Plant Science 180 (2011) 313–322

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Analysis of SCAR marker nucleotide sequences in maize (*Zea mays* L.) somaclones

E.S. Osipova[a,*], E.A. Lysenko[a], A.V. Troitsky[b,1], Yu.I. Dolgikh[a], Z.B. Shamina[a], S.A. Gostimskii[c,2]

[a] *K.A. Timiryazev Institute of Plant Physiology, Russian Academy of Sciences, Botanicheskaya 35, Moscow 127276, Russia*
[b] *A.N. Belozersky Institute of Physicochemical Biology, Moscow State University, Moscow 119899, Russia*
[c] *Department of Genetics, Moscow State University, Moscow 119899, Russia*

ARTICLE INFO

Article history:
Received 10 July 2010
Received in revised form 4 October 2010
Accepted 5 October 2010
Available online 16 October 2010

Keywords:
Somaclonal variation
SCAR
Hot spots
Transposon
Zea mays L.

ABSTRACT

SCAR (sequence characterized amplified region) markers allow the reliable identification of unique somaclonal variations. Six SCAR markers were developed previously and were thought to be exclusively characteristic of eight maize somaclones. However, we detected two of these markers in maize lines and a cultivar unrelated to the progenitor line of the somaclones. Therefore, we sequenced these markers and performed bioinformatic searches to understand the molecular events that may underlie the variability observed in the somaclones. All changes were found in noncoding sequences and were induced by different molecular events, such as the insertion of long terminal repeat (LTR) transposon(s), precise miniature inverted repeat transposable element (MITE) excision, microdeletion, recombination, and a change in the pool of mitochondrial DNA. For example, the SCAR marker QR is represented by the two variants QR-A and QR-2. The sequences of the two variants were similar, except for a 457-bp fragment found only in QR-A; this region was denoted as Q. Region Q was flanked by the direct 3-bp repeat 5′-TAA-3′ (target site duplication; TSD) and the inverted 14-bp repeat 5′-GGGCCTGTTTGGAA-3′ (terminal inverted repeats; TIRs). These features confer the Q region with similarity to the nonautonomic *Tourist*-like MITE. In two groups of independently produced somaclones, the same features (morphological, molecular) were variable, which confirms the theory of 'hot spots' occurring in the genome. The distribution of one of the SCAR markers was confirmed using Southern blot hybridization. The presence of the same molecular markers in the somaclones and in different non-somaclonal maize variants suggests that in some cases, the same mechanisms determine both *in vitro* and *in vivo* variability and that cell culture enhances the rate of heritable genomic changes that naturally occur in living organisms.

1. Introduction

The phenomenon of plant somaclonal variation manifests itself as an increase in the heritable variability of cultured cells and tissues in regenerant progeny. Regenerated plants that differ from original plants are called somaclonal variants [1]. The investigation of somaclonal variation could be of interest in basic research as a model of gene evolution and divergence in the population and in breeding programs as a source of genetic diversity.

Previously, we obtained two groups of *Zea mays* L. somaclones from the A188 inbred line and characterized them phenotypically [2,3]. Regenerants of the first group (R11, R14, R27, and R54) were produced from calli after two months of culture. They differed from the original plants in some quantitative traits, such as plant height, the number of tassel branches, and the number of kernel rows in the ear. Somaclones of the second group (R105, R106, R107, and R119) were produced from calli after eight months of culture. They differed from the original A188 line in quantitative characteristics and in flowering time, kernel coloration, and the capacity to form an embryogenic callus *in vitro*. The stable inheritance of the somaclonal variations we discovered was monitored over the next two to four generations. A genetic analysis showed the appearance of altered features that were inherited independently, implying that they resulted from different mutations. The inheritance of one trait (purple kernel coloration) was unstable; after crossing with the original line, it was not detected in either the first or second generation, but it revealed itself again in the third generation [2,3].

Chromosome analysis revealed no differences in either the number of chromosomes or the distribution of heterochromatin segments between the somaclones produced and the original line [4].

Isoenzyme analysis of the tested somaclones revealed numerous inheritable qualitative and quantitative changes [5,3]. Within the

Abbreviations: ISSR, inter simple sequence repeat; MITE, miniature inverted repeat transposable element; RAPD, random amplified polymorphic DNA; SCAR, sequence characterized amplified region.

* Corresponding author. Fax: +7 495 977 80 18.
E-mail addresses: oes@bk.ru (E.S. Osipova), tav@genebee.msu.su (A.V. Troitsky), sgost@list.ru (S.A. Gostimskii).
[1] Fax: +7 495 939 31 81.
[2] Fax: +7 495 939 43 09.

0168-9452/$ – see front matter.
doi:10.1016/j.plantsci.2010.10.004

sets of esterase and peroxidase isoforms, changes in tissue specificity were the most common. For example, in the progenitor line A188, the peroxidase isoform Px12 was detected in the roots only, whereas in all somaclones, this isoform was present in both roots and leaves. The isoenzyme spectra of both independently propagated groups of somaclones were similar, and changes in the same traits distinguished them from A188. An analysis of the regenerant progeny showed that most of the detected variations were inherited by all generations obtained by somaclone selfing. After crossing with the original A188 line, the peroxidase root-specific isoform Px12 was found in the leaves of all plants in the F_1 and F_2 generations without any segregation [5,3].

At present, one of the most widely used and rapid approaches for detecting DNA changes that occur during tissue culture is polymerase chain reaction (PCR)-based DNA markers. For example, polymorphic PCR fragments obtained using the technique of random amplified polymorphic DNA (RAPD) [6] or inter simple sequence repeat (ISSR) [7] could be sequenced and used for the design of sequence characterized amplified region (SCAR) markers [8], which would uniquely and reliably differentiate somaclonal variations from the original individual.

We previously tested the original A188 line and the somaclones obtained using 38 RAPD and 10 ISSR primers. No A188 plants showed variation in the RAPD and ISSR spectra for any of the primers used [9,10]. However, the PCR spectra obtained from the somaclones demonstrated some variations, i.e., 22 RAPD primers and 6 ISSR primers differentiated at least one somaclonal variant from the progenitor line [9,10]. These PCR spectra could be divided into the three following groups: those that detected differences between the individual somaclones and A188 (12 RAPD and 3 ISSR spectra), those that amplified fragments characteristic of the somaclonal groups (nine RAPD and two ISSR spectra), and those that discriminated all regenerants from A188 (one RAPD and one ISSR spectra). The somaclonal variation frequency increased as the culture time increased, and it corresponded with increasing morphological changes [9,10]. Six SCAR markers were developed based on several RAPD and ISSR fragments (M-10, Q-20, OPC-09, NO-15, QR-2, and Leb). Five of them confirmed the polymorphism found by the originating RAPD and ISSR fragments. The source Leb RAPD fragment was characteristic only of the R105 somaclone, whereas the developed Leb SCAR marker was amplified in all regenerants and in the original A188 line. The inheritance of the six SCAR markers was verified in the selfing progeny of each somaclone in the R_1–R_4 generations and in the R27 and R105 hybrids, with A188 as the parental line in the F_1 and F_2 generations. As a rule, somaclonal changes of the original genotype occurred in one allele, i.e., regenerants of the R_0 generation were heterozygous. Similar results were obtained for the OPC-09, NO-15, and Leb markers: in the R_1–R_4 generations, segregation based on the presence/absence of the DNA fragment occurred close to a Mendelian ratio. In some cases (Leb: $R_2$14, $R_2$54, and $R_4$27), the marker was lost. Thus, SCAR Leb was not unique to R105 as we thought, but this result demonstrates polymorphism in the regenerant progeny. Segregation of the Q-20, M-10, and QR-2 markers was not detected in R_1 or subsequent selfing generations [11].

