Anal Bioanal Chem (2010) 396:2125–2133
DOI 10.1007/s00216-009-3176-z

ORIGINAL PAPER

Characterization of polyadenylated *cryIA(b)* transcripts in maize MON810 commercial varieties

José Luis La Paz · Carlos Vicient · Pere Puigdomènech · Maria Pla

Received: 23 June 2009 / Revised: 18 September 2009 / Accepted: 20 September 2009 / Published online: 21 October 2009

Abstract The *Zea mays* L. event MON810 is one of the major commercialized genetically modified crops. The inserted expression cassette has a 3′ truncation partially affecting the *cryIA(b)* coding sequence, resulting in the lack of the NOS terminator, with transcription of the transgene reported to read-through 3′-past the truncation site. Here, we demonstrate that the *cryIA(b)* transgene gives rise to a variety of polyadenylated transcripts of different sizes that extend to around 1 kbp downstream the truncation site. A Stop codon at position +7 downstream the truncation site indicates the production of a transgenic protein with two additional amino acids; which is compatible with the reported size of the CryIA(b) protein in MON810. There is no evidence of the existence of other translated products. Several main 3′ transcription termination regions were detected close to the truncation site and in the transgene 3′ flanking sequence. Next to these main termination sites, we identified some sequence motifs that could potentially act as 3′-end-processing elements and drive termination of the transgene transcripts. The MON810 transgene has been introduced into different commercial varieties through breeding programs. Here, we demonstrate that there are no significant differences among the levels of transgene mRNA accumulation, major transcript sizes and 3′ termini profiles comparing a number of MON810 commercial varieties grown under similar environmental conditions. Commercial varieties of this event appear to be stable in terms of transgene expression.

Keywords GMO (genetically modified organism) · MON810 maize · Transgene mRNA · 3′-end-processing site · Polyadenylated mRNA

Abbreviations

CaMV	Cauliflower mosaic virus
BAC	Bacterial artificial chromosome
NOS	Nopaline synthetase
GMO	Genetically modified organism
mRNA	Messenger ribonucleic acid
CRM	Certified reference material
IRMM	Institute for Reference Materials and Measurements
Real-time RT-PCR	Reverse transcription coupled to real-time PCR
PCR	Polymerase chain reaction
cDNA	Complementary deoxyribonucleic acid
RACE 3′-PCR	Rapid amplification of cDNA ends

J. L. La Paz · C. Vicient · P. Puigdomènech
Departament Genètica Molecular,
Centre de Recerca en Agrigenòmica, CSIC-IRTA-UAB,
Jordi Girona, 18,
08034 Barcelona, Spain

J. L. La Paz
e-mail: jllgmp@cid.csic.es
URL: http://www.ibmb.csic.es/

C. Vicient
e-mail: cvsgmp@cid.csic.es

P. Puigdomènech
e-mail: pprgmp@cid.csic.es

M. Pla (✉)
Institut de Tecnologia Agroalimentària, Universitat de Girona,
Campus Montilivi, EPS-1,
17071 Girona, Spain
e-mail: maria.pla@udg.edu
URL: www.udg.edu

Introduction

The maize (*Zea mays*) event MON810 (trade name YieldGard®) is one of the main maize genetically modified

(GM) products approved for commercialization in the European Union. A total of nearly 60,000 ha of MON810 varieties were grown in 2006 (www.gmo-compass.org). MON810 was developed through a specific genetic modification for resistance to the European corn borer (ECB; *Ostrinia nubilalis*), a major insect pest of maize. The expression cassette used to produce MON810 incorporates the cauliflower mosaic virus (CaMV) 35S promoter, the *hsp70* maize intron, a synthetic *cryIA(b)* gene coding for a truncated delta-endotoxin derived from *Bacillus thuringiensis* that acts as a potent and highly specific insecticide, and the *Agrobacterium tumefaciens* NOS terminator. This cassette was integrated as a single copy but in previous work, we found a 3′-truncation partially affecting the coding sequence (truncation at position 2448 of the coding sequence), resulting in the complete deletion of the NOS terminator [1]. The 3′ maize genome flanking sequence has similarity to a gene putatively coding for the HECT E3 ubiquitin-ligase (BAC clone ZMMBBc0409B05) [2]. Expression analysis of *cryIA(b)* indicated a read-through transcription giving rise to different RNA variants that extended downstream from the 3′ truncation site [2]. The precise 3′ transcription termination site(s) of the transgene has not been characterized to date, nor has the possible synthesis of chimerical polyadenylated RNA species.

GM organisms (GMOs) are subject to legislation which establishes that each GMO has to be authorized before marketing to ensure consumer safety. Risk assessment for approval of new GM crops include field trials, animal nutrition and basic chemical composition studies (see, e.g., the guidelines of the Organization for Economic Co-operation and Development, www.oecd.org) but also molecular studies of the integrity of the transgene cassette, its copy number, insertion site, and transgene levels of expression and protein accumulation. In maize and other commercial crops, genetic transformation is accomplished in a single inbred line, with transgenes then introduced into different commercial varieties through breeding programs. Therefore, commercial varieties of maize MON10 contain, in principle, the same transgene in different genetic backgrounds. In March 2007, 47 MON810 maize varieties were inscribed in the EU common catalog of varieties of agricultural plant species, all coming from the same original transformation event (www.gmo-compass.org).

