant-wide effects on different processes and signaling pathways due to the interactions of biological networks [55]. These effects can be observed on the genomic, epigenomic, transcriptomic, proteomic, or metabolomic levels. To better understand interactions in plant biological networks, to improve breeding techniques, and to perform adequate risk assessments, it is important to identify and investigate such unintended effects.

Acknowledgements We would like to express our gratitude to all three anonymous reviewers for their helpful and detailed comments. Material costs for this work were partially supported by the Hochschuljubiläumsstiftung (HJST) of the city of Vienna.

Compliance with ethical standards

Conflict of interest All the authors declare that they have no competing interest.

Ethics requirements This article does not contain any studies with human or animal subjects.

References

1. ISAAA (2016) Global Status of Commercialized Biotech/GM Crops: 2016. ISAAA. http://www.isaaa.org/purchasepublications/itemdescription.asp?ItemType=BRIEFS&Control=IB052-2016. Accessed 26 Sept 2017
2. Monsanto (2017) Safety Assessment of YieldGard Insect-Protected Corn Event MON 810. Monsanto, https://monsanto.com/company/media/statements/mon810-bans/. Accessed 30 Oct 2017
3. Trtikova M, Wikmark OG, Zemp N, Widmer A, Hilbeck A (2015) Transgene expression and Bt protein content in transgenic Bt maize (MON810) under optimal and stressful environmental conditions. PLoS One 10(4):e0123011. https://doi.org/10.1371/journal.pone.0123011
4. Armstrong CL, Green CE, Phillips RL (1991) Development and availability of germplasm with high Type II culture formation response. Maize Genet Cooper Newsl 65:92–93
5. Haslberger AG (2003) Codex guidelines for GM foods include the analysis of unintended effects. Nat Biotech 21(7):739–741
6. Lowe BA, Shiva Prakash N, Way M, Mann MT, Spencer TM, Boddupalli RS (2009) Enhanced single copy integration events in corn via particle bombardment using low quantities of DNA. Transgenic Res 18(6):831–840. https://doi.org/10.1007/s11248-009-9265-0
7. Yin Z, Plader W, Malepszy S (2004) Transgene inheritance in plants. J Appl Genet 45(2):127–144
8. Kohli A, Griffiths S, Palacios N, Twyman RM, Vain P, Laurie DA, Christou P (1999) Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology mediated recombination. Plant J Cell Mol Biol 17(6):591–601
9. Schnell J, Steele M, Bean J, Neuspiel M, Girard C, Dormann N, Pearson C, Savoie A, Bourbonnière L, Macdonald P (2015) A comparative analysis of insertional effects in genetically engineered plants: considerations for pre-market assessments. Transgenic Res 24(1):1–17. https://doi.org/10.1007/s11248-014-9843-7
10. Hernandez M, Pla M, Esteve T, Prat S, Puigdomenech P, Ferrando A (2003) A specific real-time quantitative PCR detection system for event MON810 in maize YieldGard based on the 3′-transgene integration sequence. Transgenic Res 12(2):179–189
11. Rosati A, Bogani P, Santarlasci A, Buiatti M (2008) Characterisation of 3′transgene insertion site and derived mRNAs in MON810 yield gard maize. Plant Mol Biol. https://doi.org/10.1007/s11103-008-9315-7
12. Aguilera M, Querci M, Pastor S, Bellocchi G, Milcamps A, Van den Eede G (2009) Assessing copy number of MON 810 integrations in commercial seed maize varieties by 5′ event-specific real-time PCR validated method coupled to 2-ΔΔCT analysis. Food Anal Methods 2(1):73–79. https://doi.org/10.1007/s12161-008-9036-1
13. Corbisier P, Bhat S, Partis L, Xie VR, Emslie KR (2010) Absolute quantification of genetically modified MON810 maize (*Zea mays* L.) by digital polymerase chain reaction. Anal Bioanal Chem 396(6):2143–2150. https://doi.org/10.1007/s00216-009-3200-3
14. Privalle LS, Chen J, Clapper G, Hunst P, Spiegelhalter F, Zhong CX (2012) Development of an agricultural biotechnology crop product: testing from discovery to commercialization. J Agric Food Chem 60(41):10179–10187. https://doi.org/10.1021/jf302706e
15. EFSA (2011) Scientific Opinion on Guidance for risk assessment for food and feed from genetically modified plants. EFSA J. https://doi.org/10.2903/j.efsa.2011.2150
16. Guttikonda SK, Marri P, Mammadov J, Ye L, Soe K, Richey K, Cruse J, Zhuang M, Gao Z, Evans C, Rounsley S, Kumpatla SP