In the present work, we searched the genomes of four maize variants unrelated to A188 and two pea cultivars for the presence of somaclonal SCAR markers reported in previous studies. Some of the SCAR markers that were supposedly exclusive to the somaclones (M-10 and QR-2) were found in unrelated maize lines and a cultivar. Therefore, we sequenced these markers and performed bioinformatic searches to understand the molecular events underlying these genetic changes. To verify the conclusions we obtained from PCR-based methods, we analyzed the distribution of one SCAR marker using a different molecular method. We chose a group-specific marker (Q-20) and analyzed the distribution of the sequence in the initial line A188 and in both groups of somaclones using Southern blot hybridization. The data obtained from the DNA hybridization confirmed the results of the PCR-based methods.

2. Materials and methods

2.1. Original material

We compared the original maize (*Zea mays* L.) inbred line A188 (USA) and two groups of somaclones produced from different calli after two (R11, R14, R27, and R54) or eight (R105, R106, R107, and R119) months of culture. The morphological variations in the somaclones are given in Supplementary Table 1. Eight plants of the original A188 line were used in the molecular analysis. The somaclones were represented by generations R_1 (4–20 plants), R_2, R_3, and R_4 (about 20 plants in each case). We also analyzed hybrids of somaclones R27 and R105 (which differed the most from the original line in each group) with A188 as the male parental line, i.e., 19 F_1 (R27 × A188) plants, 4 F_1 (R105 × A188) plants, and 35 F_2 plants for each hybrid. In addition, we analyzed 10 plants of the four following maize variants that are unrelated to A188: inbred lines of Mangelsdorf Tester (USA) and Gk26 (Russia), cultivar 'Pannonia' (Hungary), a hybrid obtained by crossing the inbred Chinese31 to the Latin American race Cateto Sulino Grosso (Chi31 × Cateto SG), and two pea cultivars ('Kapital' and 'Rannii Zelenyi') (10 plants each).

The seeds were sterilized with the commercial bleaching reagent Belizna (sodium hypochlorite, 7–9% active chlorine) for 20 min, washed with sterile distilled water three times, and germinated on agar-solidified medium containing half the concentration of Murashige and Skoog salts [12]. Two-week-old seedlings were used for the PCR analysis.

2.2. DNA isolation and DNA amplification

DNA isolation and amplification was performed as previously reported by Osipova et al. [9].

2.3. Sequencing

DNA sequencing reactions were performed using ABI PRISM® BigDye™ Terminator v. 3.1 with subsequent analysis of the reaction products using the DNA ABI PRISM 3100-Avant automatic sequencer. The sequences were deposited in the GenBank database under accession numbers FJ748858 (M-10), FJ748859 (QR-A), FJ748860 (QR-2), FJ748861 (Q-20), FJ748862 (OPC-09), FJ748863 (Leb), and FJ748864 (NO-15).

2.4. Design of SCAR primers

Primers for SCAR analysis were designed using Oligo 4.0 software [13] and synthesized by Lytech (Russia).

2.5. Southern blot hybridization

Total DNA (4 (g) from each sample was digested with *Eco*RI, *Eco*RV, and *Xba*I restriction endonucleases (Fermentas, Lithuania) according to the manufacturer's recommendations but with an excess of enzyme activity (15 units). The digested samples were electrophoresed on a 0.7% gel containing ethidium bromide in TAE buffer and transferred to a nylon Hybond-N^+ membrane (Amersham, UK) as recommended by the manufacturer. A specific fragment amplified using the SCAR primers was used as a probe. Prehybridization, hybridization, and washing were performed according to the Amersham protocols, and the hybridization step

was performed at 65 °C. Autoradiography was performed on Kodak X-Omat AR films at −70 °C overnight.

2.6. Bioinformatic analysis of SCAR fragments

The search for similar DNA sequences was performed using the BLAST program in GenBank (http://www.ncbi.nlm.nih.gov), and MaizeGDB (http://www.maizegdb.org), the Institute for Genomic Research (TIGR) (http://maize.jcvi.org), and plantGDB (http://www.plantgdb.org) databases. The sequences were searched for known repeats with RepeatMasker (http://www.repeatmasker.org), and the TIGR Maize Repeat database (http://maize.jcvi.org/repeat_db.shtml). The secondary structures of transposons were predicted by mfold (http://mfold.bioinfo.rpi.edu) [14], and the alignment of sequences was performed using ClustalW [15].

3. Results

The somaclonal variations revealed with the six SCAR markers are shown in Fig. 1. The QR-2 fragment was amplified from DNA of all regenerants, whereas in the progenitor A188 line, the SCAR QR-2 primers originated from a larger amplicon (1428 bp) named QR-A (Fig. 1a). Fragment M-10 was common for all somaclones of the first group (Fig. 1b), and Q-20 characterized the second group of somaclones (Fig. 1c). SCARs OPC-09 (Fig. 1d), NO-15 (Fig. 1e), and Leb (Fig. 1f) were detected in specific somaclones. The inheritance of the SCAR markers was verified in the selfing progeny of each somaclone in the R_1–R_4 generations. In addition, hybrids of somaclones R27 and R105 (which differed the most from the original line in each group) with A188 as the male parental line were analyzed in the F_1 and F_2 generations. The markers were stably inherited, and all information concerning the generations tested is shown in Supplementary Table 2.

To evaluate whether the developed markers were unique for these somaclones, we analyzed four maize variants unrelated to A188. Additionally, two pea cultivars were used as an extra control for specificity. As expected, none of the primers amplified fragments from pea DNA (data not shown). Similar to A188, SCARs Q-20 and NO-15 were not detected in the four other maize variants, while OPC-09 and Leb were present in all of the genotypes tested. To our surprise, SCARs M-10 and QR-2, which were supposedly specific for A188-originated somaclones, were found in three of the four tested maize variants that were unrelated to A188 (Table 1).

To investigate the background of these genetic changes, we determined the nucleotide sequences of the SCAR markers. Database searches yielded matches to some fragments and sequences in the maize genome (Table 2). A comparison of the sequences revealed that the nature of some mutations changed the profile of PCR products in the somaclones. Thus, the M-10 fragment, which marked the first group of somaclones, was composed of two domains: the first domain (1–235 bp) exhibited high similarity to the flanking retrotransposons, and the second domain (236–806 bp) was similar to the noncoding region between *rp3*-