Here, we provide more information on commercial GMOs at the molecular level by analyzing the *cryIA(b)* transcription products in a number of MON810 commercial varieties, focusing on the 3′ ends of the transgene mRNA species. To avoid varietal differences caused by environmental and physiological factors, our experimental approach used plant leaves at the same growth stage grown under controlled conditions. Our results showed that the analyzed varieties had similar transgene mRNA levels. With a combination of Northern blot and RACE 3′-PCR we characterized the main 3′ termination sites and identified multiple polyadenylated *cryIA(b)* mRNA forms that extended downstream from the transgene truncation point. The patterns of 3′ transcription termination were similar among varieties. The possibility of adventitious 3′-end-processing sequences in the transgene 3′ flanking sequence is discussed.

Materials and methods

Plant material

Seeds from 28 commercial hybrid varieties of the MON810 maize event, listed in Table 1, were purchased in Spain. They are included in the Spanish official registry of plant varieties and were confirmed as MON810 by event specific real-time PCR [1]. Maize seeds of all 28 varieties were germinated and the seedlings grown in the greenhouse under controlled conditions (16 h/8 h photoperiod, 28 °C), harvested at the vegetative seven-leaf stage (V7) and immediately frozen in liquid nitrogen and then stored at −80 °C. Three biological replicates per variety were sampled, consisting of the two youngest fully opened leaves of each of three plantlets. Powdered certified reference material (CRM) for GM maize line MON810 (material containing <0.02% to 5% of genetically modified maize, ERM-BF-413 A, B, D, F) was purchased from the Institute for Reference Materials and Measurements (IRMM, Geel, Belgium), commercialized by Fluka (Fluka-Riedel, Geel, Belgium).

RNA isolation and cDNA synthesis

RNA was extracted from 100 mg of frozen leaves using TRIzol® reagent (Invitrogen Life Technologies, Carlsbad, CA, USA) according to the manufacturer's instructions. Total RNA was then purified with the Qiagen RNeasy MiniElute Cleanup Kit (Qiagen, Hilden, Germany) and treated with RNase-free DNase I (New England Biolabs Inc., Ipswich, MA) to eliminate any remaining DNA prior to reverse transcription (RT) and real-time PCR. The resulting RNA samples were spectrophotometrically quantified with a Nanodrop ND-1000 device (Nanodrop technologies, Wilmington, DE, USA).

First-strand cDNA was synthesized from 1 μg of total RNA using the Superscript III (Invitrogen Life Technologies) containing a mixture of random hexamers and oligodT in a total volume of 20 μL. Reverse transcription (RT) was at 25 °C for 10 min followed by 42 °C for 30 min in an Eppendorf Mastercycler S thermal cycler (Eppendorf AG, Hamburg, Germany). For each sample, cDNA was prepared at least in duplicate.

Table 1 MON810 varieties and transgene mRNA levels in leaves of each V7 plant

MON810 Maize variety	Breeder Company	Year[a]	ΔCt	
			mean	SD
Aristis Bt	Semences coop. de PAU	2003	6.20	0.19
Asturial Bt	Euralis Genetique	2006	6.21	0.04
Beles Sur	Holden Foundation Seed	2006	6.41	0.15
Campero Bt	Advanta Seeds B.V.	2004	6.30	0.08
Coventry Bt	Nickerson Sur	2003	6.87	0.08
Cuartal Bt	Euralis Genetique	2004	6.75	0.07
DKC513	R.a.g.t. Semences	1998	6.00	0.21
DKC5784-YG	Monsanto Technology LLC.	2005	6.15	0.25
DKC6041-YG	Monsanto Technology LLC.	2005	6.31	0.14
DKC6575	Monsanto Technology LLC.	2003	6.68	0.08
Foggia	Maisadour Semences	2005	6.35	0.29
Gambier Bt	Semences coop. de PAU	2004	6.68	0.10
Helen Bt	Advanta Seeds B.V.	2005	6.82	0.03
Jaral Bt	Semillas Fitó, S.A.	2004	6.66	0.11
Kaper-YG	Holden Foundation Seed	2007	6.61	0.32
Kardan	KWS SAAT AG	2006	6.60	0.06
Mas-58-YG	Maisadour Semences	2007	6.14	0.03
Mas-60-YG	Maisadour Semences	2007	6.49	0.12
Mas-73-YG	Maisadour Semences	2007	6.50	0.23
MGM155588	Maisadour Semences	2008	6.47	0.18
PR31N28	Pioneer Hi-Bred International	2006	6.44	0.21
PR32P76	Pioneer Hi-Bred International	2004	6.55	0.17
PR32R43	Pioneer Hi-Bred International	2005	6.46	0.13
PR33P67	Pioneer Hi-Bred International	2003	6.56	0.04
PR34N44	Pioneer Hi-Bred International	2005	6.75	0.22
Protect	Golden Harvest/J.C. Robins	2004	6.74	0.05
Sancia Bt	Holden Foundation Seed	2006	6.55	0.14
SF1112T	Semillas Fitó, S.A.	2005	6.39	0.23
OVERALL			6.49	0.22

Transgene expression values were obtained by real-time RT-PCR (targeting positions −70 to +36) of three biological replicates and are expressed as ΔCt after normalization with α-actin. Means and standard deviations (SD) are shown. Higher ΔCt values indicate lower mRNA contents, although no statistical differences were found among varieties

[a] Year of inscription in the European catalog of varieties of agricultural species

Quantitative real-time PCR

Quantitative real-time PCR assays were performed in a final volume of 20 μL containing 1× SYBR Premix Ex Taq (Takara Bio Inc., Shiga, Japan), 0.2 μM forward and reverse primers (Table 2), 0.4 μL ROX reference dye II (Takara Bio Inc.) and 2 μL of RT product as template. Real-time PCR reactions were carried out in an ABI7300 Real-Time PCR thermal cycler (Applied Biosystems, Foster City, CA, USA) with the following program: initial denaturation at 95 °C for 10 s; 40 cycles of 95 °C for 3 s and 60 °C for 35 s; and melting curve analysis at 60 to 95 °C with a heating rate of 0.5 °C/s. All samples were run in triplicate. Melting curve analyses consistently showed single peaks and the absence of artifacts.

