Pantone 460 C Pantone 457 C

Zeszyty Naukowe Uniwersytetu Rolniczego im. Hugona Kołłątaja w Krakowie

rozprawy

531

Marek Szklarczyk

The search for mitochondrial polymorphisms differentiating cytoplasmic male-sterile and male-fertile beets

zeszyt 408

WYDAWNICTWO UNIWERSYTETU ROLNICZEGO W KRAKOWIE

Zeszyty
Naukowe
Uniwersytetu
Rolniczego
im. Hugona
Kołłątaja
w Krakowie
nr 531

ISSN 1899-3486

rozprawy

zeszyt 408

Marek Szklarczyk

The search for mitochondrial polymorphisms differentiating cytoplasmic male-sterile and male-fertile beets

Praca wykonana w Zakładzie Genetyki,
Hodowli Roślin i Nasiennictwa
Instytutu Biologii Roślin i Biotechnologii
Uniwersytetu Rolniczego w Krakowie

Wydawnictwo Uniwersytetu Rolniczego w Krakowie
Kraków 2016

Pracę opiniowali:
Prof. dr hab. Waldemar Marczewski (Instytut Hodowli i Aklimatyzacji Roślin PIB)
Prof. dr hab. Monika Rakoczy-Trojanowska (SGGW w Warszawie)

Redaktor Naczelny Wydawnictwa
Prof. dr hab. inż. Józef Bieniek

Redaktor Naukowy
Prof. dr hab. inż. Edward Kunicki

Korekta językowa
Lynn Holmes

Wydano za zgodą Rektora Uniwersytetu Rolniczego w Krakowie

Projekt okładki: Monika Wojtaszek-Dziadusz

Publikacje Wydawnictwa UR w Krakowie można nabyć w siedzibie Wydawnictwa.
Prowadzona jest również sprzedaż wysyłkowa (tel. 12 662 51 60).
Księgarnia internetowa: www.wydawnictwo.ur.krakow.pl

Wydawnictwo UR w Krakowie
31-425 Kraków, al. 29 Listopada 46
tel. (12) 662 51 57, 662 51 59
e-mail: wydawnictwo@ur.krakow.pl
www.wydawnictwo.ur.krakow.pl

Ark. wyd. 7. Ark. druk. 6,75. Nakład 120 egz.

Druk i oprawa: DRUKMAR, 30-080 Zabierzów, ul. Rzemieślnicza 10

ISSN 1899-3486

Contents

Acknowledgements

The research received support from the following sources:

- Kutnowska Hodowla Buraka Cukrowego, grant no. (internal) BZ 576 – genotyping with RAPD and vectorette markers,
- Faculty of Biotechnology and Horticulture (University of Agriculture in Krakow), statutory fund no. (internal) DS 3500 – DNA sequencing, identification of SCAR markers, proteomic analyses of sugar beet accessions,
- Polish Ministry of Science and Higher Education, grant no. N N310 437338 – proteomic analyses of table beet accessions.

Grants BZ 576 and N N310 437338 were awarded to Marek Szklarczyk.

The author acknowledges the help of the following persons: Mrs. Anna Herda M.Sc., Mr. Wojciech Wesołowski M.Sc., Mrs. Beata Domnicz M.Sc., Dr. Magdalena Simlat, Eng. Przemysław Gierski and Mrs. Magdalena Jopek M.Sc. Moreover, explicit thanks are addressed to Mr. Mirosław Łakomy M.Sc., and Dr. Adam Sitarski from KHBC as well as to Dr. Leszek Róg from KHiNO Polan for providing the breeding stocks for analysis.

Abbreviations

2-DE	two-dimensional gel electrophoresis
AFLP	amplified fragment length polymorphism
ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CMS	cytoplasmic male sterility
cpDNA	chloroplast DNA
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
E value	expect value
g	gravitational acceleration
IAM	iodoacetamide
IEF	isoelectric focusing
IPG	immobilized pH gradient
IPTG	isopropyl β-D-1-thiogalactopyranoside
kb	kilo-base pair
kDa	kilodalton
LC-MS/MS	liquid chromatography-tandem mass spectrometry
mRNA	messenger RNA
MS	male-sterile
mtDNA	mitochondrial DNA
NCBI	National Center for Biotechnology Information
NGS	next-generation sequencing
nt	nucleotide
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PC1, PC2	principal component 1, 2
PCA	principal component analysis
PCR	polymerase chain reaction

pI	isoelectric point
PTM	post-translational modification
QTL	quantitative trait locus
RAPD	random amplified polymorphic DNA
rRNA	ribosomal RNA
RT	reverse transcriptase
SCAR	sequence-characterized amplified region
SDS	sodium dodecyl sulfate
SNP	single nucleotide polymorphism
STS	sequence-tagged site
TBE	Tris-borate buffer
tRNA	transfer RNA
u	unit of enzyme activity
UMP	universal mitochondrial primer
UV	ultraviolet
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1. Introduction

The first description of cytoplasmic male sterility (CMS) in beets was provided by Owen [1942, 1945] who discovered male-sterile plants in sugar beet variety US-1. According to his model of inheritance, male-sterile plants carried a sterilizing cytoplasm – the S cytoplasm, and the recessive alleles within two nuclear loci designated *X*/*x* and *Z*/*z*. Therefore, the genotype of a male-sterile plant can be depicted as S*xxzz*. Any combination of the S-cytoplasm and the dominant alleles – *X* and *Z* – resulted in a certain degree of male fertility. Plants with the normal (N) cytoplasm were always male-fertile.

In contrast to the yellow turgid anthers of male-fertile plants, the anthers of male-sterile sugar beet flowers are shrunken with coloration changing from white to brown or even black [Kaul 1988]. CMS in sugar beet is of the sporogenous type – male organ development is affected at the stage of tetrad formation or shortly after microspore release. This is related to abnormal tapetum behavior – its cells enlarge as a result of extensive vacuolization. The enlarged tapetal cells depress developing microspores which disintegrate into a blackened mass deposited within the anther locules. As a result, pollen produced by male-sterile plants is not viable [Kinoshita 1971, Majewska-Sawka et al. 1993].

The discovery of CMS in sugar beet was very influential in the breeding of this crop [Skaracis and De Biaggi 2005]. The development of CMS lines enabled a large-scale production of hybrid seed which, since then, has dominated the variety market. Reproduction of the CMS maternal components is based on maintainer lines which carry the N-cytoplasm and the homozygous recessive alleles within the *X*/*x* and *Z*/*z* loci (genotype N*xxzz*). However, the maintainer genotypes, known also as O-types (Owen-types) or zero-types (zero restorers), are very rare in sugar beet populations and therefore, substantial effort is required in their selection. This is achieved by test-crossing a candidate maintainer plant with a male-sterile plant. If the seed harvested from the latter produces only male-sterile progeny, the O-type character of the candidate is evidenced. Since the candidate plant is also subjected to self-pollination, its genotype is further represented by the resulting S1 seed. This mate-

rial is then used for a few cycles of backcrossing to the male-sterile progeny which produces a pair of equivalent lines – the CMS (MS) line and the maintainer (O-type) line. Ideally, these lines should differ genetically only with respect to the harbored cytoplasm, with the S-cytoplasm in the CMS line and the N-cytoplasm in the corresponding maintainer. During commercial hybrid seed production, the CMS lines are crossed with pollinators selected for their appropriate combining ability. The hybrid seed is harvested from the maternal CMS component. After development of this for sugar beet the same breeding concept was used for related crops – table (red) and fodder beet. In both cases the sterilizing cytoplasm was transferred from sugar beet. At least for table beet this was accompanied by the transfer of maintainer *x* and *z* alleles, apparently underrepresented within the respective germplasm. These efforts resulted in the release of the first CMS table beet inbreds in the 60's [Goldman and Navazio 2007] and the first fodder beet hybrids in 1975 [Henry 2010].

Most of the beet CMS research has been performed for sugar beet accessions. As the trait is generally related to alterations of the mitochondrial genome, the major interest was directed at mitochondrial DNA (mtDNA) organization and expression. A comprehensive comparison of the S- and N-cytoplasmic mtDNAs became feasible when their complete nucleotide sequences were available [Kubo et al. 2000, Satoh et al. 2004]. This comparison revealed four transcriptionally active CMS-specific open reading frames (ORFs): *Satp6*, *Scox2–2*, *Sorf324* and *Sorf119*. However, transcription profiles of these sequences were not altered in the context of fertility restoration. Moreover, for *Scox2–2*, *Sorf324* and *Sorf119* the respective translation products were detected neither in the S- nor N-cytoplasmic mitochondria. Interestingly, for *Satp6,* a translation product was found which corresponded to the CMS-specific amino-extension of this ORF (referred to as *Satp6* presequence or *preSatp6*). The 35 kDa preSATP6 protein was detected in the S-cytoplasmic mitochondria, but not in the mitochondria of the normal cytoplasm [Yamamoto et al. 2005]. The protein is likely released from a larger precursor (comprising both the preSATP6 and the core SATP6 portion) and deposited in the mitochondrial membrane fraction where it forms a homomultimeric complex of 200 kDa. Although the accumulation of preSATP6 was not affected upon fertility restoration, this protein has been considered the best candidate for the sugar beet sterility determinant. It is possible that the restorer genes act downstream of preSATP6 or their action is restricted only to a small group of cells – e.g., tapetal cells [Wesołowski et al. 2015].

In parallel to mitochondrial investigations the nuclear CMS-controlling factors have also received considerable attention. Early works were aimed at determining the chromosomal location of the *X/x* and *Z/z* loci. Reports by Schondelmaier and Jung [1997] as well as Hjerdin-Panagopoulos et al. [2002] accordingly mapped these genes on chromosomes III and IV, respectively. The latter authors handled the analyzed genetic factors as quantitative trait loci (QTLs). Positional cloning of the *X* restorer at first led to identification of the linked DNA markers [Hagihara et al.

2005] and then to sequence characterization of both *X* and *x* alleles [Matsuhira et al. 2012]. The *X* allele (referred to as *Rf1* in the original paper) exhibits quadruplication of a gene encoding a homolog of the yeast OMA1 metalloprotease. Of the quadruplicated ORF series (*bvORF18*, *bvORF19*, *bvORF20*, *bvORF21*) only the *bvORF20* sequence restored the fertile condition in a transgenic assay. The *x* (*rf1*) allele contains a single homologous ORF – *bvORF20L* which has an 85% identity to *bvORF20*. Recently Honma et al. [2014] reported on the identification of AFLP markers linked to the *Z* restorer (also referred to as *Rf2*). The informative AFLPs were converted into sequence-tagged site (STS) markers which enabled molecular mapping of the target locus. According to expectations the investigated restorer gene was localized on chromosome IV.

2. Research objectives

The general goal of this research was to develop a new molecular marker strategy to search for polymorphisms in the plant mitochondrial genome (mtDNA) as well as to validate this strategy using the available experimental approaches. This goal was addressed by using a model of cytoplasmic male-sterile and male-fertile beets which substantially differ with respect to the organization of the mitochondrial genome. In addition to the work on a genomic level, a proteomic study was undertaken in order to elucidate whether the identified mtDNA polymorphisms were associated with changes in the accumulation of mitochondrial proteins.

The specific research tasks included:

- identification of plasmotype-specific RAPD markers,
- identification of plasmotype-specific vectorette PCR markers,
- sequence characterization of the identified RAPD and vectorette PCR markers,
- to demonstrate representation of the identified marker sequences in accessions carrying the normal (N) and sterilizing (S) cytoplasm,
- design and verification of plasmotype-specific SCAR markers, and
- determining the plasmotype-specific features of the mitochondrial proteome.

The research hypothesis tested in this study was that the vectorette PCR technique may have the potential for identifying polymorphisms in the plant mitochondrial genome. In order to assess the effectiveness of the vectorette PCR, the search for mtDNA polymorphisms was also performed with the use of RAPDs – once a widely-used method of genotyping.

3. Materials and methods

3.1. Plant material

The accessions represented 24 CMS-maintainer pairs of cultivated beet (*Beta vulgaris* L. subsp. *vulgaris*) (Table 1). The sugar beet accessions were supplied by Kutnowska Hodowla Buraka Cukrowego Ltd. (KHBC) (Straszków, Poland). The two fodder beet lines were provided by Małopolska Hodowla Roślin Ltd. (MHR) (Kraków, Poland). The majority of the table (red) beet lines originated from a breeding program carried out in the Department of Horticulture, University of Wisconsin-Madison (USA). A further two table beet lines were developed in the Department of Genetics, Plant Breeding and Seed Science, University of Agriculture in Krakow (Poland). The four table beet lines used in the proteomic analyses were provided by Krakowska Hodowla i Nasiennictwo Ogrodnicze Polan (KHiNO Polan) (Kraków, Poland).

The plant material was provided either in the form of leaves or storage roots. The leaf material was used for extraction of total DNA while the storage roots were used for isolation of the mitochondria which, in turn, served as a source of mitochondrial DNA and proteins (see below).

Table 1. The list of breeding lines included in the study and the analyses in which they were used

CMS line	Maintainer line	Origin*	Analyses performed				
			RAPD	Vectorette PCR	NGS	SCAR	Proteomic
Sugar beet							
NS 041	NO 041	KHBC	+	–	–	+	–
NS 042	NO 042		+	–	–	+	–
NS 043	NO 043		+	–	–	+	–
NS 044	NO 044		+	–	–	+	–
NS 002	NO 002		–	+	–	–	–

Table 1. cont.

CMS line	Maintainer line	Origin*	Analyses performed				
			RAPD	Vectorette PCR	NGS	SCAR	Proteomic
NS 014	NO 014	KHBC	–	+	–	–	–
NS 023	NO 023		–	+	–	–	–
NS 031	NO 031		–	–	+	+	–
NS 032	NO 032		–	–	–	+	–
NS 033	NO 033		–	–	–	+	–
S 00188	O 00188		–	–	–	+	–
S 00500	O 00500		–	–	–	+	–
S 001299	O 001299		–	–	–	+	–
S 001353	O 001353		–	–	–	+	–
FMS 121	FO 121		–	–	–	–	+
FMS 1218	FO 1218		–	–	–	–	+
Fodder beet							
MSPT 68	LOT 68	MHR	–	+	–	–	–
Table (red) beet							
218A	218B	UWM	–	–	–	+	–
279A	279B		–	–	–	+	–
357A	357B		–	–	–	+	–
W411A	W411B		–	–	–	+	–
AR79A	AR79B	UAK	–	–	–	+	–
Re 1A	Re 1B	KHiNO	–	–	–	–	+
Re 3A	Re 3B		–	–	–	–	+

* KHBC – Kutnowska Hodowla Buraka Cukrowego, MHR – Małopolska Hodowla Roślin, KHiNO – Krakowska Hodowla i Nasiennictwo Ogrodnicze Polan, UWM – University of Wisconsin-Madison, UAK – University of Agriculture in Krakow.

3.2. Isolation of total genomic DNA

A fragment of a leaf was frozen in liquid nitrogen and pulverized using a mortar and pestle. Approximately 100 mg of the powder was suspended in 700 µl of solution D from the RNA isolation procedure introduced by Chomczynski and Sacchi [1987]. The suspension was shaken manually with 600 µl of chloroform/isoamyl alcohol (24 : 1, v/v) for 5 min and then centrifuged at 18 000 × g for 10 min. After

that 500 µl of the upper (aqueous) phase was collected and mixed with an equal volume of cold (–20°C) isopropanol. The samples were again centrifuged at 18 000 × g for 10 min, the resulting DNA pellets were washed with 80% ethanol, dried in a vacuum and dissolved in 100 µl of sterile Milli-Q water.