(2016) Molecular characterization of transgenic events using next generation sequencing approach. PLoS One 11(2):e0149515. https://doi.org/10.1371/journal.pone.0149515
17. Ben Ali SE, Madi ZE, Hochegger R, Quist D, Prewein B, Haslberger AG, Brandes C (2014) Mutation scanning in a single and a stacked genetically modified (GM) event by real-time PCR and high resolution melting (HRM) analysis. Int J Mol Sci 15(11):19898–19923. https://doi.org/10.3390/ijms151119898
18. Neumann G, Brandes C, Joachimsthaler A, Hochegger R (2011) Assessment of the genetic stability of GMOs with a detailed examination of MON810 using Scorpion probes. Eur Food Res Technol 233(1):19–30. https://doi.org/10.1007/s00217-011-1487-8
19. Illumina (2015) Highly targeted resequencing of regions of interest. https://www.illumina.com/techniques/sequencing/dna-sequencing/targeted-resequencing/amplicon-sequencing.html. Accessed 20 Jul 2017
20. Sassa A, Kanemaru Y, Kamoshita N, Honma M, Yasui M (2016) Mutagenic consequences of cytosine alterations site-specifically embedded in the human genome. Genes Environ 38(1):17. https://doi.org/10.1186/s41021-016-0045-9
21. Kurdyukov S, Bullock M (2016) DNA methylation analysis: choosing the right method. Biology 5(1):3. https://doi.org/10.3390/biology5010003
22. ThermoFisherScientific Qubit® 3.0 Fluorometer. https://www.thermofisher.com/at/en/home/industrial/spectroscopy-elemental-isotope-analysis/molecular-spectroscopy/fluorometers/qubit/qubit-fluorometer.html. Accessed 22 Jun 2017
23. Liu N, Chen H (2010) An accurate and rapid PCR-based zygosity testing method for genetically modified maize. Mol Plant Breed 7(3):619–623
24. Druml B, Cichna-Markl M (2014) High resolution melting (HRM) analysis of DNA—its role and potential in food analysis. Food Chem 158:245–254. https://doi.org/10.1016/j.foodchem.2014.02.111
25. Madi ZE, Brandes C, Neumann G, Quist D, Ruppitsch W, Hochegger R (2013) Evaluation of Adh1 alleles and transgenic soybean seeds using Scorpion PCR and HRM analysis. Eur Food Res Technol 237(2):125–135. https://doi.org/10.1007/s00217-013-1969-y
26. Castan M, Ben Ali S-E, Hochegger R, Ruppitsch W, Haslberger AG, Brandes C (2017) Analysis of the genetic stability of event NK603 in stacked corn varieties using high-resolution melting (HRM) analysis and Sanger sequencing. Eur Food Res Technol 243(3):353–365. https://doi.org/10.1007/s00217-016-2749-2
27. Illumina (2013) 16S metagenomic sequencing library preparation. https://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html. Accessed 22 Jun 2017
28. Illumina (2016) Illumina adapter sequences. https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/experiment-design/illumina-adapter-sequences_1000000002694-01.pdf. Accessed 28 Jul 2017
29. Gruntman E, Qi Y, Slotkin RK, Roeder T, Martienssen RA, Sachidanandam R (2008) Kismeth: Analyzer of plant methylation states through bisulfite sequencing. BMC Bioinf 9(1):371. https://doi.org/10.1186/1471-2105-9-371
30. Wang J, Wang C, Long Y, Hopkins C, Kurup S, Liu K, King GJ, Meng J (2011) Universal endogenous gene controls for bisulphite conversion in analysis of plant DNA methylation. Plant Methods 7:39–39. https://doi.org/10.1186/1746-4811-7-39
31. Griffiths A, Gelbart W, Miller J (1999) Modern genetic analysis. W. H. Freeman, New York
32. Zhang D, Corlet A, Fouilloux S (2008) Impact of genetic structures on haploid genome-based quantification of genetically modified DNA: theoretical considerations, experimental data in MON 810 maize kernels (*Zea mays* L.) and some practical applications. Transgenic Res 17(3):393–402. https://doi.org/10.1007/s11248-007-9114-y
33. la Paz JL, Pla M, Papazova N, Puigdomènech P, Vicient CM (2010) Stability of the MON 810 transgene in maize. Plant Mol Biol 74(6):563–571. https://doi.org/10.1007/s11103-010-9696-2
34. Aguilera M, Querci M, Balla B, Prospero A, Ermolli M, Van den Eede G (2008) A qualitative approach for the assessment of the genetic stability of the MON 810 trait in commercial seed maize varieties. Food Anal Methods 1(4):252–258. https://doi.org/10.1007/s12161-008-9035-2
35. la Paz JL, Pla M, Centeno E, Vicient CM, Puigdomènech P (2014) The use of massive sequencing to detect differences between immature embryos of MON810 and a comparable non-GM maize variety. PLoS One. https://doi.org/10.1371/journal.pone.0100895
36. Waminal NE, Ryu KH, Choi S-H, Kim HH (2013) Randomly detected genetically modified (GM) maize (*Zea mays* L.) near a transport route revealed a fragile 45S rDNA phenotype. PLoS One 8(9):e74060. https://doi.org/10.1371/journal.pone.0074060
37. Singh CK, Ojha A, Kamle S, Kachru DN (2007) Assessment of cry1Ab transgene cassette in commercial Bt corn MON810: gene, event, construct and GMO specific concurrent characterization. Protocol Exch. https://doi.org/10.1038/nprot.2007.440
38. Zhang R, Yin Y, Zhang Y, Li K, Zhu H, Gong Q, Wang J, Hu X, Li N (2012) Molecular characterization of transgene integration by next-generation sequencing in transgenic cattle. PLoS One 7(11):e50348. https://doi.org/10.1371/journal.pone.0050348
39. Chen SL, Lee W, Hottes AK, Shapiro L, McAdams HH (2004) Codon usage between genomes is constrained by genome-wide mutational processes. Proc Natl Acad Sci USA 101(10):3480–3485. https://doi.org/10.1073/pnas.0307827100
40. Angov E (2011) Codon usage: nature's roadmap to expression and folding of proteins. Biotechnol J 6(6):650–659. https://doi.org/10.1002/biot.201000332
41. van der Salm T, Bosch D, Honée G, Feng L, Munsterman E, Bakker P, Stiekema WJ, Visser B (1994) Insect resistance of transgenic plants that express modified *Bacillus thuringiensis* cryIA(b) and cryIC genes: a resistance management strategy. Plant Mol Biol 26(1):51–59. https://doi.org/10.1007/bf00039519
42. Latham JR, Love M, Hilbeck A (2017) The distinct properties of natural and GM cry insecticidal proteins. Biotechnol Genet Eng Rev 33(1):62–96. https://doi.org/10.1080/02648725.2017.1357295
43. Khan MS, Musatafa G, Nazir S, Joyia FA (2016) Applied molecular biotechnology: the next generation of genetic engineering. Plant molecular biotechnology: applications of transgenics. Taylor & Francis Group, New York
44. Liu H, He R, Zhang H, Huang Y, Tian M, Zhang J (2010) Analysis of synonymous codon usage in *Zea mays*. Mol Biol Rep 37(2):677–684. https://doi.org/10.1007/s11033-009-9521-7
45. Giroux MJ, Shaw J, Barry G, Cobb BG, Greene T, Okita T, Hannah LC (1996) A single mutation that increases maize seed weight. Proc Natl Acad Sci 93(12):5824–5829
46. D'Amato F (1997) Role of somatic mutations in the evolution of higher plants. Caryologia 50(1):1–15. https://doi.org/10.1080/00087114.1997.10797380
47. Agapito-Tenfen SZ, Wickson F (2017) Challenges for transgene detection in landraces and wild relatives: learning from 15 years of debate over GM maize in Mexico. Biodivers Conserv. https://doi.org/10.1007/s10531-017-1471-0
48. Vilperte V, Agapito-Tenfen SZ, Wikmark O-G, Nodari RO (2016) Levels of DNA methylation and transcript accumulation in leaves of transgenic maize varieties. Environ Sci Eur 28(1):29. https://doi.org/10.1186/s12302-016-0097-2
49. Rocca E, Andersen F (2017) How biological background assumptions influence scientific risk evaluation of stacked genetically modified plants: an analysis of research hypotheses and argumentations. Life Sci Soc Policy 13(1):11. https://doi.org/10.1186/s40504-017-0057-7