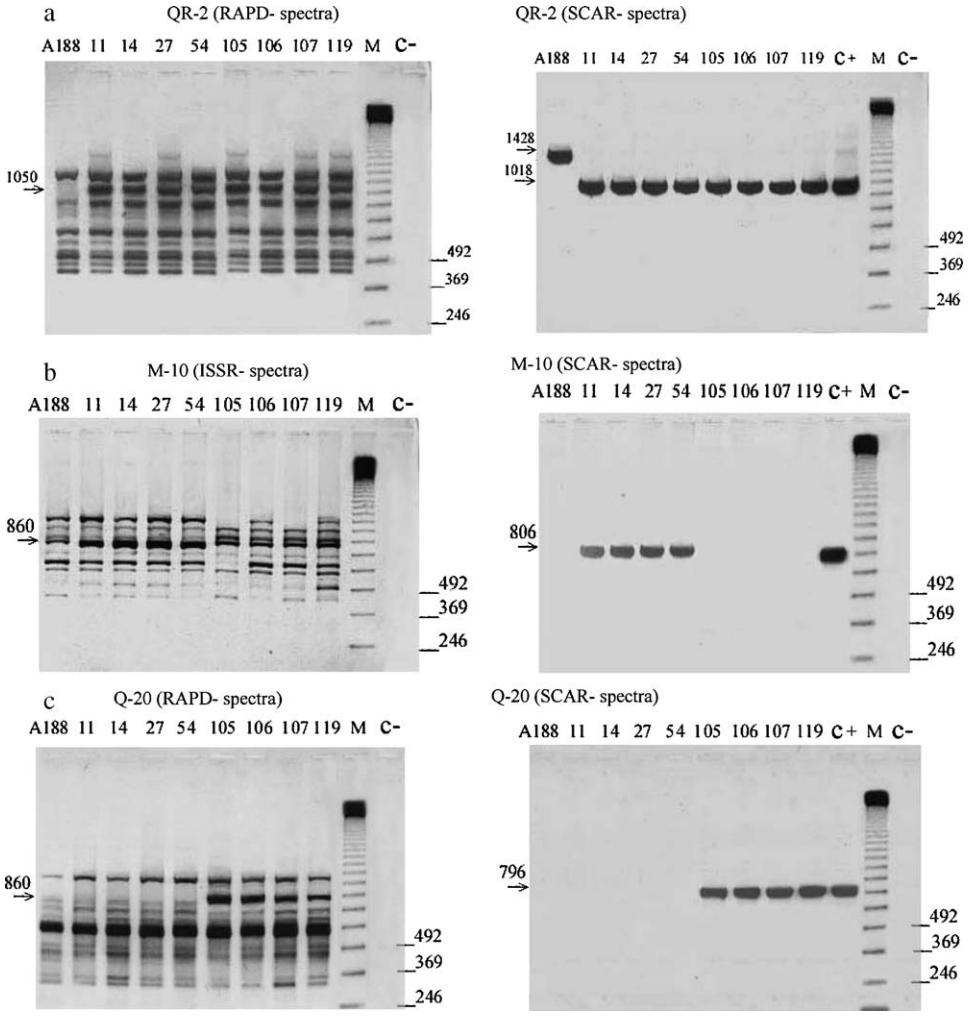


Fig. 1. Correspondence between RAPD-, ISSR-, and SCAR-spectra.

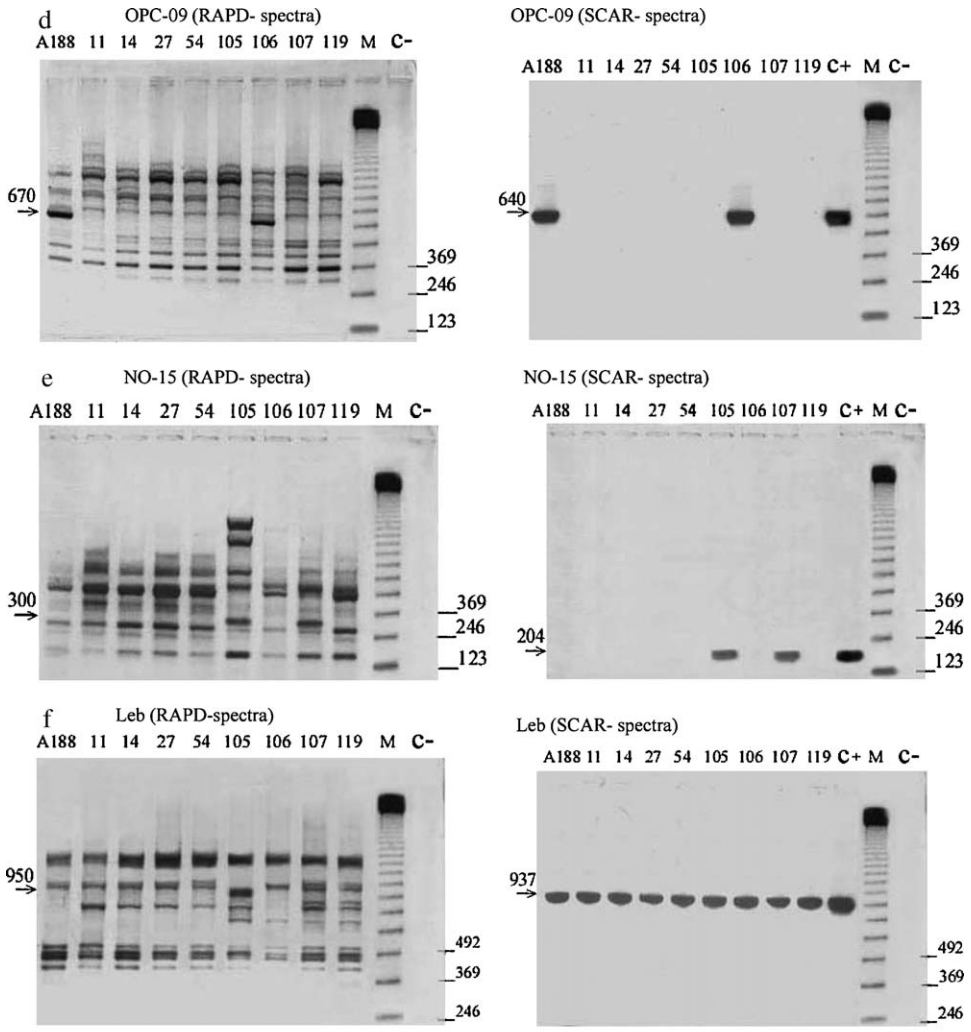


Fig. 1. (*Continued*).

1 and *rp3-2* genes. The M-10 marker may have arisen as a result of retrotransposon insertion near the *rpl26* pseudogene in the spacer region between the *rp3-1* and *rp3-2* genes.

The binding site of the NO-15 reverse primer differed from the corresponding sequence in the genome of *Z. mays* (CC421029) by 5 bp (Fig. 2). In fact, the primer could not anneal to the site in the DNA of *Z. mays* without the deletion of these 5 bp. Apparently, such a microdeletion occurred in somaclones R105 and R107, thus permitting primer annealing and synthesis of the NO-15 PCR product.

Sequences QR-2 and QR-A differed from the homologous genomic sequences of maize by two additional fragments. QR-2 contained a 47-bp fragment (826–872 bp) that was absent from all

Table 1
Presence of SCAR markers in various maize lines, hybrid, cultivar, and somaclones.

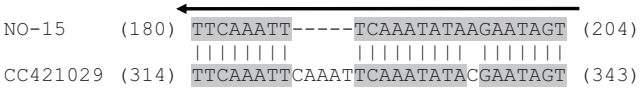
Analyzed maize variants		No. plants with marker fragment:without fragment					Fragment QR			
		Q-20	M-10	OPC-09	NO-15	Leb	–[a]	2	A	2+A
Chi31 × Cateto SG		0:10	0:10	10:0	0:10	10:0	0	0	10	0
Mangelsdorf Tester		0:10	3:7	7:3	0:10	10:0	3	7	0	0
Gk26		0:10	8:2	10:0	0:10	10:0	5	0	3	2
'Pannonia'		0:10	8:2	10:0	0:10	10:0	0	10	0	0
A188		0:8	0:8	8:0	0:8	8:0	0	0	8	0
A188 originated somaclones	Group 1[b]	–	+	–	–	+/–	–	+	–	–
	Group 2[b]	+	–	106,119	105,107	+/–	–	+	106	106,119

[a] Absence of any fragment.
[b] The fragment presence (+) or absence (–) in all somaclones of the group; if a fragment is presented in some somaclones of a group, the number of such somaclone is given. Thorough analysis is given in Supplementary Table 2.

Table 2
The analysis of SCAR marker nucleotide sequences.