3.3. Random Amplified Polymorphic DNA (RAPD)

RAPD-PCR was performed in 20 µl containing: 1x *Taq* buffer with KCl (Fermentas), 2 mM $MgCl_2$, 0.25 mM dNTPs, 0.5 µM decamer primer, 1 u of *Taq* DNA polymerase (recombinant, Fermentas) and 30 ng of total genomic DNA. The primer sequences are given in Appendix I. The reactions were carried out in an Eppendorf Mastercycler Gradient programmed for 3 min at 94°C; 40 cycles of 30 s at 93°C, 1 min at 42°C and 2.5 min at 68°C followed by 5 min at 72°C. After amplification the reaction mixtures were supplemented with the 6x Loading Dye (Fermentas) and the products were separated in 1% agarose using TBE as the running buffer, the gel also contained ethidium bromide with a concentration of 0.5 µg/ml. Electrophoresis was performed for 2 h at a field strength of 3.5 V/cm.

The obtained RAPD profiles were then coded in the form of a binary matrix in which the presence of a given product was designated with 1 and its absence with 0. This matrix was an input into Past 2.12 [Hammer et al. 2001], which was used to calculate principal components based on a variance-covariance matrix.

3.4. Isolation of mitochondrial DNA for vectorette PCR

The mitochondria were isolated from 15 g of a sliced storage root using the procedure of Steinborn et al. [1992]. The final mitochondrial pellet was lysed in 210 µl of solution D from the RNA isolation procedure of Chomczynski and Sacchi [1987]. The lysate was extracted with 180 µl of chloroform/isoamyl alcohol (24 : 1, v/v) for 5 min. The resulting emulsion was centrifuged at 10 000 × g for 10 min. 280 µl of the upper (aqueous) phase was transferred into a new tube and supplemented with an equal volume of isopropanol. After mixing, the samples were incubated for 0.5 h at –20°C and centrifuged as previously. The resulting mitochondrial (mtDNA) pellet was washed twice with 80% ethanol, dried in a vacuum and dissolved in 40 µl of sterile Milli-Q water. The samples were then further purified on Qiaprep spin columns. For this purpose the mtDNA preps were supplemented with five volumes of buffer PB. After mixing, the samples were transferred to the columns and centrifuged for 2 min at 15 000 × g. Then the columns were washed with 0.75 and 0.25

ml of buffer PE. Each wash was followed by 2 min centrifugation at 15 000 x g. In order to remove traces of the wash buffer this centrifugation was repeated with the emptied columns. For elution, 30 µl of sterile Milli-Q water was applied onto the center of the silica-gel-membrane of each column. After 1 min pause the columns were centrifuged as previously. The collected eluate was stored at –20°C. Both columns as well as buffers PB and PE originated from Qiaprep Spin Miniprep Kit (Qiagen).

3.5. Vectorette PCR

Generation of vectorette libraries. Restriction was performed in 15 µl containing: 100 ng of mitochondrial DNA (mtDNA), 1x buffer O + Tween 20 (Fermentas) and 7.7 u of *Eco*RI (Fermentas). The reaction was carried out at 37°C for 4 h and then the digestions were directly subjected to ligation. The ligations were carried out in 15 µl containing: 13 µl of the *Eco*RI digest, 0.8 pmol of the *Eco*RI vectorette unit, 1.7 mM ATP, 1.7 mM dithiothreitol (DTT) and 0.35 u of T4 DNA ligase. Except for the restriction digest all components of the ligation mixture originated from the Universal Vectorette System (Sigma-Aldrich). The samples were incubated in a thermal cycler which was programmed for two cycles of 16°C for 60 min and 37°C for 30 min then followed by a final step of 16°C for 30 min. After that the samples were incubated overnight at 4°C, diluted with 100 µl of sterile Milli-Q water and stored at –20°C.

DNA amplification. PCR was carried out in 15 µl containing: 1x PCR Buffer (Fermentas), 2.5 mM $MgCl_2$, 0.25 mM dNTPs, 1 µM vectorette (vec-st) primer, 1 µM universal mitochondrial primer (UMP), 0.75 u of *Taq* DNA polymerase (recombinant, Fermentas) and 0.75 µl of a diluted vectorette library (see above). The primer sequences are given in Appendix I. Vec-st was designed by the author on the basis of earlier sequenced vectorette PCR products, this primer is likely shorter than the vectorette primer provided in the Universal Vectorette System. PCRs were carried out in an Eppendorf Mastercycler Gradient programmed for 5 min at 94°C; 35 cycles of 45 s at 92°C, 45 s at 57°C and 2 min at 72°C followed by 10 min at 72°C. After supplementation with the 6x Loading Dye (Fermentas), the vectorette PCR products were separated in 1% agarose using TBE as the running buffer, the gel also contained ethidium bromide at the concentration of 1 µg/ml. Electrophoresis was performed for 1–2 h at a field strength of 4 V/cm.

As in the case of RAPDs, the obtained vectorette PCR profiles were coded into a binary matrix which was used for the calculation of principal components with Past 2.12 [Hammer et al. 2001]. Principal component analysis (PCA) was carried out in order to illustrate how the analyzed beet accessions were differentiated with the use of RAPDs and vectorette PCR markers.

3.6. Conventional PCR

These reactions were used for genotyping with SCAR markers. Amplification was carried out in 15 μl containing: 1x *Taq* Buffer with KCl (Fermentas), 5 mM $MgCl_2$, 0.25 mM dNTPs, 0.25 μM either primer, 1.125 u of *Taq* DNA polymerase (recombinant, Fermentas) and 5 ng/μl total genomic DNA. The primer sequences are given in Appendix I. PCRs were carried out in the Eppendorf Mastercycler Gradient programmed as for the vectorette PCR (see above). The resulting reaction products were separated in either 1% or 1.5% agarose for 90–230 min with the remaining conditions as in the case of RAPD-PCR (see above).

3.7. Cloning of PCR products

When the target PCR product was accompanied by other PCR products it was purified from the agarose gel after electrophoresis. When the accompanying products were not visible it was purified directly from the PCR mixture. Prior to the preparative electrophoresis the samples were supplemented with a 6x modified loading buffer (40% sucrose, 0.025% bromophenol blue). The separation was carried out in 1% agarose, at 3.4 V/cm, for 1.5 h using TBE as the running buffer. After electrophoresis the bands of interest were excised from the agarose under UV light and the resulting gel slices were weighed and supplemented with the appropriate volume of Membrane Binding Solution (10 μl per 10 mg of a gel slice) from the Wizard SV Gel and PCR Clean-Up System (Promega). When the accompanying PCR products were not present an equal volume of Membrane Binding Solution was added directly to the PCR reaction mixture. Further steps of purification were performed using the remaining components of the Wizard SV Gel and PCR Clean-Up System under the manufacturer's recommendations. Elution was carried out with 30 μl of sterile Milli-Q water.

Ligation was carried out in 10 μl containing: 1x Rapid Ligation Buffer, 50 ng of pGEM-T Vector, 3 μl of the purified PCR product (15–120 ng) and 3 Weiss units of T4 DNA ligase. Except for the PCR product, all components of the reaction originated from the pGEM-T Vector System I (Promega). The reaction mixtures were prepared on ice and then incubated for 1 h at room temperature and for 16 h at 4°C.

In the next step, the products of ligation were transformed into the chemically competent *Escherichia coli* DH10B cells. For their preparation, 200–300 ml of the LB broth was inoculated with 1 ml of the overnight bacterial culture and incubated with shaking at 37°C until OD_{600} reached 0.3–0.5. Then the culture was poured into 25 ml spin bottles and chilled in ice for 20 min. This was followed by a 10 min cen-

trifugation at 0–2°C at 3 140 × g. The resulting supernatant was discarded and the bacterial pellets were suspended in a 1/2 volume of 50 mM ice-cold calcium chloride (pH 8 adjusted with NaOH). After the samples were centrifuged (with the same parameters as previously) the pelleted cells were suspended in a 1/15 volume of ice-cold 50 mM $CaCl_2$, 10% glycerol, pH 8 (adjusted with NaOH). Finally, the cell suspensions were aliquoted in 100 µl portions into 1.5 ml tubes and stored at –80°C. For the purpose of transformation the whole ligation mixture (10 µl) was added to 100 µl of the competent cells and incubated in ice for 20 min. After that the samples were dipped for 1 min in a 42°C water bath, returned to the ice for 2 min, supplemented with 1 ml of LB broth and incubated with rotation for 1 h at 37°C. The cells were then collected by centrifugation at 6000 x g for 3 min, a 700 µl aliquot of the supernatant was removed and the cells were resuspended in the remaining supernatant portion. The cell suspensions were plated on LB agar supplemented with 100 µg/ml ampicillin and covered with 100 mM IPTG and 50 mg/ml X-Gal using 100 and 30 µl per standard (8.5 cm) Petri dish, respectively. After overnight incubation at 37°C the plates were stored in a refrigerator (4–8°C) allowing the blue color to develop in non-recombinant colonies.

With the use of yellow tips, bacteria from the white colonies were transferred into numbered 1.5 ml tubes filled with 100 µl of sterile Milli-Q water. After dipping in water the same tips were used to streak the bacteria onto the numbered sectors of LB agar plates with 100 µg/ml ampicillin. Therefore, a given white colony (bacterial clone) was represented by a bacterial suspension in water and a streak on the agar plate. The streaked plates were incubated overnight at 37°C and then stored in a refrigerator. The bacterial suspensions were immersed in a boiling water bath for 15 min, cooled in ice for a few minutes and stored at –20°C. These preparations (1.5 µl) were used as a template for PCR (now referred to as colony PCR) in which the reaction conditions were the same as in conventional PCR (see above). The primer combination was dependent on the nature of the cloned PCR product. In the case of RAPD-PCR products, amplification was performed with primers SP6 and T7. For the vectorette PCR products either the original primer pair was used or the combination of one original primer and one vector primer (SP6 or T7) (in this case 'original' means that which was used to obtain the cloned PCR product). When these PCR reactions yielded PCR products of the expected size (this was examined using standard agarose gel electrophoresis), the respective recombinant clone was considered for plasmid isolation and insert sequencing. The bacteria from this clone were localized on the streaked plate (see above) and used to inoculate 10 ml of the LB broth supplemented with 100 µg/ml ampicillin. After overnight growth in a shaking water bath the resulting cultures were used for plasmid DNA isolation performed with the Wizard Plus SV Minipreps DNA Purification System (Promega) according to the manufacturer's recommendations.

3.8. DNA sequencing and sequence data processing

The cloned PCR products were custom sequenced by Genomed (Warsaw, Poland) using the dye terminator method and the primer walk strategy. Homology searches were performed using the Blastn algorithm and the non-redundant nucleotide collection (nr/nt), accessed through the NCBI web page.

Mitochondrial DNA (mtDNA) preparations for next generation (high throughput) sequencing were extracted from the mitochondria isolated according to Mikami et al. [1984]. For each sugar beet accession, isolation of the mitochondria was carried out using 80 g portions from six individual storage roots. Before loading onto the sucrose density gradient the resulting six crude mitochondrial preparations were pooled together. The final mitochondrial pellets were lysed in 200 µl of buffer AP1 from the DNeasy Plant Mini Kit (Qiagen). After the addition of 2 µl of RNase A, the subsequent steps of mtDNA extraction were carried out according to the recommendations of the kit manufacturer, except for the fact that the samples were not dragged through the Qiashredder Mini columns. DNeasy Mini column elution was performed with 40 µl of sterile water. The resulting mtDNA samples were custom sequenced at the University of California-Davis Genome Center using the Illumina next-generation sequencing platform. The indexed libraries were prepared using the Truseq DNA Sample Prep Kit and the Integenx Apollo 324 robot. The libraries were sequenced in a single channel of the Genome Analyzer IIx flow cell using an 80 bp single-read run. The sequence output file contained FASTQ-formatted reads from the four analyzed libraries. This pool of reads was split into four sub-pools corresponding to the four sugar beet breeding lines from which the sequenced mitochondrial DNAs were isolated. The split of sequence reads was performed using the program Split_solexa.py on the basis of the 6 nt indexes which had been added at the stage of library preparation. Next, the sub-pools of reads were mapped onto the sequences of the cytoplasmic markers (RAPD and vectorette PCR) using Stampy v. 1.0.23 [Lunter and Goodson 2011] with the standard settings adopted for the Illumina reads (option '--solexa'). The SAM files produced with Stampy were processed using program Coverage_in_SAM_to_fasta_qual.py to calculate the coverage of the marker sequences with the mapped Illumina reads. This produced the output files which were converted using Coverage_histogram.py into a format acceptable by Gnuplot (various authors). Gnuplot v. 4.6.1 was used to generate plots showing coverage of the marker sequences with the Illumina reads. The output files from Coverage_in_SAM_to_fasta_qual.py were also processed by Fasta_qual_not_covered_regions_stats.py to calculate the percentage of non-covered sequences. Programs Split_solexa.py, Coverage_in_SAM_to_fasta_qual.py, Coverage_histogram.py as well as Fasta_qual_not_covered_regions_stats.py were written by Przemysław Gierski.

3.9. Proteomic analyses

The mitochondria were isolated from 80 g portions of individual storage roots according to the procedure of Mikami et al. [1984] except that the steps related to DNase digestion were omitted. The final mitochondrial pellet was lysed using half the volume of the sample (rehydration) buffer without DTT and ampholytes (see below). Two μl of the lysate were transferred into a new tube and served for the purpose of protein quantitation. The remaining portion of the lysate was stored at –80°C until its use in isoelectric focusing (IEF). Protein quantitation was performed using the Bradford Reagent (Sigma) according to the manufacturer's recommendations, with 25 mM Tris HCl pH 7.5 for dilution of the mitochondrial lysates and serial dilutions of BSA. IEF was performed using 7 cm IPG BlueStrips from Serva Electrophoresis. Protein samples were prepared by adding 20 μg of the mitochondrial protein (corresponding to max. 1 μl of the mitochondrial lysate) into 125 μl of the complete sample (rehydration) buffer (7 M urea, 2 M thiourea, 2% CHAPS, 0.002% bromophenol blue, 20 mM DTT, 0.5% Servalyt 3–10 carrier ampholytes). The gel strips were rehydrated overnight in the protein samples. Rehydration was carried out at 20°C in the chambers of the Serva rehydration tray and during this process the gel strips were coated with mineral oil. IEF was performed in the IEF-SYS unit from Scie-Plas thermostated to 20°C with a circulating water bath. During the run, the gel strips were coated with an overlay of mineral oil and with strips of Parafilm (Pechiney Plastic Packaging Company). Voltage was applied in a series of 150 V/0.5 h, 300 V/0.5 h, 600 V/0.5 h, 1500 V/1.5 h, 3000 V/2–2.5 h and 300 V until the system was switched off (usually after a few minutes). After IEF the gel strips were equilibrated first for 10 min in the equilibration buffer (6 M urea, 75 mM Tris HCl pH 8.8, 30% glycerol, 2% SDS, 0.002% bromophenol blue) with 1% dithiothreitol (DTT) and then for another 10 min in the equilibration buffer with 2.5% iodoacetamide (IAM). Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed using the Tris-tricine discontinuous buffer system developed by Schägger and von Jagow [1987]. The equilibrated gel strips were loaded onto the 4% T, 3% C stacking gel overlaying the 10% T, 3% C separating gel. SDS-PAGE was performed at 4–8°C in the V10-WCDC unit from Scie-Plas for 10 h with the power supply set to 300 V and 120 mA. After electrophoresis in the second dimension, the gels were silver-stained using the protocol of Jungblut and Seifert [1990] except that methanol was substituted for ethanol in the fixation solution and the sensitizing solution was devoid of glutaraldehyde. The stained gels were archived as TIFF images using a flatbed scanner with a transmissive mode. Two-dimensional (2-D) protein patterns were analyzed using Melanie 7.0 software (Genebio). The adopted software parameters were as follows: smooth = 5, sali-

ency = 4, min. area = 5. Spots which corresponded to differentiating proteins were cut out from the gels and sent for LC-MS/MS analysis to the Laboratory of Mass Spectrometry at the Institute of Biochemistry and Biophysics of the Polish Academy of Sciences (Warsaw, Poland).