We adapted a real-time PCR assay previously developed for specific detection of MON810 [1] to SYBR-Green chemistry. It targets the 3′ portion of *cryIA(b)* and extends 36 bp downstream of the *cryIA(b)* truncation site. Preliminary experiments with DNA extracted from MON810 CRM showed it was highly efficient (E=2.00±0.03 [3]) and linear (R^2=0.99±0.02). Taking into account the size of the maize haploid genome [4], the limit of quantification (95% probability) was established at around 100 target molecules, making the assay suitable for quantification purposes.

All samples were tested using six previously developed assays (D. Caparròs and M. Casado, manuscript in preparation) which targeted six maize housekeeping genes (α-actin, β-actin, β-tubulin, ubiquitin, cyclophilin and GAPDH), in order to select the most suitable gene as internal standard. This was performed with the geNORM v3.4 (Center for Medical Genetics, University Hospital Ghent, Belgium, http://medgen.ugent.be/genorm/) and NORMfinder (Molecular Diagnostic Laboratory, Aarhus University Hospital Skejby, Denmark; http://www.mdl.dk/

Table 2 Oligonucleotides used in this work

Oligonucleotide	Sequence	Position (relative to 3′ junction site)
MONF	5′-CAA GTG TGC CCA CCA CAG C-3′	−70 to −52
MONR	5′-GCA AGC AAA TTC GGA AAT GAA-3′	+16 to +36
Anchored-primer	5′-GCT AGG CCA CTG TGC C -3′	–
Anchored-oligodT	5′-GCT AGG CCA CTG TGC CT (T)17-3′	–
MON272F	5′-ATG CAA CCT CAC TGC GTT GA -3′	+272 to +292
MON486F	5′-CAC AAC AGG GAA ACG TTC CA-3′	+486 to +505
α-actin F	5′-TAC CCA ACT AAG CGC ATG CC-3′	–
α-actin R	5′-GCA TCT GAA TCA CGA AGC AGG-3′	–

publicationsnormfinder.htm) statistical algorithms. The α-actin gene was chosen to normalize *cryIA(b)* mRNA data (M value=0.475) for each sample, quantifying *cryIA(b)* mRNA levels relative to α-actin through two regression curves run in the same plate (acceptance criteria, slope between −3.1 and −3.6, R^2 above 0.98, http://gmo-crl.jrc.ec.europa.eu/doc/Min_Perf_Requir_Analyt_methods_131008.pdf). The absence of genomic DNA in the RNA samples was confirmed by the lack of real-time PCR amplification of RNA samples not subjected to reverse transcription.

Non-radioactive Northern blot analysis

Total RNAs (15 µg) were separated on a 1.0% (*w*/*v*) agarose/formaldehyde gel and transferred to Hybond N^+ membranes (GE Healthcare Life Sciences, Fairfield, CT) by capillary transfer with 20× saline sodium citrate (SSC, 3 M NaCl, 0.3 M Tri-sodium citrate). Specific riboprobes labeled with digoxigenin (DIG) were generated by in vitro transcription using the DIG RNA labeling kit (Roche) according to the manufacturer's protocol. After a 1-h prehybridization in DIG-easy buffer (Roche) and overnight hybridization in the same solution at 65 °C, the membrane was washed twice in 2× saline sodium phosphate EDTA (SSPE, 0.36 M sodium chloride, 20 mM sodium hydrogen phosphate, 2 mM EDTA, pH 7.4) at room temperature for 5 min and then washed in 0.1× SSPE at 65 °C for 15 min. The hybrids were detected by chemiluminescence using a LAS-3000 device (Fujifilm Life Sciences, Stamford, CT) with exposure times from 30 min to 1 h.

Statistical analysis

Normalized *cryIA(b)* mRNA copy number means of each variety were compared by a one-way analysis of variance (ANOVA) (α=0.05), and where statistically significant differences were found Tukey's HSD post-hoc test was run to assess individual pair differences.

Rapid amplification of cDNA ends

Rapid amplification of cDNA ends (RACE 3′-PCR) was performed according to Sambrook et al. [5]. Synthesis of the first cDNA strand was carried out as before but tail anchored-oligodT (Table 2, 10 µM) was used as the unique primer. One microliter of this cDNA was used as template for RACE 3′-PCR in a final volume of 50 µL containing 0.2 µM of each, MONF and anchor primers, 1× PCR Buffer II (Applied Biosystems), 25 mM $MgCl_2$, 0.8 mM dNTPs (200 nM each) and 0.25 U AmpliTaq Gold DNA polymerase (Applied Biosystems). Two microliters of the amplification products was subjected to a second PCR with the Anchored-oligodT primer combined with one of the following primers: MONF, MOM272F, or MON486F. RACE 3′-PCR products were analyzed by capillary micro-electrophoresis in an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) using DNA-1000 chips (capable of resolving 1.5 kbp). The electropherogram profiles were compared using the 2100 Expert software (Agilent technologies). Aliquots of the RACE 3′-PCR products obtained with MONF and MON486F primers were cloned into pCRII-TOPO (Invitrogen Life Technologies Corporation, Carlsbad, CA) and sequenced using the M13 primer and the BigDye Terminator v1.1 kit (Applied Biosystems).