Fragment designation and its size in bp	Fragment region, bp	The highest similarity	Similarity %	Probability of random similarity
M-10 (806)	1–236	LTR retrotransposons of Opie2 type DQ186876	91%	6e-84
	1–238	A great number of unclassified retrotransposons ZRSsTERTOOT32286 gi|32194078|nt41–728	96%	3.6e-45
	236–655	Noncoding region between *rp3-1* and *rp3-2t* genes AY574035 AY574035	89%	1e-150
	689–793		95%	8e-38
	308–613	Pseudogene *rpl26*, between *rp3-1* and *rp3-2t* genes AY574035[a]	88%	5e-105
OPC-09 (640)		Unidentified genomic sequences		
	1–638	CC620242	98%	0.0
	1–640	AZM5_68605	98%	4.1e-135
NO-15 (204)		Unidentified genomic sequences		
	1–193	CC421029	97%	5e-87
	1–204	AZM5_15261	91%	2.5e-33
Q-20 (796)	1–796	Noncoding region between mitochondrial *trnF-cp* and *nad2* genes of *Zea mays* subsp. *mays* genotype CMS-T (DQ490953)	99%	0.0
	1–796	*Zea perennis* mitochondrion, complete genome (DQ645538)	98%	0.0
	1–796	*Zea luxurians* mitochondrion, complete genome (DQ645537)	97%	0.0
QR-A (1428)		Unidentified genomic sequences		
	188–491	CG329708	97%	4e-145
	945–1428	CG329720	95%	0.0
	945–1428	BH879991	94%	0.0
	945–1428	AZM5_44674	95%	2.5e-95
	492–948 (Q)	Unidentified genomic sequence CC381326	98%	0.0
QR-2 (1018)		Unidentified genomic sequences		
	187–826	CG329708	98%	0.0
	187–826	AZM5_44674	98%	1.2e-159
	873–1018	AZM5_44674	92%	1.2e-159
Leb (937)		A great number of unidentified genomic sequences, possibly repeats		
	4–937	AY530952	88%	0.0
	1–937	AZM5_74043	91%	1.6e-172
	20–871	A great number of unclassified retrotransposons ZRSsTERT00201433 gi|32025499|nt1-943 Ty3-gypsy-like	65%	4.3e-58

[a] A 95% similarity of corresponding AY574035 region with the nucleotide sequence of *rpl26* (AF093540) (4e-148).



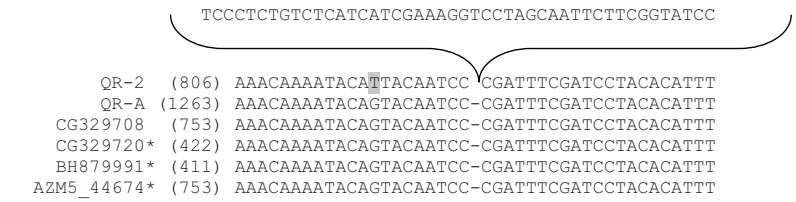
Arrow indicates primer direction (5' →3'). Nucleotides involved in primer hybridization are marked with a grey background.

Fig. 2. Fragment of the alignment of NO-15 reverse primer complementary sequence and unidentified maize genomic sequence CC421029.

other highly similar sequences in GenBank and QR-A (Fig. 3). In QR-A, a 457-bp fragment (492–948 bp) was detected and denoted as Q; this region was absent from QR-2 and other highly similar maize sequences (Fig. 4). A sequence highly similar to region Q was previously detected in the maize genome CC381326, but apart from region Q, this genomic sequence had no similarity with the QR-2/QR-A sequences.

We analyzed the sequence of region Q to determine whether it could be a transposon. Region Q was flanked by the direct 3-bp repeat 5′-TAA-3′ and the inverted repeat 5′-GGGCCTGTTTGGAA-3′ (Fig. 4). The direct repeat may represent target site duplication (TSD), whereas the 14-bp repeat is similar to the terminal inverted repeats (TIRs) of a transposon [16]. Region Q is small (454 bp without 5′-TAA-3′), A/T-enriched (62.56%), has no open reading frames, and can produce the secondary structure required for amplification [17] and transposon excision [18] (Fig. 5). All of these features confirm its similarity with the nonautonomic *Tourist*-like MITE [19,20,16]. Therefore, the absence of region Q from the somaclones that originated from A188 suggests its precise excision from the original A188 cells during *in vitro* culture.

The Leb fragment shared similarity with a large number of sequences and may be a member of a repeat family. Nevertheless, some somaclonal progenies demonstrated loss of the Leb



*The complementary sequences were used for alignment. Non-identical nucleotide is marked with a grey background. The numbers in brackets denote position of 5'-end nucleotide in a whole sequence.

Fig. 3. Fragment of the alignment of QR-2, QR-A, and genomic sequences.

*The complementary sequences were used for alignment. Light-grey background marks TIR, terminal inverted repeats of 14 nucleotides (noninverted nucleotides in the repeat remained unmarked), dark-grey background marks TSD, target site duplication of 3 nucleotides. Fragment QR-A (492-948 bp) is similar (by 98.9%) to the fragment of the CC381326 sequence (286-745 bp); other regions of CC381326 have no substantial similarity with QR-A, QR-2, and other sequences and therefore are not shown in the figure. The numbers in brackets denote position of 5'-end nucleotide in a whole sequence.

Fig. 4. Fragment of the alignment of QR-2, QR-A, and genomic sequences.

marker (Supplementary Table 2), therefore raising the question of how big genomic rearrangements must be to prevent the synthesis of PCR products from a large family of repeats. We performed multiple alignments of the Leb fragment with the most similar genomic maize sequences in GenBank and analyzed the potential annealing sites for both Leb primers in these sequences (Fig. 6). In most of them, the penultimate residue of the 3′-end of the direct primer hybridization site was occupied by different nucleotides. Some sequences also differed in the 3′-terminal nucleotide (Fig. 6). In some sequences, a number of differences were observed in the annealing site of the reverse primer. Only 3 genomic sequences out of the 10 presented in Fig. 6 (more diverged sequences not shown) had appropriate annealing sites for both of the Leb primers, and these sites could easily be disrupted by recombination.

The entire sequence of the Q-20 fragment demonstrated high similarity with mitochondrial (mt) DNA; the highly similar regions were found in the noncoding region between the *trnF-cp* and *nad2* genes in *Z. mays* subsp. *mays* genotype CMS-T (DQ490953) and in the noncoding region between the *nad5* and *atp8* genes in the related species *Z. perennis* (DQ645538, similarity 98%) and *Z. luxurians* (DQ645537, similarity 97%) (Fig. 7). Such sequences were absent from the mtDNA of A188 (DQ490952) and other *Z. mays* genotypes (AY506529, DQ490951, DQ645536, and DQ645539). Some possible reasons for the appearance of Q-20 in somaclones are analyzed in the discussion section.

Fragment OPC-09 shared high similarity (98%) with maize genomic sequences CC620242 and AZM5_68605. However, the nature of the mutations that resulted in the disappearance of this PCR product from all somaclones except R106 remains unclear.

To verify the conclusions we obtained using PCR-based methods, we analyzed the distribution of one SCAR marker using a different molecular method. We chose the group-specific marker Q-20 as a probe and verified its distribution among the progenitor line and the somaclones that originated from A188 using Southern blot hybridization. This probe hybridized with DNA from all the somaclones of group two (R105, R106, R107, and R119), whereas no signal was detected with DNA from the initial line A188 or somaclones of group one (R11, R14, R27, and R54; Fig. 8). Thus, the data obtained using DNA hybridization confirmed the results of PCR-based methods.