4. Results

4.1. Identification of cytoplasmic markers with the use of RAPD-PCR

Four pairs of type-O/MS sugar beet lines were included in the analysis: NO 041/NS 041, NO 042/NS 042, NO 043/NS 043 and NO 044/NS 044. Each line was represented by three plants. The analysis was performed with 15 decamer primers – A-01, A-02, A-03, A-05, A-06, A-08, A-09, A-10, A-11, A-12, B-01, B-02, B-03, B-04 and B-05.

The size of the obtained RAPD-PCR products ranged from 0.2 to 3.8 kb with an average length of 1.28 kb. The RAPD primers produced from 3 (primer A-05) to 14 (primer A-09) scorable DNA fragments, and on average the reactions yielded 9.3 fragments per primer. The total number of scored fragments reached a value of 139 and the majority – 93 (67%) – were polymorphic in the analyzed set of accessions. The number of polymorphic fragments per primer ranged from 2 (primer A-05) to 13 (primer A-09) with an average of 6.2. None of the RAPD primers produced solely monomorphic products.

Polymorphism was observed both between the analyzed lines and between plants within a given line. Using principal component analysis (PCA), plants of the same line were generally placed in the vicinity of each other. It was also observed that the lines of a given MS/maintainer pair occupied the neighboring areas of the PCA scatter plot. However, this was not the case for NS 044 and NO 044, which were markedly separated along the PC2 axis. Since the MS and O-type lines of different pairs were intermingled the formation of general plasmotype-specific clusters was not observed (Fig. 1).

Five RAPD-PCR products showed distribution which correlated with the plasmotype of the analyzed materials (Table 2, Fig. 2). Primer A-01 yielded a fragment of 1.3 kb which was characteristic for all the analyzed maintainer lines but missing in the MS lines. The 0.25 kb fragment produced with primer A-08 was present only in the profiles of NO 043 and NO 044. The use of the B-01 primer yielded two fragments which co-segregated with the presence of the S-cytoplasm – they were

observed only in the profiles of the male-sterile lines. The size of these fragments was estimated at 1.9 and 2.2 kb. The 2.7 kb fragment amplified with primer B-05 exhibited the opposite distribution – it was observed only in the profiles of the male-fertile lines. All of the above mentioned plasmotype-specific RAPD-PCR products did not show intra-line polymorphism.

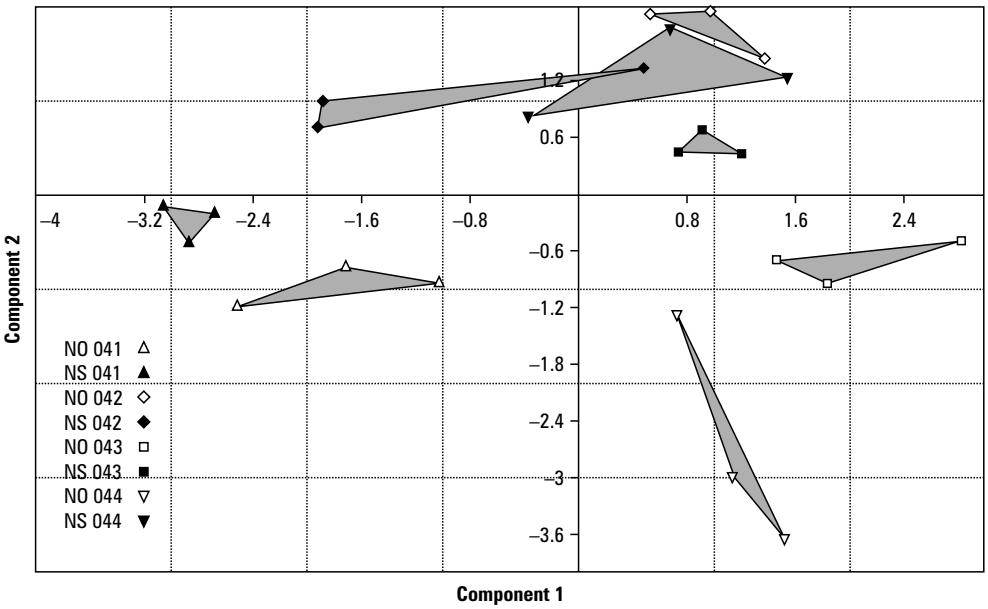


Fig. 1. Principal component analysis (PCA) for plants from eight sugar beet breeding lines (NO 041 – NS 044) based on distribution of RAPD markers. PC1 and PC2 explain 18.7 and 12.9% of the variance, respectively

Table 2. Distribution of plasmotype-correlated RAPD-PCR products [kb] in the set of eight sugar beet breeding lines

Line → Primer ↓	NO 041	NS 041	NO 042	NS 042	NO 043	NS 043	NO 044	NS 044
A-01	1.3	–	1.3	–	1.3	–	1.3	–
A-08	–	–	–	–	0.25	–	0.25	–
B-01	–	2.2	–	2.2	–	2.2	–	2.2
	–	1.9	–	1.9	–	1.9	–	1.9
B-05	2.7	–	2.7	–	2.7	–	2.7	–

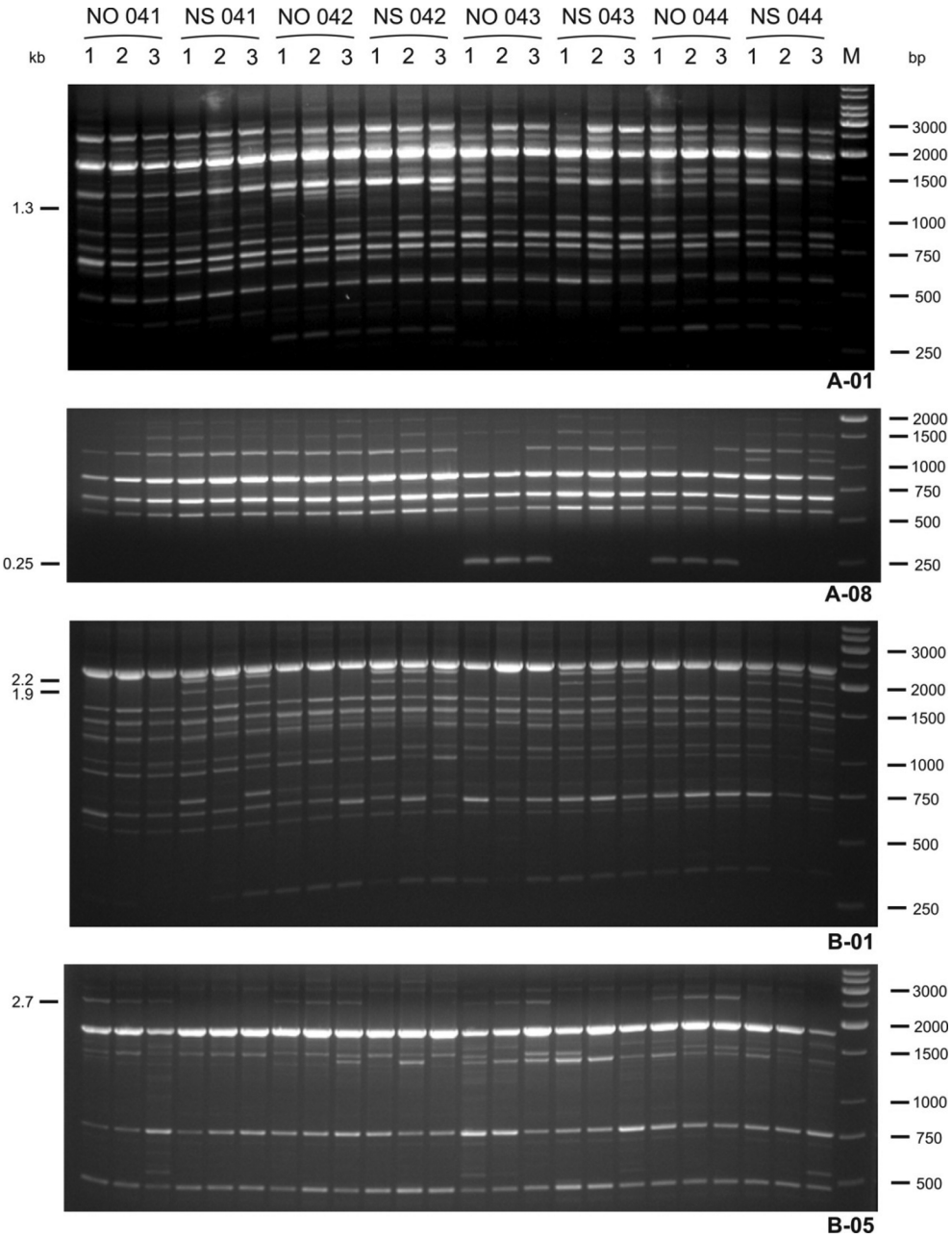


Fig. 2. RAPD-PCR profiles obtained for single plants (1–3) from eight sugar beet breeding lines (NO 041 – NS 044) with the use of primers A-01, A-08, B-01 and B-05. The sizes of the products showing co-segregation with the plasmotype are indicated on the left. M – DNA size marker

4.2. Identification of cytoplasmic markers with the use of vectorette PCR

The analysis was performed for eight breeding lines representing four MS/O pairs. These included six lines of sugar beet – NS 014, NO 014, NS 023, NO 023, NS 002 and NO 002 – as well as two lines of fodder beet – MSPT 68 and LOT 68. Each line was represented by a single plant for which the vectorette library of *Eco*RI-digested mitochondrial DNA was produced. The libraries were subjected to vectorette PCR in which the vectorette primer (vec-st) was accompanied by one of the universal mitochondrial primers (UMPs) (refer to Appendix I). Altogether, the libraries were screened with 105 UMPs (Figs. 3–6).

Four of the universal mitochondrial primers did not yield defined amplification products, and the respective electrophoretic profiles did not contain discrete and scorable bands. For the remaining 101 UMPs at least one product of this type was observed. The size of the amplified DNA fragments ranged from 0.18 to 3.8 kb with an average of 0.75 kb. Altogether, 381 fragments were scored and therefore, an average UMP gave between three to four (3.6) vectorette PCR products. Most of the UMPs yielded only monomorphic DNA fragments (66–62.9%), for 35 (33.3%) UMPs at least one of the amplified products was polymorphic. Since the total number of differentiating products was 48 (12.6%), the mean number of polymorphic fragments per UMP amounted to 0.46.

However, upon closer examination it appeared that the majority of UMPs produced from one to three common monomorphic fragments with sizes of 0.23–0.25, 0.37–0.39 and 0.51–0.53 kb. These fragments were usually weakly amplified and most likely they were produced solely with the vectorette primer which was present in all reactions. Yet another fragment was probably produced this way – its size was 1.58–1.65 kb and it was specific for the maintainer (O) lines (this product was observed in the profiles of 11 UMPs – usually in the absence of strongly amplifying fragments). When the above mentioned fragments were subtracted from the original pool the remaining ones totaled 177. Eight UMPs produced only fragments of this artifactual type and therefore, the number of non-productive UMPs increased to 12. Elimination changed the statistics of the obtained vectorette PCR products – now an average UMP gave less than two products (approx. 1.7%). Exclusively monomorphic DNA fragments were obtained for 64 (60.95%) UMPs, and for 29 (27.6%) UMPs at least one product was polymorphic. After exclusion of the 1.58–1.65 kb product, the number of polymorphic products dropped to 37 (20.9%), which reduced the mean number of polymorphic fragments per UMP to 0.35.

The majority of the polymorphic fragments showed plasmotype-specific distribution defined as the occurrence of a given vectorette PCR product in at least two lines carrying the same cytoplasm and its absence in all lines with the opposite cytoplasm type. These criteria were met for 21 (11.9%) amplified fragments (Table 3). Of these,

11 (6.2%) were specific for the S-cytoplasm – they were produced using the following UMPs: nad3-r, cob-r1, atp6-r, rps7-r1, rrn26-f1, orfX-f, orfX-r (three fragments), nad2ex2-f and nad7ex1-f. The fragments showing specificity for the N-cytoplasm totaled 10 (5.6%) and were obtained with primers: nad6-r1, cob-r1, atp6-r, atp8-f, atp9-f2, rps7-r1, rrn5-r, orfX-r, nad4ex3-f and ccb452ex2a. Therefore, four UMPs – atp6-r, cob-r1, orfX-r and rps7-r1 – generated both S- and N-plasmotype-specific products.

Table 3. Distribution of plasmotype-correlated vectorette PCR products [kb] in the set including two fodder beet (MSPT 68, LOT 68) and six sugar beet (NS 014 – NO 002) breeding lines

Line → UMP ↓	MSPT 68	LOT 68	NS 014	NO 014	NS 023	NO 023	NS 002	NO 002
atp6-r	–	–	–	1.65	–	1.65	–	–
	1.52	–	1.52	–	1.52	–	1.52	–
atp8-f	–	1.82	–	1.82	–	1.82	–	1.82
atp9-f2	–	–	–	1.48	–	1.48	–	–
ccb452ex2a	–	0.75	–	0.75	–	0.75	–	0.75
cob-r1	–	0.95	–	0.95	–	0.95	–	0.95
	0.45	–	0.45	–	0.45	–	0.45	–
nad2ex2-f	0.63	–	0.63	–	0.63	–	0.63	–
nad3-r	0.77	–	0.77	–	0.77	–	0.77	–
nad4ex3-f*	–	0.37	–	0.37	–	0.37	–	0.37
nad6-r1	–	0.6	–	0.6	–	0.6	–	0.6
nad7ex1-f	2.05	–	–	–	2.05	–	–	–
orfX-f	0.4	–	0.4	–	0.4	–	0.4	–
orfX-r	–	3.05	–	3.05	–	3.05	–	–
	2.95	–	2.95	–	2.95	–	2.95	–
	1.54	–	1.54	–	1.54	–	1.54	–
	1.35	–	1.35	–	1.35	–	1.35	–
rps7-r1*	1.23	–	1.23	–	1.23	–	1.23	–
	–	1.0	–	1.0	–	1.0	–	1.0
	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24
rrn26-f1	1.13	–	1.13	–	1.13	–	1.13	–
rrn5-r	–	0.2	–	0.2	–	0.2	–	0.2

* Due to the displayed polymorphism the 0.37 kb product obtained with nad4ex3-f and the 0.24 kb product obtained with rps7-r1 are the only approved fragments of the artifactual series which includes products of 0.23–0.25, 0.37–0.39 and 0.51–0.53 kb (see text); since the 0.24 kb rps7-r1 product displayed quantitative polymorphism (depicted by the font size difference) – it is possibly an overlap of the weak monomorphic product obtained solely with the vectorette primer and the S-cytoplasm-specific product of the rps7-r1/vectorette pair