50. Sun XQ, Li DH, Xue JY, Yang SH, Zhang YM, Li MM, Hang YY (2016) Insertion DNA accelerates meiotic interchromosomal recombination in *Arabidopsis thaliana*. Mol Biol Evol 33(8):2044–2053. https://doi.org/10.1093/molbev/msw087
51. Ayarpadikannan S, Kim H-S (2014) The impact of transposable elements in genome evolution and genetic instability and their implications in various diseases. Genom Inf 12(3):98–104. https://doi.org/10.5808/GI.2014.12.3.98
52. Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. Curr Opin Genet Dev 15(6):621–627. https://doi.org/10.1016/j.gde.2005.09.010
53. Mehrotra S, Goyal V (2014) Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. Genom Proteom Bio 12(4):164–171. https://doi.org/10.1016/j.gpb.2014.07.003
54. Weber N, Halpin C, Hannah LC, Jez JM, Kough J, Parrott W (2012) Editor's choice: crop genome plasticity and its relevance to food and feed safety of genetically engineered breeding stacks. Plant Physiol 160(4):1842–1853. https://doi.org/10.1104/pp.112.204271
55. Buiatti M, Christou P, Pastore G (2013) The application of GMOs in agriculture and in food production for a better nutrition: two different scientific points of view. Genes Nutr 8(3):255–270. https://doi.org/10.1007/s12263-012-0316-4