4. Discussion

In the present work, we analyzed the sequences of SCAR markers that distinguished the initial line A188 from somaclones. The information we obtained suggests that the genetic changes in the somaclones can be caused by different molecular processes. For instance, marker M-10 probably arose after a transposon insertion near the pseudogene *rpl26* (Table 2). Furthermore, the conversion of marker QR-2 into QR-A most likely resulted from the excision of a small MITE-type transposon (Table 2, Figs. 4 and 5). However, marker NO-15 possibly arose from a microdeletion (Fig. 2). In contrast, the appearance/disappearance of the Leb marker may have been caused by point mutations and/or recombination events. The Leb fragment is similar to a large number of sequences in the maize genome. Nevertheless, only a minor fraction of this repeat family possesses the appropriate annealing sites for these primers (Fig. 6). A repeat element often bears the appropriate annealing site for only one SCAR Leb primer, which makes amplification impossible. Recombination shuffles repeat elements and can bridge the appropriate annealing sites from two different repeat elements or disrupt connections between these sites in one repeat element.

The molecular event that led to the formation of the Q-20 marker in all somaclones from the second group is not obvious. The sequence of the somaclonal Q-20 fragment is similar to the mtDNA of the *Z. mays* subsp. *mays* genotype CMS-T (DQ490953) and of other representatives of the genus, *Z. perennis* (DQ645538) and *Z. luxurians* (DQ645537), giving an e-value of 0.0 in all three cases (Table 2). The presence of the Q-20 fragment in the mtDNA is supported by the absence of segregation in hybrids (F_2 105 × 188) of somaclones with the initial A188 line, with the somaclones used as the maternal line (Supplementary Table 2). However, no similar sequence was found in the mtDNA of other lines of *Z. mays*, including A188 (DQ490952, AY506529, DQ490951, DQ645536, and DQ645539) [21]. The rearrangement of mtDNA during *in vitro* culture has previously been described in several reports. In particular, Kemble et al. [22] observed mtDNA modification in maize regenerants, Rani et al. [23] observed polymorphism of mtDNA in coffee regenerants, and Devarumath et al. [24] observed the same polymorphism in tea regenerants. However, this cannot explain why a long sequence similar to the one that exists in other *Zea* plants arose in somaclones that originated from A188. The probability of formation of such a similar sequence by random events is zero (e = 0.0, Table 2). Another explanation is more speculative. The plant mitochondrial genome is known to be highly variable, and it is present in multiple copies; more than a single type of mtDNA occurs, even in inbred lines [25]. A minor fraction of the CMS-T mitochondrial genotype that is undetectable using PCR or sequencing may be present in the original A188 line. In this case, this type of mtDNA could propagate in somaclones of the second group, and

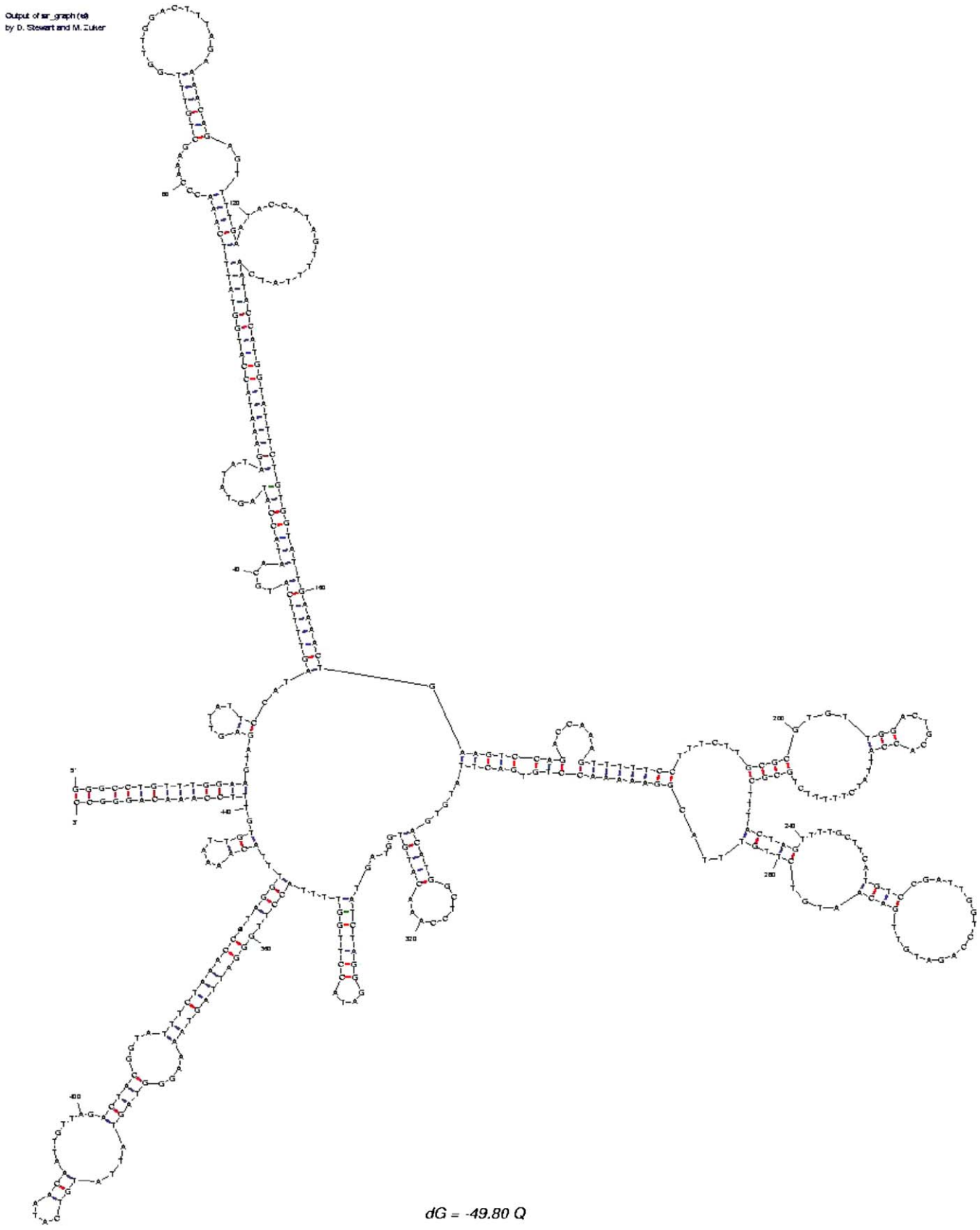
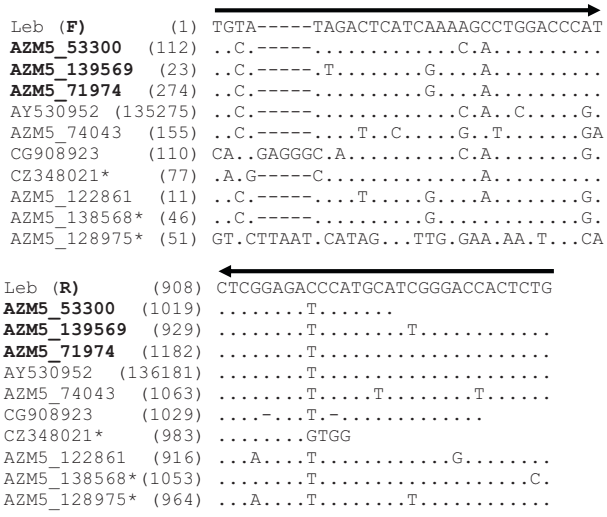


Fig. 5. Putative secondary structure of the region Q.