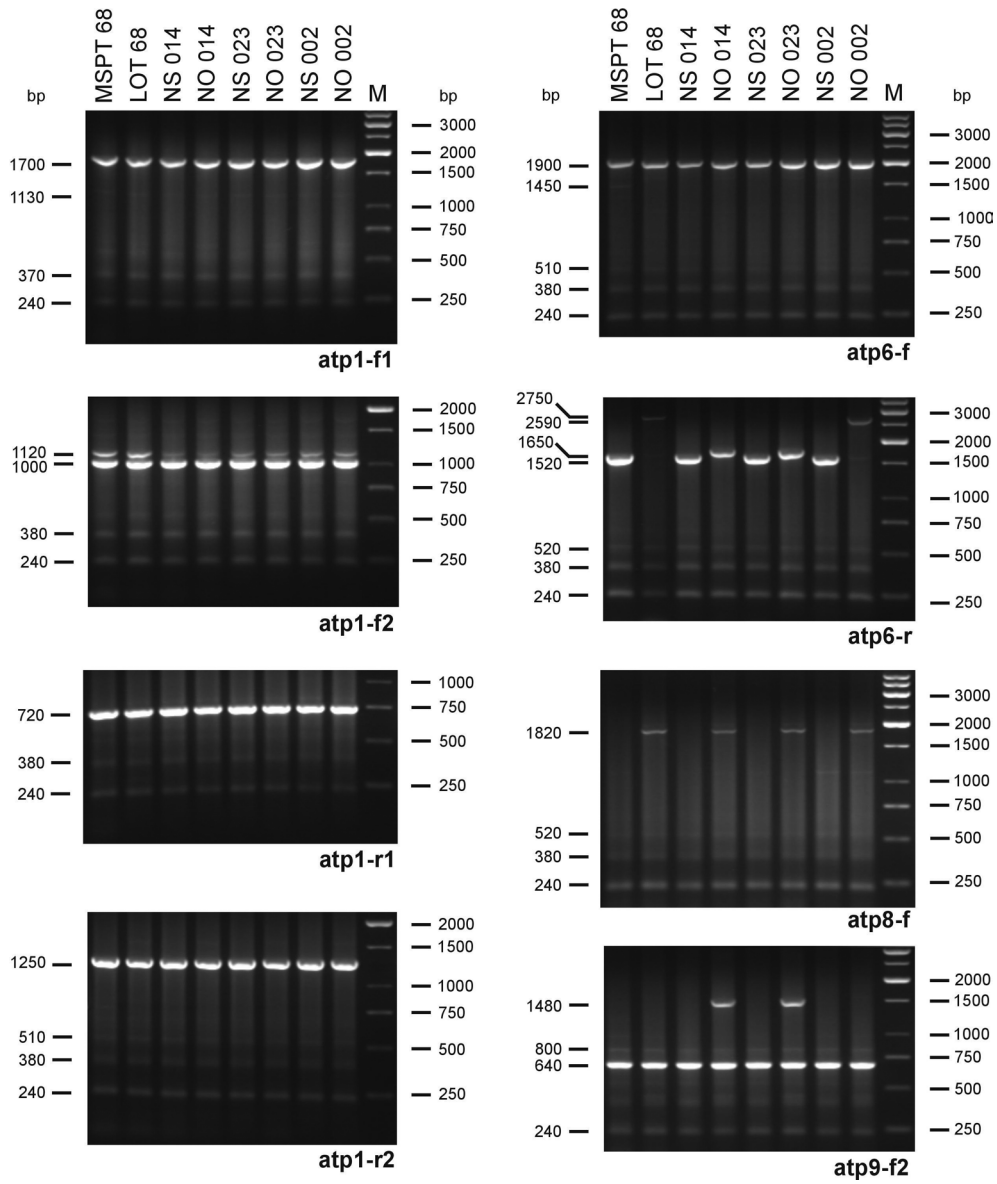


Fig. 3. Vectorette PCR profiles obtained for single plants from fodder beet (MSPT 68, LOT 68) and sugar beet (NS 014 – NO 002) breeding lines with the use of primers atp1-f1 – atp9-f2. Product sizes are indicated on the left. M – DNA size marker

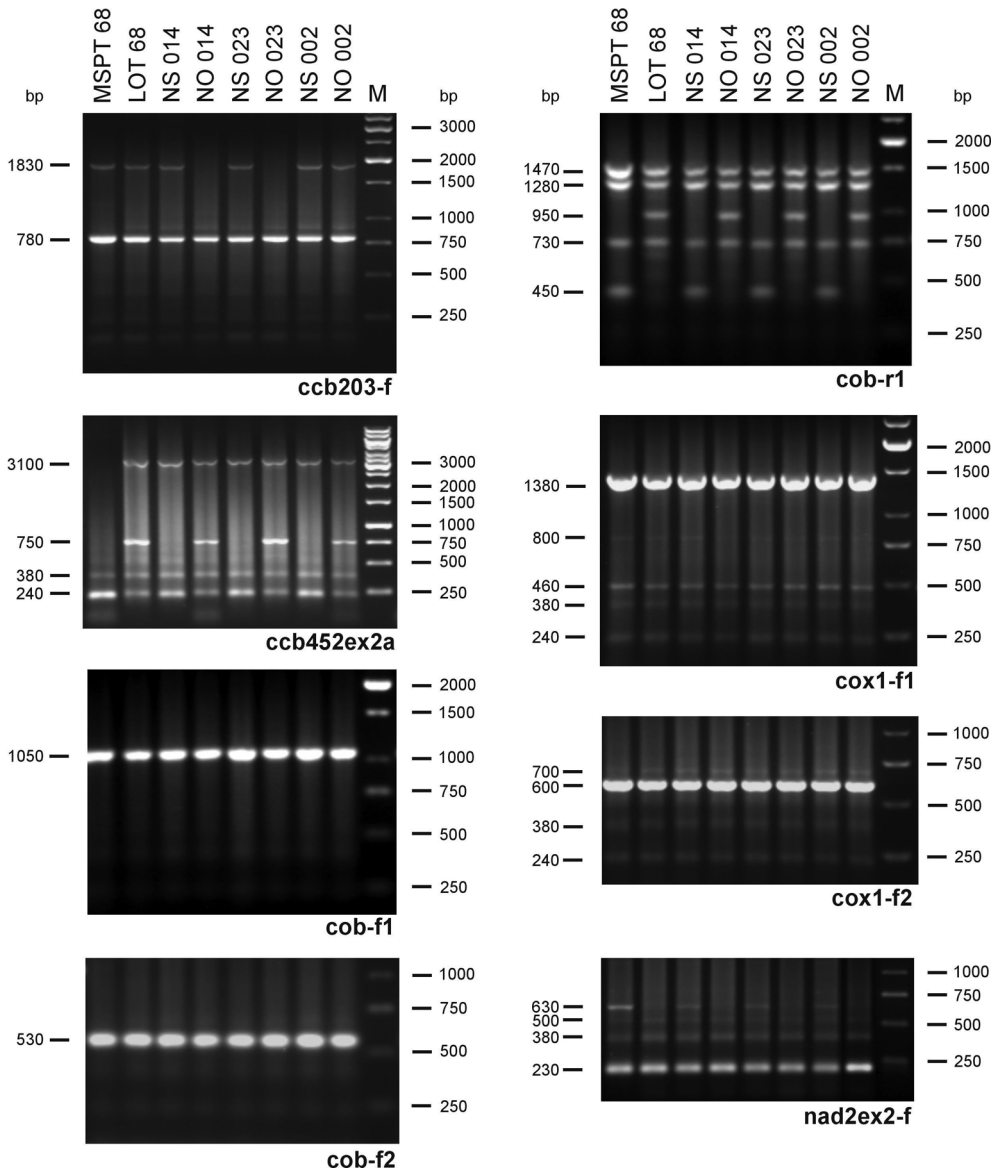


Fig. 4. Vectorette PCR profiles obtained for single plants from fodder beet (MSPT 68, LOT 68) and sugar beet (NS 014 – NO 002) breeding lines with the use of primers ccb203-f – nad2ex2-f. Product sizes are indicated on the left. M – DNA size marker

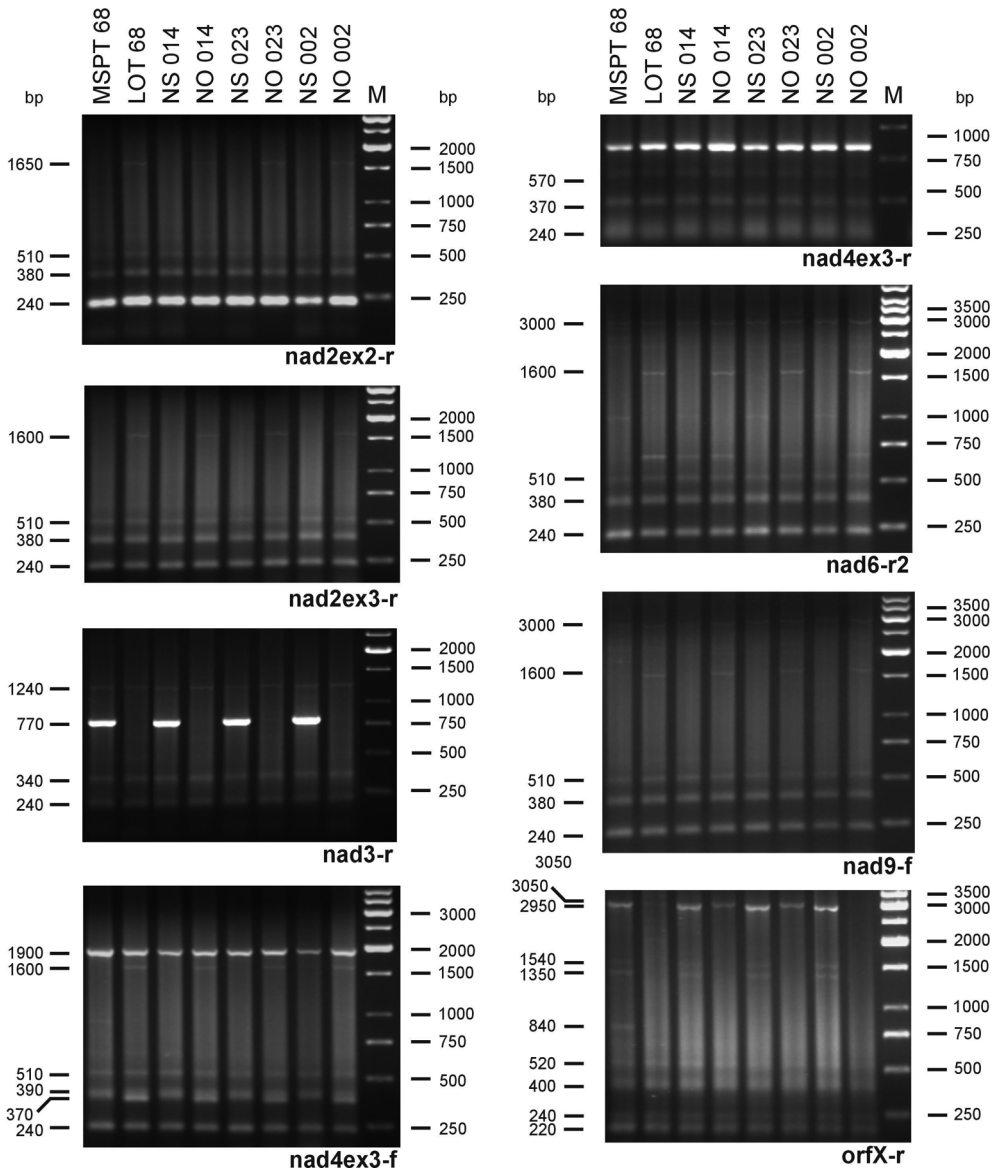


Fig. 5. Vectorette PCR profiles obtained for single plants from fodder beet (MSPT 68, LOT 68) and sugar beet (NS 014 – NO 002) breeding lines with the use of primers nad2ex2-r – orfX-r. Product sizes are indicated on the left. M – DNA size marker

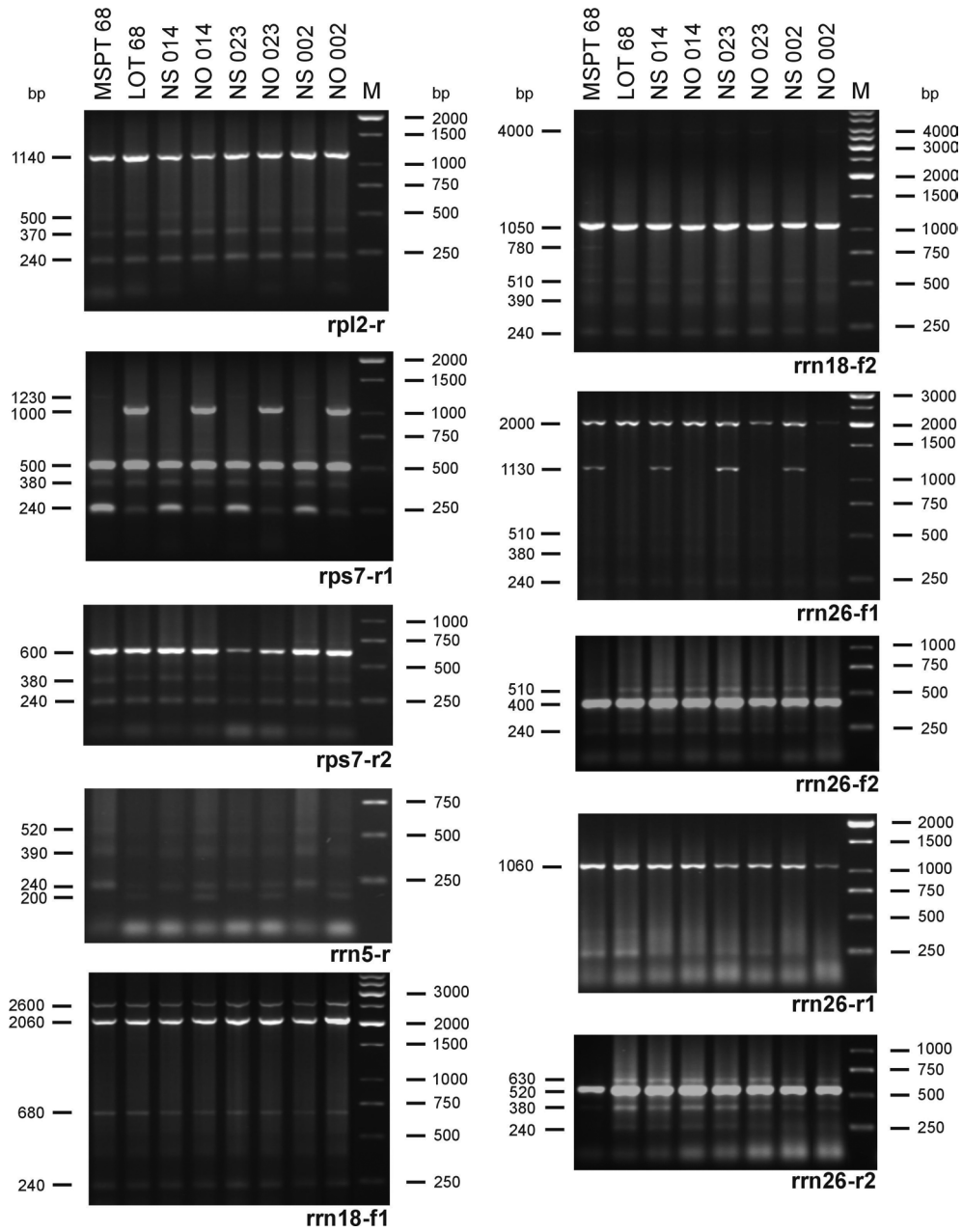


Fig. 6. Vectorette PCR profiles obtained for single plants from fodder beet (MSPT 68, LOT 68) and sugar beet (NS 014 – NO 002) breeding lines with the use of primers rpl2-r – rrn26-r2. Product sizes are indicated on the left. M – DNA size marker

The above summarization does not count the maintainer-specific 1.58–1.65 kb product which was observed for several UMPs and was likely produced only with the vectorette primer (see above). Moreover, one product – the 0.24 kb fragment obtained with the rps7-r1 UMP – displayed quantitative polymorphism, as its amplification was stronger in the MS lines. It is possible that this product represented superposition of the weak monomorphic artifactual fragment produced solely with the vectorette fragment (see above) and the S-cytoplasm-specific product of the rps7-r1/vectorette pair.

The fact that more than half of the polymorphic fragments displayed the plasmotype-specific distribution was reflected in the results of principal component analysis (PCA) which divided the analyzed accessions into two well-separated groups – one of which gathered the male-fertile O-lines (N-cytoplasmic) and the other their CMS counterparts (S-cytoplasmic) (Fig. 7). This differentiation was prominent along the first axis (principal component 1) which explained 69.6% of the total observed variation. Contrary to this effect the division between the fodder beet and sugar beet accessions was much less pronounced – it was prominent along the second axis (principal component 2) explaining only 9.9% of the observed variation.