Results

Transcription levels of *cryIA(b)* in leaves of MON810 commercial varieties

The levels of transgene mRNA accumulation were analyzed in leaves of 28 different MON810 commercial varieties by real-time PCR coupled to RT (real-time RT-PCR). In order to minimize the influence of environmental conditions, we used leaves of V7-stage plants grown in parallel under controlled conditions. Table 1 shows the *cryIA(b)* mRNA values in three biological replicates that were independently analyzed with primers MON810F/R (Table 2). The expres-

sion levels obtained showed a normal distribution (Shapiro–Wilk test, $p=0.585$) and all fell within 0.7 cycles (ΔC_T after normalization with α-actin mRNA expression, between 6.20 and 6.86): less than twofold differences, which are around the technical sensitivity [6]. Thus, all tested varieties had similar levels of transgene transcript in our experimental conditions, with an overall mean value of 115 *cryIA(b)* mRNA copies relative to *a*-actin with relative standard deviation (RSD) of 13.41%.

Northern-blot-based characterization of the *cryIA(b)* transcripts in different MON810 commercial varieties

Northern blot analysis was performed on total RNAs isolated from leaves of different maize MON810 varieties in order to determine the expression levels of the *cryIA(b)* transgene and the mRNA patterns. To minimize environmental effects we used leaves of V7 stage plants grown in parallel under controlled conditions. RNAs from leaves of five MON810 varieties (Beles Sur, DKC6041-YG, DKC 6575, PR33P67 and Protect) were analyzed, with two non-GM maize lines as control (PR33P66 and Tietar). Given that the transgene was 3′-flanked by a truncated maize gene in antisense orientation, hybridization was with riboprobes rather than conventional double-stranded probes, so that the transcribed strand could be determined. Single-stranded riboprobes were prepared in the sense and in the antisense orientation with respect to the *cryIA(b)* transgene (Fig. 1a). The antisense probe should hybridize to transcripts produced from the transgene CaMV 35S promoter; whereas sense probes should allow detecting potential transcripts corresponding to the truncated HECT E3 ubiquitin-ligase gene located 3′ of the transgene truncation point in the opposite direction. The probes corresponded to the 3′ flanking region (+1 to +597 downstream of the truncation site) and to the *cryIA(b)* coding region (positions −300 to −1 upstream of the truncation site, only sense probe).

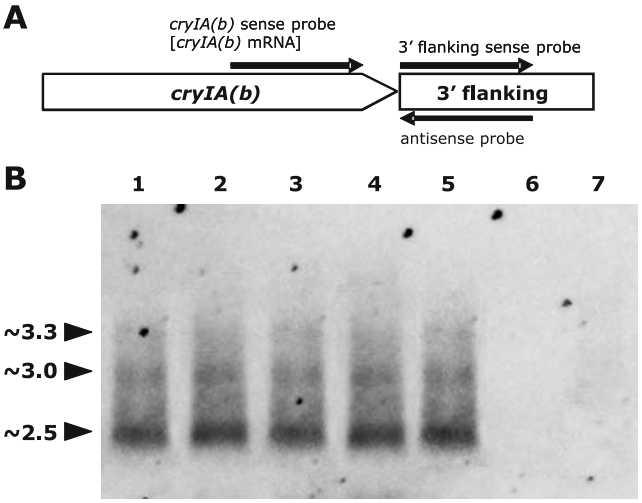


Fig. 1 Northern blot analysis of RNA extracted from leaves of V7 maize plants of five MON810 and two non-GM varieties. **a** Schematic representation of the transgene [*cryIA(b)*] and 3′ flanking regions; and the single-stranded probes; **b** Northern blot probed with *cryIA(b)* antisense probe. *Lane 1* DKC 6575, *lane 2* DKC6041-YG, *lane 3* Beles Sur, *lane 4* PR33P67, *lane 5* Asturial Bt, *lane 6* Tietar (DKC6575 near-isogenic line), *lane 7* PR33P66 (PR33P67 near-isogenic line)