***The complementary sequences were used for alignment. F and R – forward and reverse primers. Points (.) – identical nucleotides; gap ( ) – absence of information about the sequence. Arrows indicate primer direction (5'→3'). Sequences with appropriate annealing sites for both primers are marked by bold. The numbers in brackets denote position of 5'-end nucleotide in a whole sequence.**

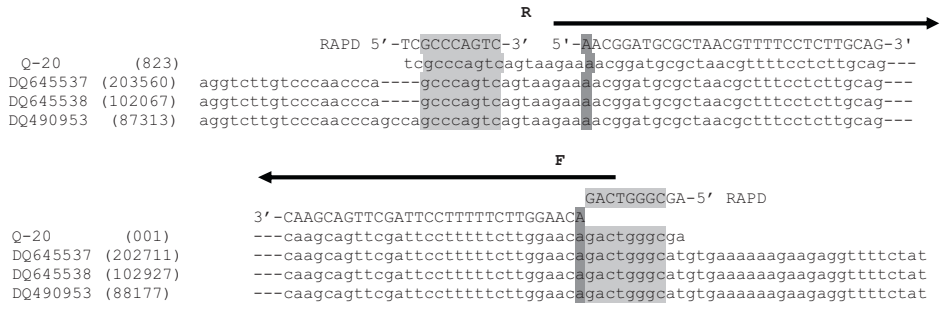
Fig. 6. Alignments of potential annealing sites for Leb primers in several highly similar genomic sequences of maize.

the Q-20 fragment was detected using SCAR primers and Southern hybridization.

The sequence of the OPC-09 fragment did not yield clues regarding the mechanism by which this fragment could form.

It is interesting to note that various somaclones within each group demonstrate similarities in morphological and molecular characteristics (M-10 for the first group and Q-20 for the second group). The reason may be that each somaclonal group originated from a single callus. However, it is more difficult to explain why we observed variability of the same traits in two groups of independently produced somaclones, namely, quantitative characteristics of the tassel and ear [2,3] (Supplementary Table 1), changes in the activity and tissue specificity of esterase and peroxidase [5,3], and changes in the QR-2 SCAR marker [11]. However, the similarity of the changes in the independently produced somaclones could be explained by the selection of definite genotypes during *in vitro* culture. It is supposed that the culture conditions facilitate the accelerated proliferation of cells with selectively advantageous mutations [26]. However, the presence of identical fragments in somaclones produced in independent experiments implies that *in vitro* culture induces specific changes in certain genome regions characterized by a high frequency of mutations. The occurrence of such 'hot spots' has previously been proposed [27]. Bohanec et al. [28] observed the presence or absence of similar RAPD fragments in homozygous onion regenerants produced from different donor plants. Polanco and Ruiz [29] detected similar AFLP fragments among *Arabidopsis* somaclones produced from different calli. Similarly, Kozyrenko et al. [30] found a similar RAPD fragment characteristic of all yellow iris regenerants that was absent in the original plant. In addition, Arnholdt-Schmitt [31] observed the loss of the same RAPD fragment in three carrot explants; however, before culture, this fragment was present in all explants. The author theorized that *in vitro* culture could induce process similar to aging, i.e., the progressive loss of repeated sequences at the end of chromosomes (telomere shortening) [31].

It is also interesting that the SCAR fragments characteristics of the somaclones are present in maize varieties unrelated to A188. Alleles that already exist in maize populations may appear in the hypervariable regions of the somaclone during *in vitro* culture. Similar cases have previously been described. Roth et al. [32] detected RFLP fragments that were absent in original soybean cultivars but appeared in tissue cultures of this cultivar and other soybean cultivars. Rostiana et al. [33] reported missing or newly acquired ISSR fragments of the same size in horseradish regenerants that originated from different calli, mother plants, and strains. Al-Zahim et al. [34] found a similar RAPD fragment in somaclones of the garlic cultivar 'California Late' and in the parental plant cultivar 'Chinese'. Linacero et al. [27] observed the appearance or absence of sim-



Light-grey background marks inverted repeats which are parts of the RAPD primer. Dark-grey background marks 5'-nucleotide of the SCAR primer. Arrows indicate primer direction (5'→ 3'). The numbers in brackets denote position of 5'-end nucleotide in a whole sequence.

Fig. 7. Fragment of the alignment of Q-20 and similar genomic sequences.

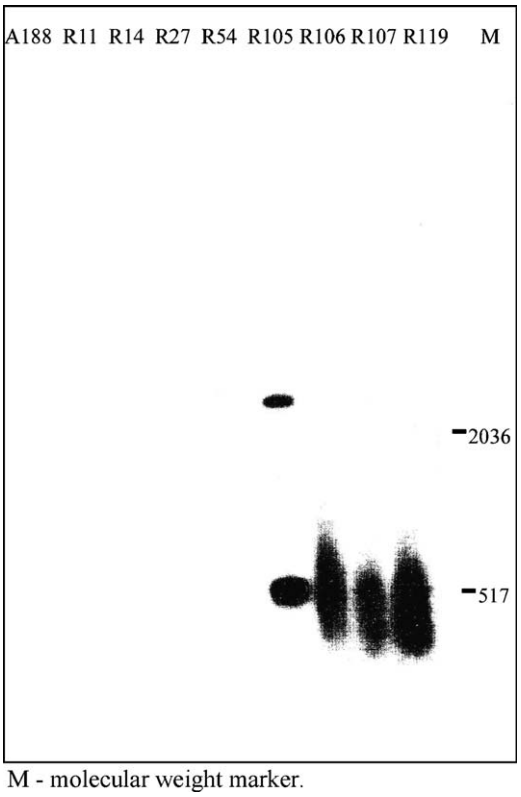


Fig. 8. Southern blot hybridization of DNA from A188 and somaclonal lines with Q-20 fragment as a probe.

ilar RAPD fragments in rye regenerants produced from different cultivars.

One of the causes of hypervariability of some genomic regions may be the sites of mobile element insertion positioned in these regions. In the present study, the markers M-10 and QR-2 were dependent on transposon insertion/excision (Table 2). The appearance of the M-10 fragment could result from the insertion of a retrotransposon near the *rpl26* pseudogene situated in the spacer region between the *rp3-1* and *rp3-2* genes. The M-10 marker was absent in all A188 plants and in all somaclones from group two, but it was present in all somaclones from group one (Table 1, Supplementary Table 2). We assume that the M-10 marker arose in group one during *in vitro* culture, but we cannot exclude the possibility that it existed in the progenitor A188 plant. The presence of the M-10 marker in the majority of Gk26 and 'Pannonia' specimens and its absence in the majority of plants we analyzed from A188, Mangelsdorf Tester, and Chi31 × Cateto SG lines implies that a site near the *rpl26* pseudogene in the *Z. mays* genome is a preferential site for retrotransposon insertion/excision. The activation of LTR retrotransposons *in vitro* has previously been reported. Examples include *Tnt1* in tobacco protoplasts [35] and regenerants [36], *Tto1* and *Tto2* in tobacco cell culture [37], *Tos10*, *Tos17*, and *Tos19* in rice cell culture [38], *Karma* in rice regenerants [39], and *Cyclop* and *Ogre* in pea cell culture [40]. The occurrence of preferential insertion sites for retrotransposons has also been reported. Miyao et al. [41] analyzed rice somaclones and reported that *Tos17* avoided the retrotransposon-rich pericentromeric region and preferred the protein kinase and disease-resistance genes.