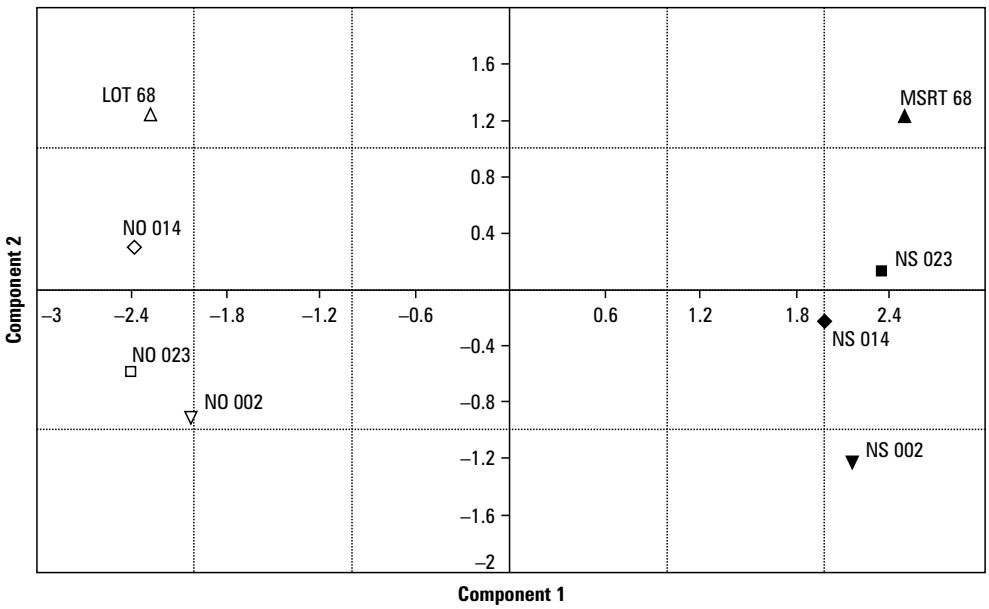


Fig. 7. Principal component analysis (PCA) for single plants from fodder beet (MSPT 68, LOT 68) and sugar beet (NS, NO) breeding lines based on distribution of vectorette markers. PC1 and PC2 explain 69.6 and 9.9% of the variance, respectively

The set of exploited UMPs included pairs of nested primers targeting the same mitochondrial gene. Such nested primer pairs were designed for genes: *atp1*, *atp9*, *cob*, *cox1*, *matR*, *nad6*, *rps7*, *rrn18* and *rrn26*. In the case of *rps7* only one nested pair was designed, for each of the remaining genes two nested primer pairs were laid down – one pair for one direction of synthesis (altogether there were 17 nested pairs). The nested character of UMPs is reflected in their names – designations f1 and f2 refer to the nested forward primers, while r1 and r2 – to the nested reverse primers. In principle, for complete mitochondrial loci, the nested UMPs should yield products with a size difference corresponding to the distance between the primer binding sites. However, for this assumption to be effective, the *Eco*RI site (*Eco*RI was used in the construction of the vectorette libraries) should not be present between the primer binding sites. In order to test the reliability of the vectorette PCR approach, products meeting the above criteria were searched among those obtained with the use of nested primers. For cutoff purposes two of the products obtained with nested UMPs were considered nested if their observed size difference fell within the interval between the theoretical size difference +10% and the theoretical size difference – 10%. Such products were identified for nine nested primer pairs representing six genes: *atp1*, *atp9*, *cob*, *cox1*, *rps7* and *rrn26* (Table 4). The majority of nested products were amplified at a high efficiency further confirming their specificity. According to expectations, the nested products were obtained mostly with the nested primer pairs devoid of the *Eco*RI site(s) between the primer binding sites (this was the case for eight pairs for which nested products were obtained). Nested products were not produced for eight UMP pairs – three of those had one or two *Eco*RI recognition sequences between the anchor sites, another four pairs were characterized by low amplification efficiency – at least for one primer of a nested pair.

Table 4. The summary of the vectorette PCR products obtained with nested universal mitochondrial primers (UMPs) for the set of two fodder beet and six sugar beet breeding lines

Pairs of nested universal mitochondrial primers (UMPs)	Expected size relation of the nested products*	Theoretical size difference of the nested products [bp]**	Observed products [kb] with nested products shown in bold***	Observed size difference of the nested products [kb]	*Eco*RI site between the UMP binding sites
atp1-f1	1 > 2	758	**1.7** 1.13	0.7	–
atp1-f2			1.12 **1.0**		
atp1-r1	1 < 2	561	**0.72**	0.53	–
atp1-r2			**1.25**		

atp9-f1	1 > 2	29	1.8 **0.67**	0.03	–
atp9-f2			1.48 0.8 **0.64**		
atp9-r1	1 > 2	28	0.41	n.a.	–
atp9-r2			0.84 0.78 0.6 0.41 0.2		
cob-f1	1 > 2	530	**1.05**	0.52	–
cob-f2			**0.53**		
cob-r1	1 < 2	441	**1.46 (a)** **1.28 (b)** **0.95 (c)** 0.73 0.45	0.48 (a) 0.42 (b) 0.45 (c)	–
cob-r2			2.3 **1.94 (a)** **1.7 (b)** **1.4 (c)**		
cox1-f1	1 > 2	848	**1.38** 0.8 0.46	0.78	–
cox1-f2			0.7 **0.6**		
cox1-r1	1 < 2	632	1.02 0.54 **0.43**	0.59	+
cox1-r2			**1.02** 0.54 0.43		
matR-f1	1 > 2	1025	0.93 0.54 0.22	n.a.	++
matR-f2			0.53		
matR-r1	1 < 2	1069	0.36 0.19	n.a.	++
matR-r2			0.35		
nad6-f1	1 > 2	158	–	n.a.	–
nad6-f2			–		

Table 4. cont.

Pairs of nested universal mitochondrial primers (UMPs)	Expected size relation of the nested products*	Theoretical size difference of the nested products [bp]**	Observed products [kb] with nested products shown in bold***	Observed size difference of the nested products [kb]	*Eco*RI site between the UMP binding sites
nad6-r1	1 > 2	111	0.6	n.a.	–
nad6-r2			3.0 1.6		
rps7-r1	1 < 2	131/122****	1.23 1.0 **0.49** 0.24	0.13	–
rps7-r2			**0.62**		
rrn18-f1	1 > 2	786	2.6 2.06 0.68	n.a.	–
rrn18-f2			4.0 1.05 0.78		
rrn18-r1	1 < 2	898	0.66 0.45	n.a.	–
rrn18-r2			2.05 0.75		
rrn26-f1	1 > 2	1592	**2.0** 1.13	1.6	–
rrn26-f2			**0.4**		
rrn26-r1	1 < 2	1608	1.06	n.a.	+
rrn26-r2			0.63 0.52		

* e.g. for primers atp1-f1 and atp1-f2 *1 > 2* means that the product obtained with primer atp1-f1 should be larger than the product obtained with primer atp1-f2

** calculated on the basis of sequence data

*** two products (synthesized with different primers of the same nested pair) were considered nested if their observed size difference fell within the interval between the theoretical size difference + 10% and the theoretical size difference – 10%

**** for sugar and sea beet, respectively

– depending on the column either no specific products or no *Eco*RI site

\+ one *Eco*RI site

++ two *Eco*RI sites

n.a. not applicable

a, b, c differentiation between three pairs of nested products obtained for the same pair of nested primers

4.3. Sequence analysis of the identified RAPD and vectorette markers

From the marker screen described above, a selection of the identified cytoplasm-specific PCR products was subjected to cloning in a plasmid vector, and sequencing. Sequence data were produced for four RAPD markers, which were obtained with primers A-08, B-01 and B-05 (Table 2). Further on in the text these markers are referred to as N_A-08, N_B-05, S_B-01_c3 and S_B-01_c29. Designations c3 and c29 refer to the longer and shorter S-cytoplasm-specific B-01 product, respectively. Sequence data were also obtained for nine vectorette markers amplified using the following universal mitochondrial primers (UMPs): atp6-r, atp8-f, atp9-f2, cob-r1, nad3-r, rps7-r1 and rrn26-f1. Further on in the text the respective marker sequences are referred to as: N_atp6-r, S_atp6-r, N-atp8-f, N_atp9-f2 (with two versions of different lengths designated c38 and c39), N_cob-r1, S_cob-r1, S_nad3-r, N_rps7-r1 (only the N-cytoplasm-specific product was cloned and sequenced) and S_rrn26-f1.

RAPD marker N_A-08

A very similar mitochondrial sequence was found in three accessions of *Beta vulgaris* subsp. *maritima* – this homology matched the region extending from nucleotide 5 to 251, over this distance the differences were limited to one substitution and one or two indels. The homologous sequences from *B. v. maritima* were located 0.49 kb upstream of *orf105e*. Another homology covering the majority of the marker sequence was found in mtDNA of *Asclepias syriaca* – within the RAPD product it corresponded to nucleotides 101–230. On the right terminus of this segment, 29 nucleotides (positions 202–230) showed also duplicated homology in the mitochondrial genome of *Vicia faba*. The Blast search returned yet another short plant mitochondrial homology – nucleotides 41–76 were similar to a stretch from mtDNA of TK81-MS sugar beet. This was the only match of this origin and moreover, no homology to TK81-O mtDNA was found – therefore, as an entity, the N_A-08 marker sequence was not present in the reference sugar beet mitochondrial genomes. The homologous sequences from *Asclepias syriaca*, *Vicia faba* and TK81-MS sugar beet occupied genomic locations which were different from each other and from that exhibited by *Beta vulgaris* subsp. *maritima* (Fig. 8).

gtgacgtaggGTCGGTAGACACAAGCCGTGACGTAGGGTTgaagactcaagtctagtagagacgccgggaagaag
gaagataactcaagaagcagtctttTAGGCCCAGGTAGCTTGCTTGTAATCTCAGTCAAGCATATTCGCAATCCT
TGGGCAAGGAGGTAGTTTACTACTCGACCAAAGAAAGACCGTCAAATGGGGCTTATATACGCAACCCACCTTCTC
ATCTCTTATTCCTTTTCTGACCTACGTC

xxxxx – sequence of the A-08 primer
xxxxx – reverse complement sequence of the A-08 primer, incomplete – two terminal nucleotides missing (ac)
xxxxx – homology to TK81-MS mtDNA
xxxxx – homology to *Asclepias syriaca* mtDNA
xxxxx – homology to *Vicia faba* mtDNA

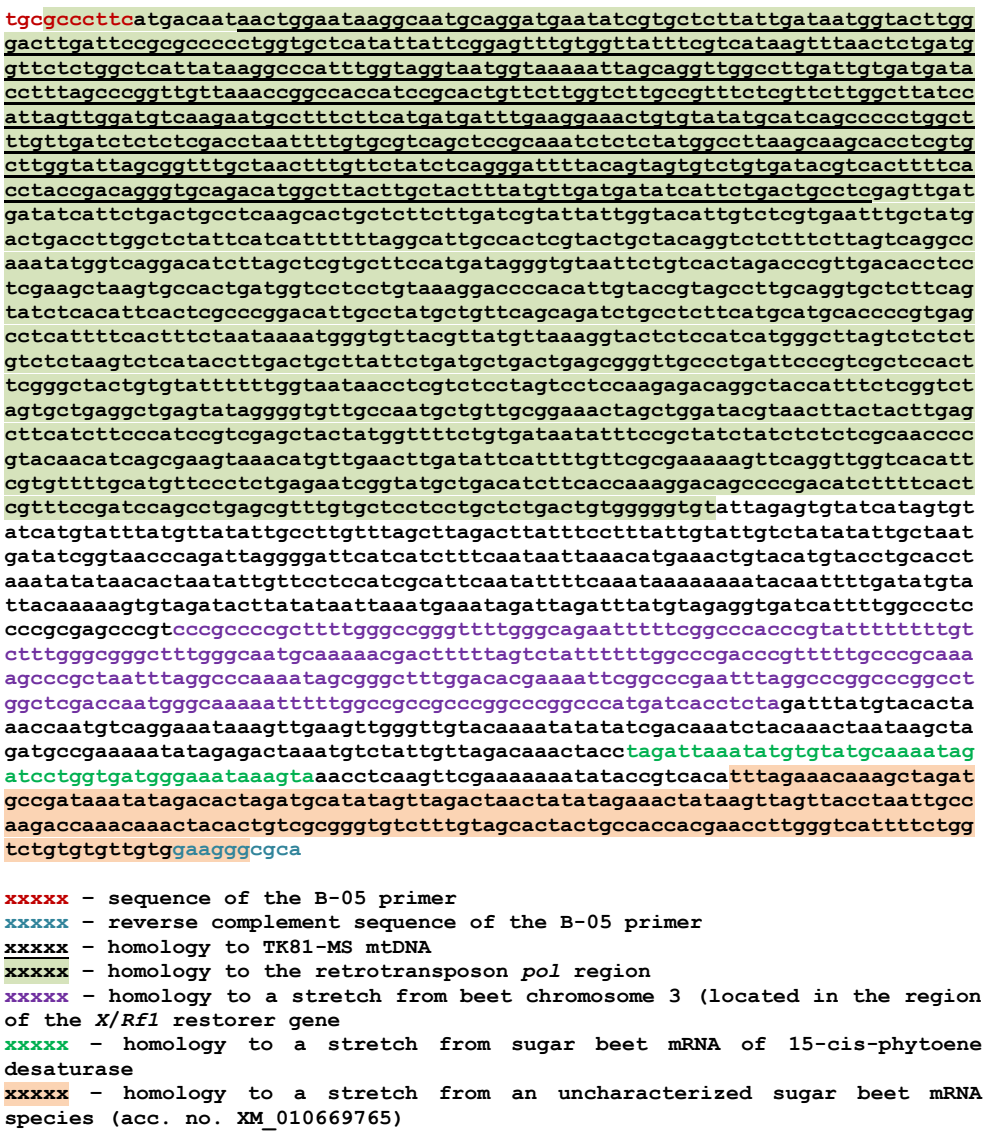
Fig. 8. Sequence features of the N_A-08 RAPD marker

RAPD marker N_B-05

The left part of the sequence, extending for approx. 1500 nt from the left terminus, showed homology to the *pol* region from the plant Ty1-Copia-like retrotransposons. This region contained portions representing both the reverse transcriptase (RT) and RNase H-coding sequences. Homology between the marker and the sugar beet retrotransposon copies was rather low – 68%. The left segment in this *pol*-like part of N_B-05 was also homologous to a two-copy fragment of mtDNA from TK81-MS and a single copy mitochondrial fragment from the male-sterile genotype G of *Beta vulgaris* subsp. *maritima*. This marker segment extended from nucleotide 19 to 592 and represented the reverse transcriptase portion of *pol*. Homology between the marker and this mitochondrial RT-like sequence was also rather low – 70%. In the right terminal region the N_B-05 marker displayed three short homologies to sugar beet nuclear DNA:

- positions 1889–2160 – homology to a stretch located in the region of the *X* (*Rf1*) restorer gene on chromosome 3 [Matsuhira et al. 2012],
- positions 2299–2349 – homology to a stretch from mRNA coding for 15-cis-phytoene desaturase,
- positions 2382–2569 – homology to a stretch from an uncharacterized mRNA species.

Homology between these stretches and the respective marker sequences was at a level of 84–85% (Fig. 9).