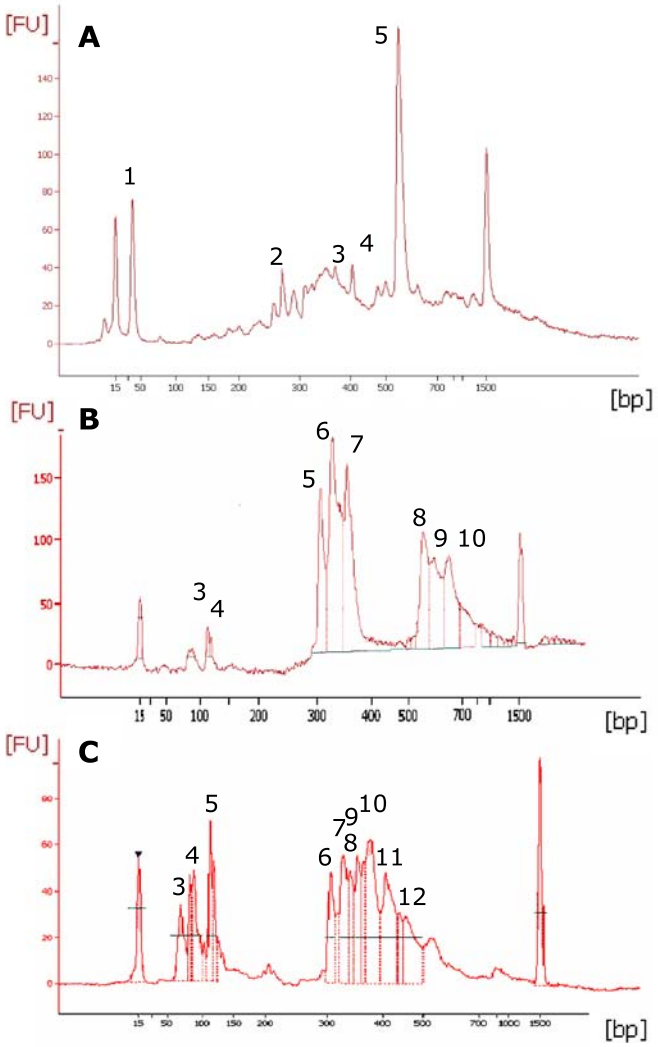
Northern blot analysis using the transgene antisense probe (to detect *cryIA(b)* transcripts) produced a prominent band of approximately 2,500 nt, with a smear and some other bands of higher MW (around 3,000 and 3,300 nt, Fig. 1b). All analyzed MON810 varieties had the same hybridization pattern with this probe, while there was no hybridization with non-GM RNA samples. The ~2,500 nt band may correspond to a transcript which ends shortly downstream the *cryIA(b)* truncation. The high MW smear suggests the presence of a range of different-sized transgene mRNA due to variability in the end point of the transgene mRNA.

In an additional experiment, hybridization with a 3′ flanking region sense probe gave no detectable expression of the HECT E3 ubiquitin-ligase gene, nor was there transcription of the transgene non-coding strand using the *cryIA(b)* sense probe, in any of the tissues and varieties analyzed (not shown). This was especially relevant for MON810 varieties in which this sequence is adjacent to a truncated gene in the opposite orientation.

Characterization of *cryIA(b)* 3′-end-processing sites in MON810

Our Northern blot analyses indicated certain complexity at the 3′-end of transgene mRNAs. We used a RACE 3′-PCR strategy coupled to capillary micro-electrophoresis to map 3′ termination sites of *cryIA(b)* polyadenylated transcripts in MON810 plants. The experimental design included oligo(dT)-primed reverse transcription and PCR amplification with a forward primer at position −79 from the *cryIA(b)* truncation site (MONF). This ensures only transgene transcripts are detected. The first PCR products were re-amplified with oligo(dT) and one of the following primers: MONF, MON272F or MON486F (positions −70, +272 and +486 bp, respectively, Table 2) to facilitate detection and size determination of longer transcripts. The amplification products were resolved by capillary micro-electrophoresis and the size of all peaks with intensities above the threshold was determined.

As illustrated in Fig. 2, the results obtained with all three forward primers were fully compatible and indicated multiple transgene transcript ends. They showed that transgene transcripts either terminated around the truncation site or extended downstream of it: these transcripts were different sizes in the 200 to around 1000 nt range (from the



3' RACE PCR peak#			
MONF	MON272F	MON486F	cDNA size*
1	4		
2			210
3			344
	3		390
	4		396
4			412
5			484
		3	574
	5		583
		4	599
	6		601
		5	604
	7		626
		6	792
		7	818
		8	826
	8		828
		9	839
		10	864
	9		873
		11	894
	10		927
		12	942

* length from the truncation site (nt). Note that a bias of around 10 nt can be expected (a relative application is used to calculate the size of the fragments).

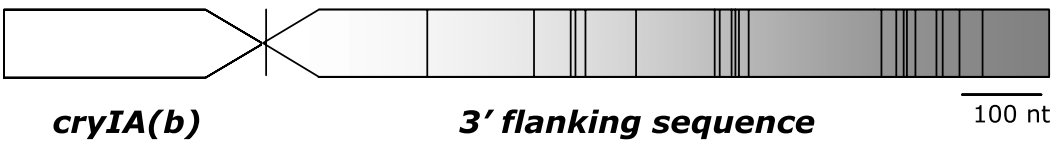


Fig. 2 Representative examples of electropherograms of RACE 3′-PCR products obtained with mRNA extracted from V7 leaves of the MON810 variety DCK513. The anchored-oligo(dT) reverse primer was used in combination with MONF (**a**), MON272F (**b**) or MON486F (**c**) forward primers. All analyses were performed at least in triplicate. Peaks were automatically numbered in each electropherogram and do not correspond to the same sequence. **d** Mean lengths of the electropherogram peaks were used to calculate the length of the mRNA molecules from the truncation point. Thus, peaks with similar sizes (calculated from the truncation) are likely to correspond to similar mRNA molecules. **e** Schematic representation of the length of *cryIA(b)* mRNA molecules analyzed by RACE 3′-PCR. *Vertical bars* represent the 3′ termination sites of *cryIA(b)* mapped by RACE 3′-PCR coupled to capillary micro-electrophoresis. *White and gray* shades in **d** and **e** correspond to regions closed the truncation site and farther downstream, respectively

truncation) and many extended around 550–625 or 800–1000 nt 3′ of *cryIA(b)*. This is in agreement with the band patterns observed in the Northern blot analyses.