The conversion of QR-A/QR-2 markers is probably dependent on the presence/absence of region Q (Fig. 4). The SCAR QR-A (Q+) marker was found in the original A188 line, the Chi31 × Cateto SG hybrid, and the Gk26 line. Interestingly, the SCAR QR-2 (Q−) marker was present in all somaclones that originated from A188 and the Mangelsdorf Tester, 'Pannonia', and Gk26 lines (simultaneously with Q+) (Table 1). Region Q resembles the nonautonomic *Tourist*-like MITE [19,20,16] but does not share similarities with other MITE sequences, which might be due to the diversity of MITE nucleotide sequences. Pairwise alignments of *Tourist* elements within the same organism indicate that a nucleotide similarity below 70% is not uncommon [42]. The transposition of nonautonomous MITE *Tourist* during *in vitro* culture has previously been reported [43,44] as was the detection of a mobile element in the RAPD fragment [45]. Excision of the MITEs *Tourist* and *Stowaway* can occur precisely, or some nucleotides of a transposon (footprints) may remain in the genomic site after an excision [46,44,47,48]. Region Q disappeared without any footprints (Fig. 4). There is a probability that Q was excised in the source A188 plant material. If this was the case, the eight somaclone families obtained from the two independent experiments should have arisen from A188 (Q−). For this assumption to be likely, A188 (Q−) should be widespread in the population. However, Q− was not found in the analyzed A188, but it was found in other maize variants, such as Mangelsdorf Tester, Gk26, and 'Pannonia'. Therefore, the assumption that region Q was excised from A188 (Q+) during *in vitro* culture seems more probable to us.

Previous investigations on the specificity of RAPD and ISSR fragments for somaclones and cell culture have demonstrated that both the coding and noncoding genome regions undergo genetic changes *in vitro* [49–54]. In the present study, we report that changes in the DNA sequence of the somaclones occurred only in noncoding regions. The large amount of noncoding sequences in the genome (up to 80% in maize [55,56]) may account for the higher probability of mutations observed only in these regions. Another reason may be that the maize genome is well studied. Sequences that share similarity with known genes are not always real genes. For example, the LTR M-10 fragment of the retrotransposon is located next to a sequence with 95% similarity to the gene *rpl26* (Table 2), but this sequence is a pseudogene and is therefore noncoding.

To summarize, we conclude that diverse processes could result in changes in RAPD and ISSR fragments, such as the insertion/excision of transposon(s) (M-10, QR-A/QR-2), microdeletion (NO-15), insertion (QR-2), recombination (Leb), and mtDNA modification (Q-20), as observed in noncoding maize sequences.

Acknowledgments

The authors are grateful to D. Ivanov and "Genome" IMB RAS (http://www.genome-centre.narod.ru) for their help in sequencing, F. Rosov for help with Southern hybridization, E. Kochieva for providing us with ISSR primers, Z. Kokaeva, O. Lebedeva, and L. Patrushev for providing us with RAPD primers, A. Alekseevsky and S. Spirin for their advice on bioinformatics, and E. Khavkin for constructive criticism.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2010.10.004.

References

[1] P.J. Larkin, W.R. Scowcroft, Somaclon variation—a novel source of variability from cell culture for plant improvement, Theor. Appl. Genet. 60 (1981) 197–214.
[2] Yu.I. Dolgikh, S.N. Larina, Yu.V. Bolonkina, Analysis of somaclonal variability of inbred maize line A188, Russ. J. Genet. 28 (1992) 74–82.
[3] Yu.I. Dolgikh, Somaclonal variability of plants and possibility of its practical application (on the maize model), Dissertation, K.A. Timiryazev Institute of Plant Physiology, Russian Academy of Sciences, 2005.
[4] D.M. Ataeva, M.V. Dubrovina, Yu.I. Dolgikh, S.N. Larina, Z.B. Shamina, Variability of heterochromatic chromosome regions of the first seed generation of maize regenerants, Cytol. Genet. 28 (1994) 15–20.