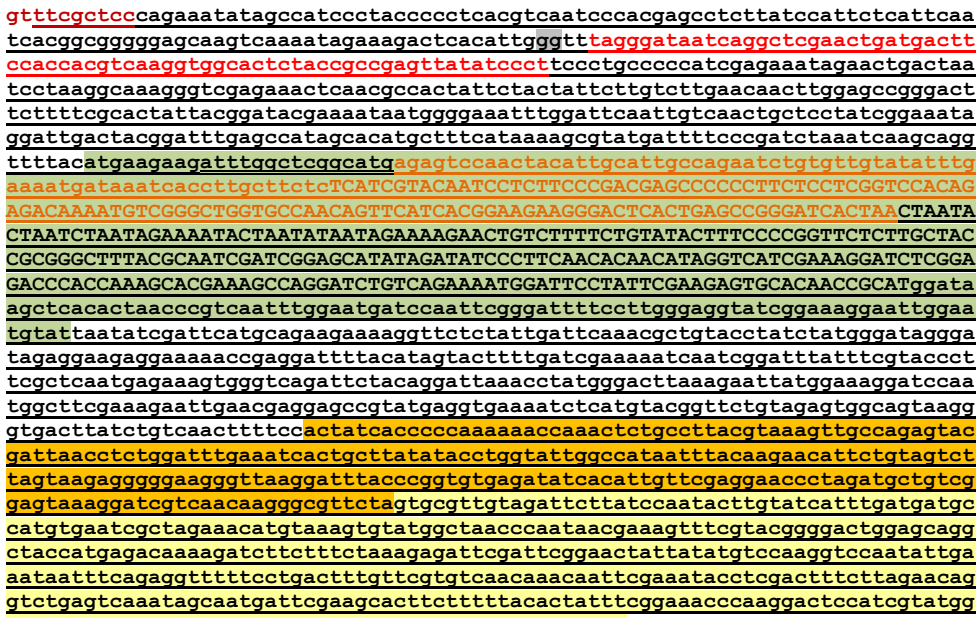


xxxxx – sequence of the B-05 primer
xxxxx – reverse complement sequence of the B-05 primer
xxxxx – homology to TK81-MS mtDNA
xxxxx – homology to the retrotransposon *pol* region
xxxxx – homology to a stretch from beet chromosome 3 (located in the region of the *X/Rf1* restorer gene
xxxxx – homology to a stretch from sugar beet mRNA of 15-cis-phytoene desaturase
xxxxx – homology to a stretch from an uncharacterized sugar beet mRNA species (acc. no. XM_010669765)

Fig. 9. Sequence features of the N_B-05 RAPD marker

RAPD marker S_B-01_c29

The longest homologous sequence block was found in mtDNA of TK81-MS between *orf136b* and *trnW* (acc. no. BA000024, nucleotides 463119–465040). It covered almost the whole marker sequence – nucleotides 3–1923. It contained one annotated sequence – *orf174*. Within this region eight sequence differences were noted – one indel and seven substitutions. In the mtDNA of TK81-O sugar beet, *Beta vulgaris* subsp. *maritima* and *Beta macrocarpa*, this marker sequence was represented by two sequence blocks which corresponded to marker regions 3–480 and 466–1923. Altogether, fourteen substitutions differentiated the analyzed marker sequence from the above mentioned Blast hits (which were identical) – three of these substitutions were located within the shorter region of homology and eleven within the longer region. Moreover, within the shorter homology the marker sequence was devoid of a 28-nt stretch which was present in the returned hits. In the reference mitochondrial genomes mentioned above, the shorter block extended between *orf106a*/*orf113a*/*orf496a* (depending on the beet accession and apparently due to the fact that the sequences located upstream of the shorter homologous block diverged in the *Beta vulgaris* germplasm) and *orf146*. This block contained gene coding for valine tRNA (*trnV*). In the reference mtDNAs the longer homologous block was located between *orf146* (its 3' end was represented in the marker sequence) and *rps7*. This block contained *orf114a*. *TrnV* and *rps7* are of chloroplast origin and accordingly both homologous blocks were also found in the *Beta vulgaris* chloroplast genome (acc. nos. AB032426, EF534108) – they were located between the 16S rRNA gene (*rrn16*) and exon 3 of *rps12* (with *rps7* further downstream). Their mutual arrangement was the same as the one observed in mtDNA of TK81-O sugar beet, *Beta vulgaris* subsp. *maritima* and *Beta macrocarpa*. Upon comparison with the chloroplast genome, the marker sequence showed 18/19 polymorphisms (depending on the cpDNA record) – 6 and 12/13 within the shorter and longer homologous region, respectively. The plastid genome annotations indicated the presence of *trnV* (included also in the mitochondrial annotations – see above) and exon 2 of *rps12* within the shorter and the longer block, respectively. These features as well as extensive chloroplast homologies from other plant species indicated the chloroplast origin of the analyzed marker sequence (Fig. 10).



xxxxx – sequence of the B-01 primer
xxxxx – reverse complement sequence of the B-01 primer
xxxxx – TK81-O/*B. v. maritima*/*B. macrocarpa* mitochondrial homology and *B. vulgaris* plastid homology
xxxxx – overlap between blocks of TK81-O/*B. v. maritima*/*B. macrocarpa* mitochondrial homology and *B. vulgaris* plastid homology
xxxxx – homology to *trnV*
xxxxx – homology to *orf174*
xxxxx – *rps12* exon 2
xxxxx – homology to *rps12* intron 2 (3′ part is missing)
xxxxx – *orf146* (5′ part is missing)
XXXXX – *orf114a*
xxxxx – nucleotides flanking a minor insertion present in the O-type mtDNA (ttcgaagatcttagttgctagactaagg)

Fig. 10. Sequence features of the S_B-01_c29 RAPD marker

RAPD marker S_B-01_c3

This marker showed two regions of homology to the S_B-01_c29 RAPD fragment (described above). They corresponded to the two homologous blocks – the shorter and the longer – distinguished upon comparison of the S_B-01_c29 marker with mtDNAs from TK81-O sugar beet, *Beta vulgaris* subsp. *maritima* and *Beta macrocarpa*. The major difference between S_B-01_c3 and S_B-01_c29 was the 414 bp region which separated the mentioned homologous blocks in the former and was missing in the latter. Thus, this indel was responsible for the size difference between

the two RAPD markers. In S_B-01_c3 the 414 bp region at its right terminus carried a stretch of 15 nt which was an imperfect direct repeat of the 15 right terminal nucleotides of the shorter homologous block – therefore, it is likely that recombination between these direct repeats led to deletion of the 414 nt region from the sequence represented by marker S_B-01_c29. In addition to this indel polymorphism, the marker sequences were differentiated by 17 SNPs. Accordingly with this description, the highest-scoring Blast hits were delivered by two *Beta vulgaris* cpDNA records (acc. nos. AB032426, EF534108). In sugar beet cpDNA the S_B-01_c3-homologous region extended between the *rrn16* gene and exon 3 of *rps12* – it included the *trnV* gene, exon 2 of *rps12* as well as the 5' part of *rps12* intron 2 (Fig. 11).

gtttcgctccagaaatatagccatccctacccccctcacgtcaatcccacgagcctcttatccattctcattcaa
tcacggcgggggagcaagtcaaaatagaaaaactgacattgggtttagggataatcaggctcgaactgatgactt
ccaccacgtcaaggtgacactctaccgctgagttatatccctttccctgcccccatcgagaaatagaactgactaa
tcctaaggcaaagggtcgagaaactcaacgccactattctactattcttgtcttgaacaacttggagccgagact
tcttttcgcactattacggatacgaaaataatggggaaatttggattcaattgtcaactgctcctatcggaaata
ggattgactacggatttgagccatagcacatgctttcataaaatcgtacgattttcccgatctaaatcaagcagg
ttttacatgaagaagatttggctcggcatgttctatttgatataggtaggagaagaacccgactcggtattcaaa
aaaaaaatagaggaagcagaaccaagtcaagatgatacggatcaacccctttcttcttgcgacaaagatcttaccc
tttccaaaggaagttccatctcttttccatttccattcaagagttcttatgtgtttccacgccccttgaaaccc
cgaaaaatggacaaattccttttcttaggaatacataccgcactcgtcactccaaaaaggataatggtaacccca
ccattaaccacttcatttatgaatttcatagtaatagaaatacatgtcctaccgagacagaatttggaacttgct
atcctcttgcttagcaggcaaagacttacctccgtggaaaggatgattcattccattcggatcgacatgagagtc
caactacattgcattgccagaatctgtgttgtatatttgaaaatgataaatcaccttgcttctctcatcgtacaa
tcctcttcccgacgagccccccttctcctcggtccacagagacaaaatgtcgggctggtgccaacagttcatcac
ggaagaagggactcactgagccgggatcactaactaatactaatctaatagaaaatactaatataatagaaaaga
actgtcttttctgtatactttccccggttctcttgctaccgcggactttacgcaatcgatcggatcatatagata
tcccttcaacacaacataggtcatcgaaaggatctcggagacccaccaaagcacgaaagccaggatctttcagaa
aatggattcctattcgaagagtgcacaaccgcatggataagctcacactaacccgtcaatttggaatgatccaat
tcgggattttccttgggaggtatcggaaaggaattggaatgtaataatatcgattcatgcagaagaaaaggttct
ctattgattcaaacgctgtacctatctatgggatagggatagaggaagaggaaaaaccgaggattttacatagta
cttttgatcgaaaaatcaatcggatttatttcgtacccttcgctcaatgagaaagtgggtcagattctacaggat
taaacctatgggacttaaagaattatggaaaggatccaatggcttcgaaagaattgaacgaggagccgtatgagg
tgaaaatctcatgtacggttttgtagagtggcagtaagggtgacttatctgtcaacttttccactatcaccccca
aaaaaccaaactctgccttacgtaaagttgccagagtacgattaacctctggatttgaaatcactgcttatatac
ctggtattggccataatttacaagaacattctgtagtcttagtaagagggggaagggttaaggatttacccggtg
tgagatatcacattgttcgaggaaccctagatgctgtcggagtaaaggatcgtcaacaagggcgttctagtgcgt
tgtagattcttatccaagacttgtatcatttgatgatgccatgtgaatcgctagaaacatgtaaagtgtatggct
aacccaataacgaaagtttcgtaaggggactggagcaggctaccatgagacaaaggatcttctttctaaagagat
tcgattcggaactattatatgtccaaggtccaatattgaaagaatttcagaggtttttcctgactttgttcgtgt
caacaaacaattcgaaatacctcgactttcttagaacaggtctgagtcaaatagcaatgattcgaagcacttctt
tttacactatttcggaaacccaaggactccatcgtatggatatggaaaatacaggatttccaatcctagcaggaa
aaggagcgaaac

xxxxx – sequence of the B-01 primer
xxxxx – reverse complement sequence of the B-01 primer
xxxxx – homology to the S_B-01_c29 RAPD marker
xxxxx – imperfect direct repeats
xxxxx – *trnV*
xxxxx – *rps12* exon 2
xxxxx – homology to *rps12* intron 2 (3' part is missing)

Fig. 11. Sequence features of the S_B-01_c3 RAPD marker

Vectorette marker N_atp6-r

An identical sequence (100% homology – identity) was present in the mitochondrial genome of *Beta macrocarpa* (acc. no. FQ378026, nucleotides 4443–6157). In the sea beet (*Beta vulgaris* subsp. *maritima*) male-fertile genotypes A and B as well as in sugar beet TK81-O the homologous sequences were split into two blocks – in the analyzed vectorette marker they corresponded to nucleotide positions 1–927 and 837–1715. The first block was non-coding (339568–340494) while the second (200951–201829) contained the *atp6* ORF without 15 3'-terminal nucleotides (the nucleotide coordinates are given for TK81-O). Three homologies were found upon comparison with TK81-MS – they corresponded to sequence stretches: 1–226, 837–932 and 971–1715. In TK81-MS mtDNA (acc. no. BA000024) the first stretch was duplicated (87619–87844, 280135–280360) and located between *nad2* exon 5 and *orf143*. The second stretch represented a repeat sequence found in slightly

tcctttcttccctctcgacagttagctagaggaagcttcgcttctcgtacttgaactgctgtagacaagaagtca
aagccctagggaaaaagcgtagcgcttattgtttaacaggtacgtagcctctttcgagaggcgaagaatagctat
ctttcagaaagaagaagagcagcaaaccttttagtactaccagactcgaggcgaaaagtggatttaccagaaag
aatcaataggagggtgaaagttgcttcgtccttagcccggaactcgttcctcacgtagactgaacttcttctttc
ttagcgcctcgtttttagtggaagacgactgaactcgctccgctttgtaggcaaggggcaaagccaaccacagcc
atcacctttcagaagaggacatgctcttagcagatttacagcataaccgccctagcttagctccttctttctaa
agcgcaaaagctaaagcgctacatagagtttattatagggcgaggcttcctaaatcccatttagagggctggta
tcccgctctatacgattgtcagctaagtaaatagatgatgaggcaaatgaagaaagaaagaaagcgcttttaaag
tcgtacacgaaatcaggcttcaccttagctactaagaggctgcctctggcaacgactttcatttttccaaagatt
tcattcttaccggcatacgctaggagcatcactctctagctaagagctttctttccttcttccgcagttagggta
gcagtcctctttgtttcctttctccctcactacatctgtttctatctcaatatcataagagaagaaagacaagtc
gctgtcttcatatgcccaaagactcccatgcctttcttggtcggaccaacccaaccggtgatttccgacaagtct
ttcttcatttttgagcaagaagcggaactacagggatgaaatgaaaagtgtttattACGATTACGCCCAACAGCC
CACTTGAGCAATTTTCCATTCTCCCATTGATTCCTATGAAAATAGGAAACTTGTATTTCTCATTCACAAATCCAT
CTTTGTTTATGCTGCTAACTCTCAGTTTGGTCCTACTTCTGCTTCATTTTGTTACTAAAAAGGGAGGAGGAAACT
CAGTACCAAATGTTTGGCAATCCTTGGTAGAGCTTATTTATGATTTCGTGCTGAACCTGGTAAACGAACAAATAG
GTGGTCTTTCCGGAAATGTTAAACAAAAGTTTTTCCCTTGCATCTTGGTCACTTTTACTTTTTTGTTATTTCGTA
ATCTCCAGGGTATGATACCCTATAGCTTTACAGTTACAAGTCATTTTCTCATTACTTTGGGTCTTTCATTTTCCA
TTTTTATTGGCATTACTATAGTGGGATTTCAAAGAAATGGGCTTCATTTTTTAAGCTTCTCATTACCTGCAGGAG
TCCCGCTGCCGTTAGCACCTTTTTTAGTACTCCTTGAGCTAATCCCTCATTGTTTTCGCGCATTAAGCTCAGGAA
TACGTTTATTTGCTAATATGATGGCCGGTCATAGTTCAGTAAAGATTTTAAGTGGGTTCGCTTGGACTATGCTAT
GTATGAATGATCTTTTATATTTCATAGGAGATCTTGGTCCTTTATTTATAGTTCTTGCATTAACCGGTCTTGAAT
TAGGTGTAGCTATATTACAAGCTCATGTTTTTACGATCTTAATCTGTATTTACTTGAATGATGCT

xxxxx – reverse complement sequence of the atp6-r UMP
xxxxx – homology to TK81-MS mtDNA
xxxxx – homology to a sequence often located approx. 20 bp upstream of plant *nad9* ORFs
XXXXX – *atp6* ORF (incomplete - 3′ terminus missing)

Fig. 12. Sequence features of the N_atp6-r vectorette marker

different variants in several localities (13897–13993, 421285–421360, 441857–441930, 469963–470045[1]). Interestingly, variants of this repeat were also present in mtDNAs of other plants – both as a unique sequence (e.g. in *Silene latifolia*) or as a repeat (e.g. in *Hevea brasiliensis* or in *Rhazya stricta*). In the latter case they also (as in beet) had diverse genomic contexts – among them this sequence stretch could be found approx. 20 bp upstream of the *nad9* ORF. This location was also characteristic for species with a single copy of this sequence indicating the primary character of this location and its functional importance. The last stretch corresponded to nucleotides 422551–423295 – the canonical *atp6* part of the composite *preSatp6-Satp6* locus (Fig. 12).

Vectorette marker N_atp8-f

The whole sequence showed close to 100% homology (identity) with a fragment of the TK81-O mtDNA (acc. no. BA000009, nucleotides 263978–262149) – there were only two substitutions and three indels. The marker contained a 5' part of the *rps12* ORF, the *nad3* ORF and a 3' part of *orf100e*. Almost identical hits were returned from the mitochondrial genomes of *Beta vulgaris* subsp. *maritima* (genotypes A, B and E) and *Beta macrocarpa*. A sequence extending from nucleotide 1 to 974 was also represented in mtDNA of TK81-MS (acc. no. BA000024), however, in a few stretches, their nucleotide coordinates were the following: 335542–334679, 423540–423386, 334565–334504, 423954–423897, 234172–234126, 382522–382568 (given in the order of growing Expect value). The *nad3* ORF was preceded by a conservative stretch of 60 nucleotides – its genomic location in other plants was the same as in this beet vectorette marker (Fig. 13).