A total of 8 independent RACE 3′-PCR products were cloned and sequenced (GenBank under Acc. # FN552460 to FN552467). Comparison of their sizes and those of the different peaks in Fig. 2 allows deducing that they probably corresponded to peaks 3 and/or 4 (MONF primer); and peaks 3, 4, 8, and/or 9 (MON486F primer). The obtained sequences corresponded to the transgene and 3′ flanking sequence and all perfectly matched the genomic sequence with no insertions or deletions and, as expected, they all had poly(A) tails. This established the synthesis of polyadenylated transgene transcripts with 3′

Table 3 Homology-predicted 3′-end-processing elements in the 3′ flanking region (bp +1 to +1000 relative to the truncation site) of the *cryIA(b)* transgene in MON810 maize

Sequence	3′-end-processing element	Element position[a]	cDNA clones[b]
GATGTA	FUE[c]	128	
AATAAT	NUE[d]	361	FN552460 (5 nt)
			FN552465 (10 nt)
			FN552466 (5 nt)
			FN552467 (9 nt)
AATAGA	NUE	572	FN552461 (21 nt)
			FN552462 (13 nt)
TTTGTA	FUE	599	
AATGAA	NUE	716	
AAAATT	NUE	818	FN552463 (16 nt)
			FN552464 (11 nt)

Only putative NUE and FUE sequences are displayed. Indicated are the GenBank Acc. No. of sequenced cDNA clones derived from RACE 3′-PCR that had poly(A) tracts placed downstream each NUE. The distance between the poly(A) tract of each cDNA clone and the most proximal NUE is shown in brackets

[a] nt from the truncation point

[b] GenBank acc # of sequenced cDNA clones and position of the last nt relative to the closest NUE (in brackets)

[c] Far upstream element

[d] Near upstream element

termini close to the truncation and up to around 1 kb downstream of this point.

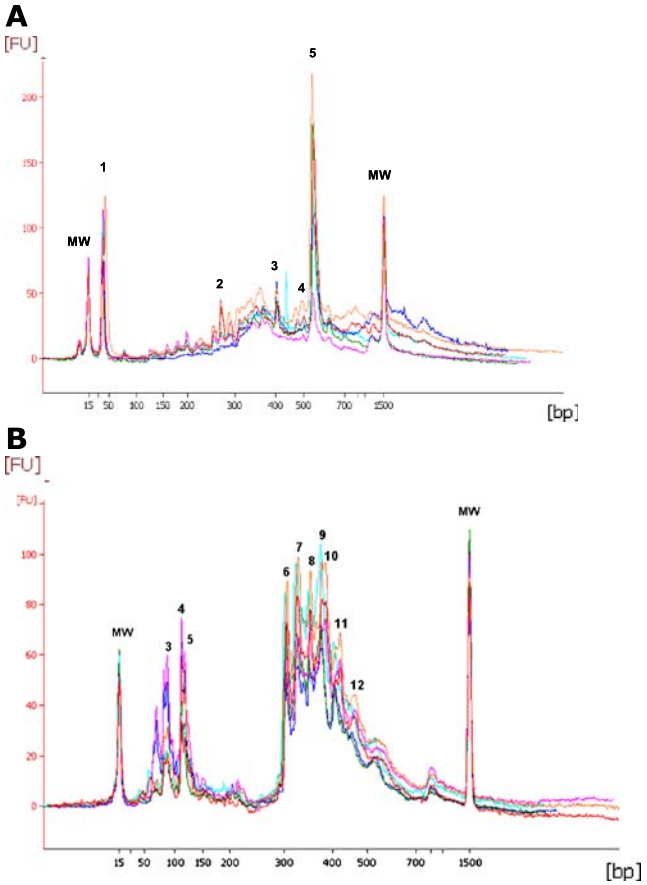
We searched for sequence homologies with 3′-mRNA-end-processing elements within the proximal 3′ flanking region of the MON810 *cryIA(b)* insert (nt +1 to +1314, i.e. up to the 5′ ending of HECT E3 ubiquitin-ligase exon 8 sequence), a sequence with 56.8% A/T residues. A number of putative 3′-end-processing elements were found (Table 3): two far-upstream elements (FUE), four near-upstream elements (NUE) and various poly(T) tracts, whose position relative to the polyadenylation sites of *cryIA(b)* mRNA molecules analyzed in this work makes them potential candidates for driving termination of the transgene transcripts [7, 8]. One of the sequenced mRNAs had a conserved cleavage site (CS, 5 consecutive T residues). Other homologous sequences were identified farther downstream (data not shown).

Transgene 3′-end-processing profiles in MON810 commercial varieties

To check to what extent *cryIA(b)* 3′-end-processing could generate variability among commercial varieties of the MON810 event, we compared the RACE 3′-PCR products obtained with mRNA extracted from V7 stage leaves of 6 different MON810 commercial varieties, i.e., Aristis Bt, Jaral Bt, PR33P67, DKC6575, DKC513, and DKC6041-YG. We focused on RACE 3′-PCRs with primers MONF and MON486F, which gave most information (see Fig. 2). All six varieties analyzed produced fully overlapping profiles (Fig. 3), indicating that they had basically the same mRNA species. The peaks were consistently detected in leaves from all varieties tested with only three minor peaks (No 2, 3, and 4, MONF) below the threshold in some varieties. The 15 identified peaks surpassing the established threshold (Fig. 3) were different lengths (ANOVA p value, 0.000 with a mean square error value placed below 0.005%) and all showed basically the same sizes in all varieties: the small differences observed could be easily explained by differences in the length of the poly(A) tails.