[5] E.E. Khavkin, M.V. Zabrodina, Heritable variations in peroxidase and esterase isoenzyme patterns of maize somaclones, Russ. J. Plant Physiol. 41 (1994) 754–761.
[6] J.G.K. Williams, A.R. Kibelik, K.J. Livak, Y.A. Rafalski, S.V. Tingey, DNA Polymorphisms amplified by arbitrary primers are useful as genetic markers, Nucleic Acids Res. 18 (1990) 6531–6535.
[7] R.V. Kantety, X.P. Zeng, J.L. Bennetzen, B.E. Zehr, Assessment of genetic diversity in Dent and Popcorn (*Zea mays* L.) inbred lines using inter-simple sequence repeat (ISSR) amplification, Mol. Breed. 1 (1995) 365–373.
[8] I. Paran, R.W. Michelmore, Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce, Theor. Appl. Genet. 85 (1993) 985–993.
[9] E.S. Osipova, Z.G. Kokaeva, A.V. Troitskij, Yu.I. Dolgikh, Z.B. Shamina, S.A. Gostimskij, RAPD analysis of maize somaclones, Russ. J. Genet. 37 (2001) 80–84.
[10] E.S. Osipova, Revealing of the somaclonal variability by molecular markers (RAPD, ISSR) in maize, Dissertation, K.A. Timiryazev Institute of Plant Physiology, Russian Academy of Sciences, 2003.
[11] E.S. Osipova, O.V. Koveza, A.V. Troitskij, Yu.I. Dolgikh, Z.B. Shamina, S.A. Gostimskij, Analysis of specific RAPD and ISSR fragments in maize (*Zea mays* L.) somaclones and development of SCAR markers on their basis, Russ. J. Genet. 39 (2003) 1412–1419.
[12] T. Murashige, F. Skoog, A revised medium for rapid growth and bio assay with tobacco tissue cultures, Plant Physiol. 15 (1962) 473–482.
[13] W. Rychlik, Selection of primers for polymerase chain reaction, Mol. Biotechnol. 3 (1999) 129–134.
[14] M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Res. 31 (2003) 3406–3415.
[15] J.D. Thompson, D.G. Higgins, T.J. Gibson, W. Clustal, Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[16] X. Zhang, N. Jiang, C. Feschotte, S.R. Wessler, Distribution and evolution of PIF- and Pong-like transposons and their relationships with Tourist-like MITEs, Genetics 166 (2004) 971–986.
[17] Z. Izsvak, Z. Ivics, N. Shimoda, D. Mohn, H. Okamoto, P.B. Hackett, Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification, J. Mol. Evol. 48 (1999) 13–21.
[18] K. Fujino, H. Sekiguchi, Site specific cytosine methylation in rice nonautonomous transposable element nDart, Plant Mol. Biol. 67 (2008) 511–518.
[19] T.E. Bureau, S.R. Wessler, Tourist: A large family of small inverted repeat frequently associated with maize genes elements, Plant Cell 4 (1992) 1283–1294.
[20] S.R. Wessler, T.E. Bureau, S.E. White, LTR-retrotransposons and MITEs: important players in the evolution of plant genomes, Curr. Opin. Genet. Dev. 5 (1995) 814–821.
[21] J.O. Allen, C.M. Fauron, P. Minx, L. Roark, S. Oddiraju, G.N. Lin, L. Meyer, H. Sun, K. Kim, C. Wang, F. Du, D. Xu, M. Gibson, J. Cifrese, S.W. Clifton, K.J. Newton, Comparisons among two fertile and three male-sterile mitochondrial genomes of maize, Genetics 177 (2007) 1173–1192.
[22] R.J. Kemble, R.B. Flavell, R.I.S. Brettell, Mitochondrial DNA analyses of fertile and sterile maize plants derived from tissue culture with the texas male sterile cytoplasm, Theor. Appl. Genet. 62 (1982) 213–217.
[23] V. Rani, K.P. Singh, B. Shiran, S. Nandy, S. Goel, R.M. Devarumath, H.L. Sreenath, S.N. Raina, Evidence for new nuclear and mitochondrial genome organizations among high-frequency somatic embryogenesis-derived plants of allotetraploid *Coffea arabica* L. (Rubiaceae), Plant Cell Rep. 19 (2000) 1013–1020.
[24] R. Devarumath, S. Nandy, V. Rani, S. Marimuthu, N. Muraleedharan, S. Raina, RAPD, ISSR and RFLP fingerprints as useful markers to evaluate genetic integrity of micropropagated plants of three diploid and triploid elite tea clones representing *Camellia sinensis* (China type) and *C. assamica* ssp. assamica (Assam–India type), Plant Cell Rep 21 (2002) 166–173.
[25] M. Szklarczyk, M. Simlat, B. Jagosz, G. Ba, The use of cytoplasmic markers in onion hybrid breeding, Cell Mol. Biol. Lett. 7 (2002) 625–634.
[26] T.A. Ezhova, A.M. Bagrova, S.A. Gostimski, Cell selection as a possible reason for the specificity of somaclonal variation in pea, Plant Breed. 114 (1995) 520–524.
[27] R. Linacero, F.E. Alves, A.M. Vázquez, Hot spots of DNA instability revealed through the study of somaclonal variation in rye, Theor. Appl. Genet. 100 (2000) 506–511.
[28] B. Bohanec, M. Jakse, A. Ihan, B. Javornik, Studies of gynogenesis in onion (*Allium cera* L.): induction procedures and genetic analysis of regenerants, Plant Sci. 104 (1995) 215–224.
[29] M.C. Polanco, M.L. Ruiz, AFLP analysis of somaclonal variation in *Arabidopsis thaliana* regenerated plants, Plant Sci. 162 (2002) 817–824.
[30] M.M. Kozyrenko, E.V. Artiukhova, E.V. Boltenkov, L.S. Lauve, Somaclonal variability of *Iris pseudacorus* L. according to RAPD and cytogenetic analyses, Biotechnol. Russ. 2 (2004) 11–22.
[31] B. Arnholdt-Schmitt, Genomic instabilities in tissue culture – a physiological normality? GEB – Giessener Elektronische Bibliothek (2003). URL: http://geb.uni-giessen.de/geb/volltexte/2003/1243/.
[32] E.J. Roth, B.L. Frazier, N.R. Apuya, K.G. Lark, Genetic variation in an inbred plant: variation in tissue cultures of soybean [*Glycine max* (L.) Merrill], Genetics 121 (1989) 359–368.
[33] O. Rostiana, M. Niwa, W. Marubashi, Efficiency of inter-simple sequence repeat PCR for detecting somaclonal variation among leaf-culture regenerated plants of horseradish, Breed Sci. 49 (1999) 245–250.
[34] M.A. Al-Zahim, B.V. Ford-Lloyd, H.J. Newbury, Detection of somaclonal variation in garlic (*Allium sativum* L.) using RAPD and cytological analysis, Plant Cell Rep. 18 (1999) 473–477.
[35] M.A. Grandbastien, A. Spielmann, M. Caboche, Tnt1. a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics, Nature 337 (1989) 376–380.
[36] Q.H. Le, D. Melayah, E. Bonnivard, M. Petit, M.A. Grandbastien, Distribution dynamics of the Tnt1 retrotransposon in tobacco, Mol. Genet. Genom. 278 (2007) 639–651.
[37] H. Hirochika, Activation of tobacco retrotransposons during tissue culture, EMBO J. 12 (1993) 2521–2528.
[38] H. Hirochika, K. Sugimoto, Y. Otsuki, H. Tsugawa, M. Kanda, Retrotransposons of rice involved in mutations induced by tissue culture, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 7783–7788.
[39] M. Komatsu, K. Shimamoto, J. Kyozuka, Two-step regulation and continuous retrotransposition of the rice LINE-type retrotransposon Karma, Plant Cell 15 (2003) 1934–1944.
[40] P. Smykal, L. Valledor, R. Rodriguez, M. Griga, Assessment of genetic and epigenetic stability in long-term *in vitro* shoot culture of pea (*Pisum sativum* L.), Plant Cell Rep. 26 (2007) 1985–1998.
[41] A. Miyao, K. Tanaka, H. Hirochika, Target specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich region of the genome, Plant Cell 15 (2003) 1771–1780.
[42] T.E. Bureau, S.R. Wessler, Mobile inverted-repeat elements of the Tourist family are associated with the genes of many cereal grasses, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 1411–1415.
[43] N. Jiang, Z. Bao, X. Zhang, H. Hiroshika, S.R. Eddy, S. McCouch, S.R. Wessler, An active DNA transposon family in rice, Nature 421 (2003) 163–167.
[44] K. Kikuchi, K. Terauchi, M. Wada, H.Y. Hirano, The plant MITE mPing is mobilized in anther culture, Nature 421 (2003) 167–170.
[45] E. Alves, I. Ballesteros, R. Linacero, A.M. Vazquez, RYS1, a foldback transposon, is activated by tissue culture and shows preferential insertion points into the rye genome, Theor. Appl. Genet. 111 (2005) 431–436.
[46] G. Petersen, O. Seberg, Phylogenetic evidence for excision of *Stowaway* miniature inverted-repeat transposable elements in Triticeae (Poaceae), Mol. Biol. Evol. 17 (2000) 1589–1596.
[47] R.J. Mason-Gamer, Multiple homoplasious insertions and deletions of a Triticeae (Poaceae) DNA transposon: a phylogenetic perspective, BMC Evol. Biol. 7 (2007) 92, URL: http://www.biomedcentral.com/1471-2148/7/92, doi:10.1186/1471-2148r-r7-92.
[48] G. Yang, F. Zhang, C.N. Hancock, S. Wessler, Transposition of the rice miniature inverted repeat transposable element mPing in *Arabidopsis thaliana*, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 10962–10967.
[49] H. Yang, Y. Tabei, H. Kamada, T. Kayano, F. Takaiwa, Detection of somaclonal variation in cultured rice cells using digoxigenin-based random amplified polymorphic DNA, Plant Cell Rep. 18 (1999) 520–526.
[50] P. Bogani, A. Simoni, P. Lio, A. Germinario, M. Buiatti, Molecular variation in plant cell populations evolving *in vitro* in different physiological contexts, Genome 44 (2001) 549–558.
[51] X.J. Leroy, K. Leon, J.M. Hily, P. Chameil, M. Branchard, Detection of *in vitro* culture-induced instability through inter-simple sequence repeat analysis, Theor. Appl. Genet. 102 (2001) 885–891.
[52] W. Guo, Y. Li, L. Gong, F. Li, Y. Dong, B. Liu, Efficient micropropagation of *Robinia ambigua* var. *idahoensis* (Idaho Locust) and detection of genomic variation by ISSR markers, Plant Cell Tiss. Org. 84 (2006) 343–351.
[53] W.L. Guo, L. Gong, Z.F. Ding, Y.D. Li, F.X. Li, S.P. Zhao, B. Liu, Genomic instability in phenotypically normal regenerants of medicinal plant *Codonopsis lanceolata* Benth. et Hook. f., as revealed by ISSR and RAPD markers, Plant Cell Rep. 25 (2006) 896–906.
[54] F. Ngezahayo, Y. Dong, B. Liu, Somaclonal variation at the nucleotide sequence level in rice (*Oryza sativa* L.) as revealed by RAPD and ISSR markers, and by pairwise sequence analysis, J. Appl. Genet. 48 (2007) 329–336.
[55] P. SanMiguel, A. Tikhonov, Y.K. Jin, N. Motchoulskaia, D. Zakharov, A. Melake-Berhan, P.S. Springer, K.J. Edwards, M. Lee, Z. Avramova, J.L. Bennetzen, Nested retrotransposons in the intergenic regions of the maize genome, Science 274 (1996) 765–768.
[56] B.C. Meyers, S.V. Tingey, M. Morgante, Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome, Genome Res. 11 (2001) 1660–1676.