Vectorette marker N_atp9-f2_c38

In this vectorette fragment the stretch of the first 158 nt displayed homology to the pEV1 macrosatellite from the beet nuclear genome [Schmidt et al. 1991, Santoni and Berville 1992]. The last six nucleotides of this stretch represented an *Eco*RI site (gaattc). The sequence located below this restriction site showed plant mitochondrial homologies. The above mentioned sequence features as well as the use of *Eco*RI for library construction jointly indicate that the analyzed vectorette marker is chimeric and likely arose from ligation of the macrosatellite stretch in between the vectorette cassette and the fragment of mitochondrial origin described below.

[1] Between: *orf177* and *atp1*, *orf214* and composite *atp6*, *rps4* and *rpl5*, *nad9* and *orf224*, respectively

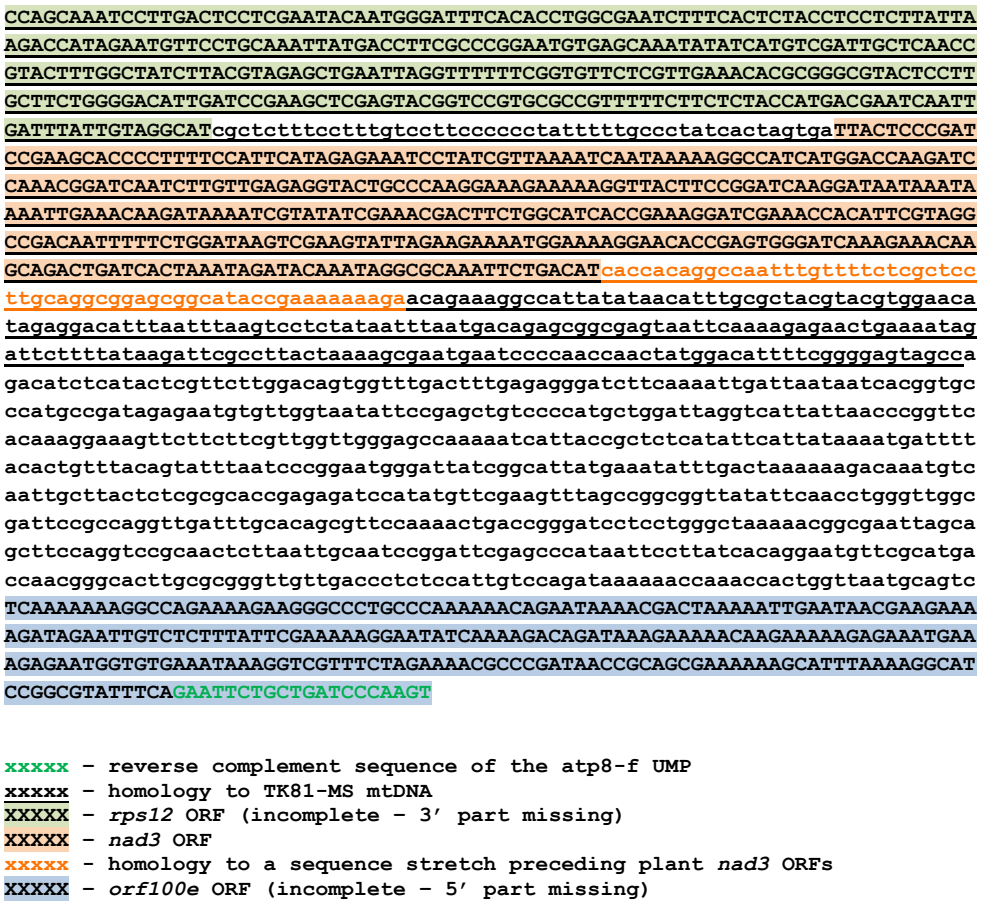


Fig. 13. Sequence features of the N_atp8-f vectorette marker

The largest homologous blocks were found in mtDNA of *Beta vulgaris* subsp. *maritima* male-sterile genotypes E and G (acc nos. FQ014226 and FP885871) (for genotype E nucleotides 139269–140766) – they corresponded to nucleotide positions 153/154–1650. The majority of these homologous sequences was non-coding, only the atp9-f2 primer terminus contained the 3' terminus of *orf105e*. The following sequence stretches showed homology to mtDNA of TK81-MS (acc. no. BA000024): 153–413 (its TK81-MS counterpart was duplicated: nt pos. 93226–93486, 285742–286002; the last approx. 40 nt – in Fig. 14 marked with dark green shading – showed homologies in four other localities), 687–731 (187102–187146), 1401/1404/1405–1441/1442/1445 (three copies: 164263–164299, 10745–10790, 478075–478113). Upon comparison of this marker with TK81-O mtDNA the pattern of homologies was similar – the most notable difference was that the longest homologous TK81-O stretch

was much shorter and in the marker sequence it corresponded to nucleotides 153–224.

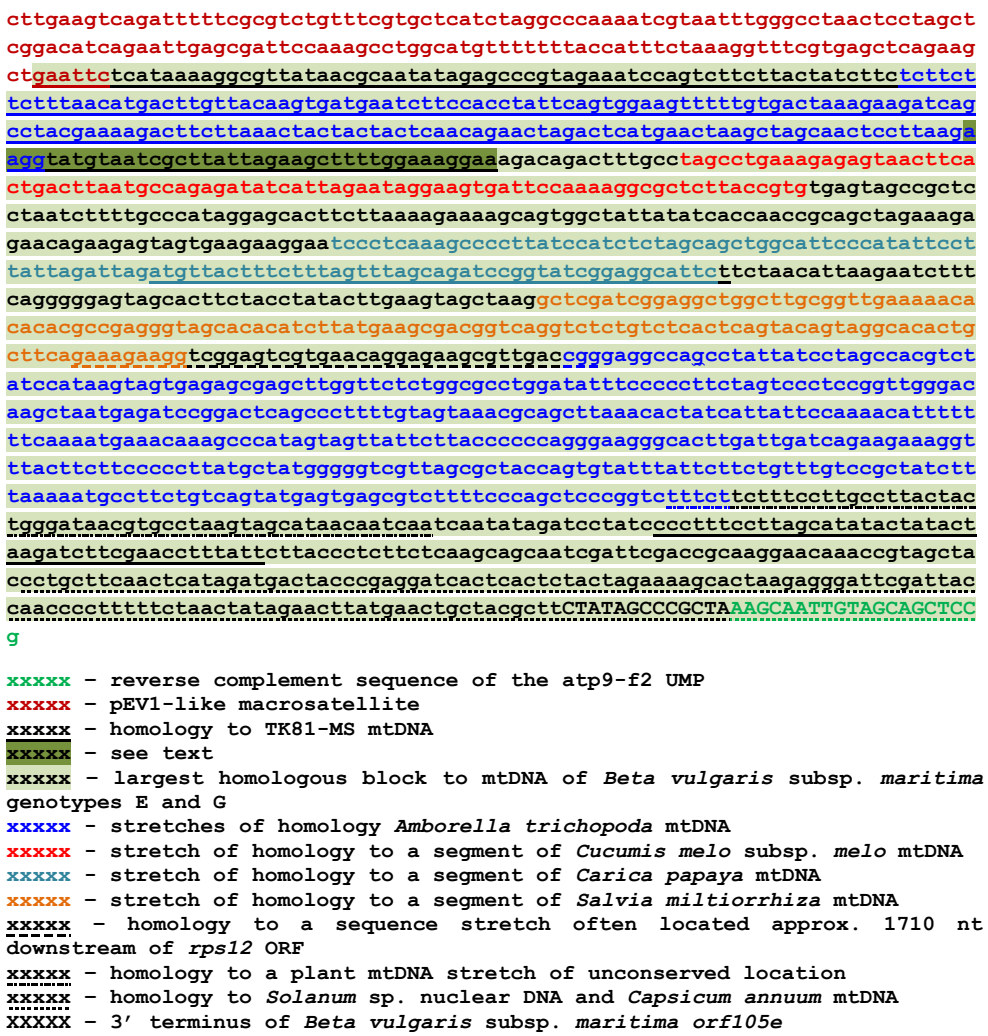


Fig. 14. Sequence features of the N_atp9-f2_c38 vectorette marker

This marker sequence was also abundant in plant mitochondrial homologies from outwith *Beta* sp. Nucleotides 220–378 displayed homology to a stretch of *Amborella trichopoda* mtDNA. Shorter homologies matching within this region were also found in the mitochondria of *Vaccinium macrocarpon* and *Hevea brasiliensis*. Nucleotides 428–512 showed homology to a stretch of *Cucumis melo* subsp. *melo* mtDNA located between the gene coding for tRNA-Asn(GTT) and the *rps10* ORF.

The internal part of this sequence – nucleotides 431–505 – was also similar to a *Salvia miltiorrhiza* mtDNA stretch located between *orf178a* and *orf137*. Nucleotides 626–730 showed homology to a *Carica papaya* mtDNA fragment located between the *atp6* ORF and gene *trnS(GGA)* of chloroplast origin. Nucleotides 792–914 were homologous to a *Salvia miltiorrhiza* mtDNA stretch located between *orf116d* and *orf100c*. Nucleotides 906–946 were similar to a sequence stretch in some species located approx. 1710 nt downstream of the *rps12* ORF. For nucleotides 944–1331 homology was found in mtDNA of *Amborella trichopoda*, a slightly shorter homologous block was detected in *Mimulus guttatus* and, moreover – *Butomus umbellatus* mtDNA had a stretch similar to the terminal part of this sequence (nucleotides 1226–1319). No conservation was found in the genomic location of these homologous sequences. Nucleotides 1327–1383 displayed homology to a plant mtDNA stretch of unconserved location.

Additional homologies were found for the sequence stretch located at the atp9-f2 primer terminus – for nucleotides 1502–1650. The homologous sequence was found in the nuclear genome of *Solanum* sp. (in chromosomes 6 and 8 of *Solanum pennellii* and in chromosome 3 of *Solanum lycopersicum*) as well as in mtDNA of *Capsicum annuum*. In the latter case the sequence location was between *orf102a* and *orf108a*. Both *Solanum* sp. and *Capsicum annuum* belong to the family *Solanaceae* (Fig. 14, Table 5).

Table 5. Mitochondrial non-*Beta* sp. homologies of the N_atp9-f2_c38 vectorette marker

Marker region*	Species displaying homology	Location of homologous sequences	Copy number
220 – 378	*Amborella trichopoda, Vaccinium macrocarpon***, Hevea brasiliensis***	species-specific	1
428 – 512	*Cucumis melo, Salvia miltiorrhiza***	species-specific	1
626 – 730	*Carica papaya*	between *atp6* and (cp-origin) *trnS(GGA)*	1
792 – 914	*Salvia miltiorrhiza*	between *orf116d* and *orf100c*	1
906 – 946	many	approx. 1710 nt downstream of *rps12*	1 or 4***
944 – 1331	*Amborella trichopoda*, *Mimulus guttatus***, *Butomus umbellatus***	species-specific	1
1327 – 1383	many	species-specific	1
1502 – 1650	*Capsicum annuum*	between *orf102a* and *orf108a*	1

* first nucleotide – last nucleotide

** shorter homologies matching within the specified marker region

*** only in *Hevea brasiliensis*, in this species there were two versions of this stretch – they had a slightly different sequence and a different genomic context – each version was present in two copies due to duplication of large (tens kb) genome segments

Vectorette marker N_atp9-f2_c39

The almost entire sequence of this marker was represented in the N_atp9-f2_c38 marker – there were only three substitutions differentiating the respective regions of these markers. However, in contrast to N_atp9-f2_c38, the N_atp9-f2_c39 marker was devoid of the macrosatellite sequence at the vectorette primer terminus. In addition, the N_atp9-f2_c39 marker displayed a 24 nt extension at the custom primer terminus. This extension had likely resulted from the use of an alternative prim-

tcataaaaggcgttataacgcaatatagagcccgtagaaatccagtcttcttactatcttctcttcttctttaac
atgacttgttacaagtgatgaatcttccacctattcagtggaagtttttgtgactaaagaagatcagcctacgaa
aagacttcttaaactactactactcaacagaactagactcataaactaagctagcaactccttaagaaggtatgt
aatcgcttattagaagcttttggaaaggaaagacagacatttgcctagcctgaaagagagtaacttcactgactta
atgccagagatatcattagaataggaagtgattccaaaaggcgctcttaccgtgtgagtagccgctcctaatctt
ttgcccataggagcacttcttaaaagaaaagcagtggctattatatcaccaaccgcagctagaaagagaacagaa
gagtagtgaagaaggaatccctcaaagccccttatccatctctagcagctggcattcccatattccttattagat
tagatgttactttctttagtttagcagatccggtatcggaggcattcttctaacattaagaatctttcaggggga
gtagcacttctacctatacttgaagtagctaaggctcgatcggaggctggcttgcggttgaaaaacacacacgcc
gagggtagcacacatcttatgaagcgacggtcaggtctctgtctcactcagtacagtaggcacactgcttcagaa
agaaggtcggagtcgtgaacaggaaaagcgttgaccgggaggccagcctattatcctagccacgtctatccataa
gtagtgagagcgagcttggttctctggcgcctggatatttccccctttctagtccctccggttgggacaagctaat
gagatccggactcagccctttttgtagtaaacgcagcttaaacactatcattattccaaaacatttttttcaaaat
gaaacaaagcccatagtagttattcttacccccccagggaagggcacttgattgatcagaagaaaggtttacttct
tccccctttatgctatgggggtcgttagcgctaccagtgtatttattcttctgtttgtccgctatctttaaaaatg
ccttctgtcagtatgagtgagcgtctttttcccagctcccggtctttctttctttccttgccttactactgggataa
cgtgcctaagtagcataacaatcaatcaatatagatcctatccccttttccttagcatatactatactaagatctt
cgaacctttattcttaccctcttctcaagcagcaatcgattcgaccgcaaggaacaaaccgtagctacctgctt
caactcatagatgactacccgaggatcactcactctactagaaaagcactaagagggattcgattaccaaccccct
ttttctaactatagaacttatgaactgctacgcttCTATAGCCCGCTAAAGCAATTGTCGCAGCTCCAGCTGAAG
CAATTGTAGCAGCTCCG

xxxxx – reverse complement sequence of the atp9-f2 UMP
xxxxx – homology to TK81-MS mtDNA
xxxxx – see text
xxxxx – largest homologous block to mtDNA of *Beta vulgaris* subsp. *maritima* genotypes E and G
xxxxx – stretches of homology *Amborella trichopoda* mtDNA
xxxxx – stretch of homology to a segment of *Cucumis melo* subsp. *melo* mtDNA
xxxxx – stretch of homology to a segment of *Carica papaya* mtDNA
xxxxx – stretch of homology to a segment of *Salvia miltiorrhiza* mtDNA
xxxxx – homology to a sequence stretch often located approx. 1710 nt downstream of *rps12* ORF
xxxxx – homology to a plant mtDNA stretch of unconserved location
xxxxx – homology to *Capsicum annuum* mtDNA and *Solanum* sp. nuclear DNA (the latter homologies may not reach to the very end of the marker – they may be a few nucleotides shorter)
XXXXX – 3′ terminus of *Beta vulgaris* subsp. *maritima orf105e*

Fig. 15. Sequence features of the N_atp9-f2_c39 vectorette marker

ing site – this can be deduced from the fact that *Beta vulgaris* subsp. *maritima* genotype E (acc. no. FQ014226) (it showed continuous homology over the entire sequence of the N_atp9-f2_c39 marker) in the respective mtDNA region had two neighboring stretches of homology to the atp9-f2 (custom) primer. Since the primer sequence was 20 nt long and the primer-like stretches were separated with a 4-nucleotide motif, the 24-nt extension of N_atp9-f2_c39 is the one to be expected. Despite this difference the sequence at the atp9-f2 primer terminus retained homology to the above mentioned sequences from plants of the *Solanaceae* family. Accordingly, the pattern of the remaining homologies was the same as for N_atp9-f2_c38 (Fig. 15).