Discussion

Over the last few years, there has been growing interest in the detailed study of transgene loci in GM plants, in relation to stability and expression [9, 10]. Commercial GM crops have also been the object of investigation, though they are authorized following strict regulations to ensure they have no negative health or environmental effects. We have previously shown that the maize MON810 transgene is truncated at position 2448 of the *cryIA(b)* coding sequence and lacks a terminator [1], and Rosati et al. [2] have shown that the transgene transcription proceeded 3′ past the truncation point. The aim of the present study was to investigate the main transgene transcripts in MON810. Since *cryIA(b)* is 3′ flanked by a sequence homologous to maize HECT E3 ubiquitin-ligase gene exon 8, in reverse



C

3' RACE-PCR peak No.		cDNA size[1]	
MONF	MON486F	Mean (nt)	RSD (%)
1		3.6	3.01
2		210.5	0.94
3		343.8	0.69
4		412.5	0.91
5		484.2	3.78
	3	574.3	1.65
	4	598.7	1.40
	5	604.3	1.36
	6	791.7	2.19
	7	817.8	2.02
	8	825.7	3.16
	9	839.3	2.70
	10	863.7	1.63
	11	894.0	4.34
	12	941.6	5.18

[1] length from the truncation site. Note that a bias of around 10 nt can be expected (a relative application is used to calculate the size of the fragments)

◀**Fig. 3** Comparison of profiles obtained in capillary micro-electrophoresis-RACE 3′-PCR experiments performed with primers MONF/Anchored-oligodT (**a**) or MON486F/Anchored-oligodT (**b**) and mRNA from six MON810 varieties: Aristis Bt (*green*), DKC513 (*light blue*), DKC6041-YC (*orange*), DKC6575 (*red*), Jaral Bt (*blue*) and PR33P67 (*pink*). **a** and **b**: superposition of electropherograms; **c** calculated size of the identified peaks (bp); and calculated size of transgene cDNA molecules from the *cryIA(b)* truncation site [mean value (nt) and relative standard deviation, RSD (%)]

orientation [2], our strategy was initially based on Northern blot RNA analysis and hybridization with single-stranded riboprobes, to identify transcripts corresponding to the coding or non-coding strand of the target gene. The *cryIA (b)* transgene gave rise to a number of transcripts of different sizes that either terminated immediately downstream of the truncation site or extended 3′ past this site up to around 1 kb. Longer molecules had a pattern of a few main sizes and multiple intermediate lengths. This confirms the transcription read-through 3′ past the *cryIA (b)* truncation point; and showed that a transgene lacking a terminator sequence can produce a fairly complex pattern of transcripts.

The most extensively cultured GM event, Roundup Ready soybean, contains a fragment of the *epsps* gene downstream from the T-NOS terminator [11] that has been shown to be partly transcribed [12]. Sense and antisense transcripts in plants can also be generated by transcriptional readout from retrotransposon long terminal repeats (LTR) [13–15], which constitute more than 70% of the maize genome [16]. It should be noted that the sequence generated at position +7 to +9 downstream from the MON810 transgene truncation site corresponds to an amber termination codon (GenBank Acc. # AY32643, AF490398). Translation termination at this site is in agreement with immunological analyses showing the accumulation of a CryIA(b) protein of the expected length (63 kD) in MON810 plants (http://www.agbios.com/dbase.php?action=Submit&evidcode=MON810). Thus, there is no evidence of the existence of other translated products. We could not detect RNA corresponding to the *cryIA(b)* non-coding strand or to the HECT E3 ubiquitin-ligase exon 8 sequence in leaves of V7 MON810 or non-GM maize varieties, indicating the absence of other chimerical RNA molecules. As most eukaryotes, plants contain multiple HECT E3 genes [reviews in 17, 18] with seven members identified in *Arabidopsis* [19].

Messenger RNA 3′ end formation (pre-mRNA processing and polyadenylation) generates mature mRNAs in eukaryotes. Poly(A) tracts are fundamental determinants of the function of a eukaryotic mRNA; they are linked to processes that determine the translatability and lifetime of mRNAs and participate in gene expression modulation [review in 20, 21]. With a RACE 3′-PCR approach we showed the formation of polyadenylated transgene mRNAs which either terminated in the region of the truncation site or extended 3′ downstream it to reach a variety of positions up to around 1 kb downstream. A representation of transcripts were sequenced: they were 100% homologous to the described genomic *cryIA(b)* and 3′ flanking sequence

and had no deletions or insertions. These results further add to the recently published existence of transgene mRNA transcripts in MON810 extending downstream from the truncation point [2]. They show that most *cryIA(b)* mRNA sequences extend less than 1 kb from the transgene (Northern blot experiments) and indicate they can be polyadenylated (RACE 3′-PCRs).

In plants, the polyadenylation control signals contained in the 3′ untranslated regions are more variable than in mammals. The 3′-end-processing signals in plants consist of FUE(s), one or more A-rich regions known as NUE, and U-rich regions located close to a CS [7, 8]. Bioinformatics studies identified new CS downstream elements (DE) in *Arabidopsis* [22], and recently, Dong et al. [23] proposed a general model of the conservative elements of the 3′-end processing in rice based on cDNA sequence data. Genome-level analyses confirmed the presence of these elements in most transcripts in *Arabidopsis* [24] and rice [25], although their sequences were somewhat variable. The *cryIA(b)* 3′ flanking sequence had several putative 3′-end processing elements that could potentially drive termination of the transgene transcripts. As an example, NUE at nt +361 is 5 to 10 nt upstream of the poly (A) of 4 sequenced transcripts, which are 6 to 1 nt upstream of a putative DE. Moreover, the +818 nt NUE is 13 and 21 nt upstream of two sequenced mRNAs and their corresponding CSs are 6 to 1 nt upstream of a putative DE. Plant genes often have multiple polyadenylation sites and multiple corresponding NUEs, with a single FUE controlling various NUEs [8]. It could be speculated that this was the case of the transgene in MON810, with FUEs at +128 and +599 being associated to NUEs +361 and +572; and +818, respectively. It should be noted that the *cryIA(b)* 3′-flanking sequences are the product of a genomic rearrangement and that the HECT E3 ubiquitin-ligase exon 8 sequence was most probably not intended to act as terminator of a gene in the opposite sense. However, our results suggested that all four NUEs identified within the 840 bp 3′ flanking the *cryIA(b)* transgene in MON810 drive termination and processing of transgene mRNA molecules.

The MON810 event is widely grown and commercialized as different varieties, which are genetically diverse but they all harbor the same truncated transgene in the same chromosomal context. Interestingly, all six analyzed MON810 commercial varieties showed the same *cryIA(b)* transcript pattern. In addition, transgene RNA levels were examined by real-time RT-PCR in leaves of 28 GM varieties. In similar tissues, environmental conditions and developmental stage, *cryIA(b)* was similarly expressed in a wide representation of MON810 varieties cultivated and/or commercially available in Europe between 1997 and 2007, comprising a broad illustration of the existing genetic and phenotypic diversity, seed companies, breeding procedures and year of introduction onto the market. Although some differences have been reported for other plant tissues, geographical areas and varieties [26], these results are in agreement with our previous results showing similar *cryIA(b)* mRNA levels in 5 MON810 commercial varieties cultured in vitro [27]. Moreover, *cryIA(b)* 3′ termini profiles were the same in six different MON810 commercial varieties (Aristis Bt, Jaral Bt, PR33P67, DKC6575, DKC513, and DKC6041-YG). Thus, commercialized MON810 varieties appear to be stable in terms of transgene expression.

Acknowledgments This work was financially supported by the EC FP6 project “Co-Extra” and by the Spanish MEC project with ref. AGL2007-65903/AGR.

References

1. Hernández M, Pla M, Esteve T, Prat S, Puigdomènech P, Ferrando A (2003) Transgenic Res 12(2):179–189
2. Rosati A, Bogani P, Santarlasci A, Buiatti M (2008) Plant Mol Biol 67:271–281
3. Rasmussen R (2001) In: Meuer S, Wittwer C, Nakagawara K (eds) Rapid Cycle real-time PCR, methods and applications. Springer, Berlin,G
4. Arumuganathan K, Earle ED (1991) Plant Mol Biol Rep 9:208–218
5. Sambrook J, MacCallum P, Russell D (2000) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press
6. Bubner B, Gase K, Baldwin IT (2004) BMC Biotechnol 4:14–20
7. Rothnie HM (1996) 32(1-2):43-61
8. Li Q, Hunt G (1997) Plant Physiol 115:321–325
9. Kohli A, Twyman R, Abranches R, Wegel E, Stoger E, Christou P (2003) Plant Mol Biol 52:247–258
10. Zhang J, Cai L, Cheng J, Mao H, Fan X, Meng Z, Chan KM, Zhang H, Qi J, Ji L, Hong Y (2008) Transgenic Res 17(2):293–306
11. Windels P, Taverniers I, Depicker A, Van Bockstaele E, De Loose M (2001) Eur Food Res Technol 213:107–111
12. Rang A, Linke B, Jansen B (2005) Eur Food Res Technol 220:438–443
13. Kashkush K, Feldman M, Levy AA (2003) Nat Genet 33:102–106
14. Kashkush K, Khasdan V (2007) Genetics 177:1975–1985
15. Hernández-Pinzón I, de Jesús E, Santiago N, Casacuberta JM (2009) J Mol Evol in press. doi:10.1007/s00239-009-9204-y
16. Vitte C, Bennetzen JL (2006) Proc Natl Acad Sci USA 103 (47):17638–17643
17. Hellmann H, Estelle M (2002) Science 297(5582):793–797
18. Kee Y, Huibregtse JM (2007) Biochem Biophys Res Commun 354(2):329–333
19. Downes BP, Stupar RM, Gingerich DJ, Vierstra RD (2003) The Plant Journal 35(6):729–742
20. Hunt AG (2008) Curr Top Microbiol Immunol 326:151–177
21. Edmonds M (2002) Prog Nucleic Acid Res Mol Biol 71:285–389
22. Loke JC, Stahlberg EA, Strenski DG, Haas BJ, Wood PC, Li QQ (2005) Plant Physiol 138(3):1457–1468
23. Dong H, Deng Y, Chen J, Wang S, Peng S, Dai C, Fang Y, Shao J, Lou Y, Li D (2007) Gene 389:107–113
24. Ji G, Zheng J, Shen Y, Wu X, Jiang R, Lin Y, Loke J, Davis K, Reese G, Li Q (2007) BMC Bioinformatics 8:43
25. Shen Y, Ji G, Haas BJ, Wu X, Zheng J, Reese GJ, Li QQ (2008) Nucleic Acids Res 36(9):3450–3161
26. Nguyen HT, Jehle JA (2007) J Plant Dis Protect 114:820–887
27. Coll A, Nadal A, Palaudelmàs M, Messeguer J, Melé E, Puigdomènech P, Pla M (2008) Plant Mol Biol 68:105–117