Vectorette marker N_cob-r1

Almost the whole sequence (without the first two Cs and the terminal sequence agtcttcttcat of the custom primer) showed close to 100% homology (identity) (there were only two substitutions) with a non-coding fragment of the TK81-O mtDNA (acc. no. BA000009, nucleotides 169777–168837) located between *orf115a* and *orf317*. Identical hits were returned for the mitochondrial genomes of *Beta vulgaris* subsp. *maritima* (genotypes A, B, E and G) and *Beta macrocarpa*. This vectorette marker contained three stretches of homology to the TK81-MS mtDNA – they corresponded to nucleotide positions: 3–486 (almost 100% identity – there were only two substitutions), 876–943 (100% identity) and 854–884 (93.5% identity – two substitutions). In TK81-MS (acc. no. BA000024) the first two stretches were duplicated, their respective nucleotide coordinates were the following: 215519–216002/401175–400692 and 217590–217657/399104–399037. The third shortest stretch was present in one copy (nucleotides 434104–434074). This sequence also returned hits to other plant mitochondrial genomes, these hits mostly matched the approximate nucleotide positions: 98–192, 231–297, 327–528 and 772–840. The available sequence data either did not indicate conservancy in genomic location of these sequence stretches or were not sufficient to determine if such conservancy existed (Fig. 16).

Vectorette marker N_rps7-r1

The whole sequence showed almost 100% homology (identity) (there was only one substitution) with a fragment of the TK81-O mtDNA (acc. no. BA000009, nucleotides 179130–178077) located between *orf214* and the *cob* ORF, and including 35 nucleotides from the 3' end of the latter. Directly below the *cob* stop codon there was a conservative stretch of approx. 150 bp – in other plants it also adjoined the 3' terminus of the *cob* ORF. When the marker was compared with mtDNAs from some other plant species this homology extended further downstream of the *cob* ORF – up to nucleotide 377 (counted from the *cob* stop codon). Further below was

a stretch of 44 nucleotides which was homologous to a fragment of the conservative sequence located approx. 130 bp upstream of the plant *nad6* ORFs. Apart from the rps7-r1 primer sequence, there were no other *rps7* sequences in the analyzed vectorette marker. The largest block of homology with the TK81-MS mtDNA (acc. no. BA000024) extended from nucleotide 1 to 1034 (in TK81-MS it corresponded to nucleotides 445142–446175). In the marker sequence there were also several shorter stretches which were represented by the TK81-MS homologies (often multiple): 73/76–99/100 (four copies), 439–468 (one copy), 854/855/857/858–878/891/892/900 (ten copies), 900–1054 (two copies) and 1022–1054 (one copy) (Fig. 17).

ccttttggttttccctcttccttgagctcttcttctagttctagcacggccctctcgcttatggcctaaaaagt
ctctctcagtccaaggtactatcctctagtccgtcttgtaggaaagaacgtatgtaagtcaagcactatcttcat
tctatctcttagtctagttggtagcagtctaaggattcgcagctctccaaaggtaggttaatcagtatatcccca
ttcccttttcttccttgctttcgattaggggtccctgccctccattctctcaagtacctctccgcttcctctgctt
gctttcatgtagtttcttccggctagactttcttcttgtcttgacttttcttttctattcgttgacgtacttttca
ctctttcgaggatgtaacaagtccttctttccagtccgctagcagtccatgcctaacagtcttcgaagccgtatc
agggccaattccctttctttgagattttcactcgggtgagagggaaggaattaggttgcaggggagccccattcg
cgatcgtctttgtaggagaagcccttgagaaattctcctattaccttacctgctgcccggggacaattcctttgt
atcgaggggctttcttccttcctaggtctattccttctctccgactattcggtatgtccgcccagtcgtaaaag
cttttccttgagactgagagtaagcaccagcgggagaatcttccttcctctttcattatgttgccctgggacaca
ttccttcgaccgtggtagaaaaggtaaagagatatgagagtgaggctagcttttggcatcccctaagccctaagag
aatggatgtgcttagacgactcttggaactcgtctttacgcccttcgttcaaaattccgagaccaaaagacgact
ccacgatcctgcgcagctgacctccctcagctatttaggcgccagtcttcttcat

xxxxx – reverse complement sequence of the cob-r1 UMP
xxxxx – homology to TK81-MS mtDNA
xxxxx – plant mitochondrial homologies

Fig. 16. Sequence features of the N_cob-r1 vectorette marker

Vectorette marker S_atp6-r

An almost identical sequence was present in TK81-MS (acc. no. BA000024, nt pos. 421697–423295) – the difference was limited to one indel and three substitutions. The sequence constituted the majority of the *preSatp6-Satp6* ORF (only 306 nt from the 5' end and 48 nt at the 3' end were missing). A similar sequence block was also present in sea beet male-sterile genotype E (94% identity). In comparison with TK81-O mtDNA, homology was limited to 744 nt extending from the atp6-r primer terminus. This part corresponded to the canonical *atp6* sequence. The remaining sequence of this vectorette marker was unique – except for the mentioned hits from male-sterile sugar and sea beets – it did not show reasonable homologies to any other database-deposited sequences (Fig. 18).

CTAATTCTTACACGGATGAGACTGAGAACACCTGAtcagtgaaaaattcttacaccaagaattgacgagtgagta
ttacaccaagaattgacaagcggattcgttttttctagttggctatgttgatatagcttagatagggaaaagaaa
gtctactatagggtaagggtgaactttcaaatccggttgttcccgaaactccccgtcccttactcgtcttttgaa
agcaagaacattcgcatagaagagtctctgtatcatgcatgctcctccacttttgaaaaaaaaaaagaaccttct
atcctctaaaaaagtgggcagtatgtattctgcatgaatatgcttctgtactgggaatatctcatagtgagaagg
cccagagatcataaaggatgagatgagaatgaagtcaacatactctgtttcgagagagatttgagatcgtaagt
aactcagtgagtgctctctgtaagtaagaaagcagttatgtaggcaccaactccgagcttctcttccagggctaa
gacgttctatgatgagatgaaaagcttatcaggggcaaatccaaaaaatggaggaggggtggaggcaggcctttc
tggcaaggctaagttcataaggactcgaatggatacagagtatccgcaagagaaattccatgaagtcctctcgag
tctcaaagggcagcaggaaaatagctttgattccaaagaaggaagattggcttaggttacgggaagaaagaaagt
gaaggtctaggccaatatatataatgctattgtaataaaatttctatcgctttaaaggcgtccttcgatcgagcc
ttgagaaaggggataaaggttaggagtcagaaaaggtagttgacttacttagtgcttgcattaagggcctgacca
atcctctagttgaatcagaaatgaacttctctatcgatagaggcaagcgatggaggcacgtaagtgaagctcaag
cagcaaaagcggttaagcaagtaattctctttttatttgtctatcaaaagcgcttgccattcgttggttattag
gggc

xxxxx – reverse complement sequence of the rps7-r1 UMP
xxxxx – the longest block of homology to TK81-MS mtDNA
XXXXX – *cob* ORF (incomplete – only a 3′ terminus is present)
xxxxx – homology to a sequence adjoining the 3′ terminus of the plant *cob* ORFs
xxxxx – further reaching homology to a sequence adjoining the 3′ terminus of the plant *cob* ORFs
xxxxx – homology to a sequence located approx. 130 bp upstream of the plant *nad6* ORFs

Fig. 17. Sequence features of the N_rps7-r1 vectorette marker

Vectorette marker S_cob-r1

Almost the whole sequence (without 12 nt – agtcttcttcat – corresponding to the 5' part of the cob-r1 primer) showed 100% homology (identity) with a non-coding duplicated fragment of TK81-MS mtDNA (acc. no. BA000024, nucleotides 217291–217657 and 399403–399037) located between *orf221* and the gene coding for tRNA-Lys (recognizing codon AAA). A similar sequence was present in the mitochondrial genome of *Millettia pinnata*, and as shortened versions – in *Gossypium hirsutum* (approx. 360 bp corresponding to nucleotides 5–327), *Hevea brasiliensis* (approx. 100 bp similar to nucleotides 5–105) and *Boea hygrometrica* (approx. 90 bp homologous to nucleotides 2–95). Genome locations of these sequences were not conserved. In TK81-O and in other *Beta* species only the right terminal part of this sequence (nucleotides 300–367) was present – it was located between *orf115a* and *orf317* (Fig. 19).

GCCCCTTTGGCTATATTGATCTTTTCGGTTTGCTTCTTTCCCAAAAATGGACGATTACGGTATTCGCTAGAGAAC
TGTTCTTTTTTTCTCCTATTACTTTGGGGGGATTACCGGCTTTTGGATACTCCGTGGTGTGAATGGGTACTCTAT
GTTTGTGGCCTCTTAACTTTATTCACGCTTCTTCAAGCGATGGTCCATCGAAAGGAGGAAGGGGTCGGCAAGTTT
CTTCTACCGGCTGCCTTCCTAATTAGTGACCTCGTGCTCCTTTGGCTTCACTCGGGGGCGATTCAGAATCCGCTG
GGAAAGGCGGCCCTAACGGGCGCCGCAGTCCTAATACTCTTTAGCCAAGCCTCCGCCGAGTCCACAAGGCAGCGG
GAGCTTGCAAATGGGGCGTATCTCATGTGTATCCCTCCTATACTATATATCATAGGAGGGGCCACGGATTGGCCC
CGATTCGTAAACGAACTCACCTTTTTTATTCTTATCCTCTGGCTAGTCTTAGTTTTTTGGGTTGTTACTTTGGAG
GACGAAAAAGAGCAGAAGGATCTGCTCCCGCTCCTATCTGCGATCTTCTCCGCTTTGGCGTCGCCGAGTTACGGC
CCCGACGTCTGGGTAAAAGAGTGGCTCGCCGCCTTTTCCTTAACCATTGGGGCGGTTTTTTTGTTTGAGAAGGTA
GGCTACCGCAAAGATACCCTTCCAGAAACGTCCCGTACTTGGGAACCTTGGTTTTGCTTTTTTGCTGTTCTTTCT
TTTGGTGTTGCACTTAGTTTTTTTCATGCATCACCCGCCCTAGGGGTGGGAACTGAAAGTGAAGCTTTTCACGAT
GCGATCAATCCAGCAGCGGGACGGATCCCCAGCCCCCTAGAGCAATTTTCCATTCTCCCATTGATTCCTATGAAA
ATAGGAAACTTGTATTTCTCATTCACAAATCCATCTTTGTTTATGCTGCTAACTCTCAGTTTGGTCCTACTTCTG
CTTCATTTTGTTACTAAAAGGGAGGAGGGAACTCAGTACCAAATGTTTGGCAATCCTTGGTAGAGCTTATTTATG
ATTTCGTGCTGAACCTGGTAAACGAACAAATAGGTGGTCTTTCCGGAAATGTTAAACAAAAGTTTTTCCCTTGCA
TCTTGGTCACTTTTACTTTTTTGTTATTTCGTAATCTCCAGGGTATGATACCCTATAGCTTTGCAGTTACAAGTC
ATTTTCTCATTACTTTGGGTCTTTCATTTTCCATTTTTATTGGCATTACTATAGTGGGATTTCAAAGAAATGGGC
TTCATTTTTTAAGCTTCTCATTACCTGCAGGAGTCCCGCTGCCGTTAGCACCCTTTTTAGTACTCCTTGAGCTAA
TCCCTCATTGTTTTCGCGCATTAAGCTCAGGAATACGTTTATTTGCTAATATGATGGCCGGTCATAGTTCAGTAA
AGATTTTAAGTGGGTTCGCTTGGACTATGCTATGTATGAATGATCTTTTATATTTCATAGGAGATCTTGGTCCTT
TATTTATAGTTCTTGCATTAACCGGTCTTGAATTAGGTGTAGCTATATTACAAGCTCATGTTTTTACGATCTTAA
TCTGTATTTACTTGAATGATGCT

xxxxx – reverse complement sequence of the atp6-r UMP
XXXXX – *preSatp6-Satp6* ORF (incomplete - the 5′ part and 3′ terminus are missing)
xxxxx – homology to TK81-O mtDNA

Fig. 18. Sequence features of the S_atp6-r vectorette marker

tctccgtgatcgtagttcccaattttcttactatgagatggcctctgtcagccttaccagctgtccgagatgagc
gtccaaaccgaatcgatcttgcggccccccgctgcgctcctcccatcgaaaaaactcggacctgaagccaactgg
ggcgtcgttatatttgtgaagcttcattcccttcaccttaaagcggacggcaataggtaattcaggctcttgctt
gatcgtgacacctgttgggaccctcgaaggcccccctgatctcttggctcttgtgtagacaatatggtaagaaga
aaattccgagaccaaaagacgactccacgatcctgcgcagctgacctccctcagctattaggcgccagtcttct
tcat

xxxxx – reverse complement sequence of the cob-r1 UMP
xxxxx – homology to an mtDNA segment of *Hevea brasiliensis*
xxxxx – homology to TK81-O mtDNA

Fig. 19. Sequence features of the S_cob-r1 vectorette marker

Vectorette marker S_nad3-r

The whole sequence showed close to 100% homology (identity) with a fragment of TK81-MS mtDNA located between genes *cox2–2* and *rps12* (acc. no. BA000024, nucleotides 334359–335178) (there were only two substitutions and one indel). The right part of this marker contained the complete *nad3* ORF. Its upstream sequence contained two conservative blocks – their approx. locations are 186–264 and 402–461 (these are the nucleotide positions of the marker). In plant mitochondrial genomes the larger block overlapped (with 22/23 bp) the 5' part of *nad5* exon 1. This held true for TK81-MS as well, but for another location – in this mtDNA fragment *nad5* exon 1 as a whole was missing – except for the first 23 bp present in the 186–264 block. In other plants the smaller block (402–461) preceded the *nad3* ORF – just as in this mtDNA fragment. In TK81-O (acc. no. BA000009) five mtDNA fragments showed homology to this marker (given in the order of growing E value):

- 263121–263614 (321–819), contained the *nad3* ORF and the adjoining shorter conservative block;
- 171828–171905 (187–264), corresponded to the larger conservative block – accordingly it overlapped the 5' part of *nad5* exon 1;
- 200792–200727 (409–474), contained the majority of the shorter conservative block and a few nucleotides from the 5' end of the *nad3* ORF;
- 263060–263121 (146–207), contained the sequence preceding the larger conservative block and the left terminal stretch of this block;
- 200359–200313 (146–192), the sequence features were as for the previous fragment (Fig. 20).

ataatttatatataaagaagagatgcccaggccaccgaagctaacaaaaggctaaggatagccatttagttgttt
cttcttatgccgactaaagctaactgcgccacagctggaatctccctggaaatcctcacttaagcctgcagcgaa
tcttataaaagaatctattttcagttctcttttgaattactcgccgctctgtcattatctgatttttggttgttt
gatcacactcgaaattatgtatttacttatcgtctttttttgaaagatcttttgaaaaaaagcgaggaaagtgt
ctattttccagcgatactaaaaattatagaggacttaacttaaattaaatgtcctctatgttccacgtacgtagc
gcaaatgttatataatggcctttctgtccttttttttcggtatgccgctccgcctgcaaggagcgagaaaacaaat
tggtctgtggtgATGTCAGAATTTGCGCCTATTTGTATCTATTTAGTGATCAGTCTGCTAGTTTCTTTGATCCCA
CTCGGTGTTCCTTTTCCATTTTCTTCTAATACTTCGACTTATCCAGAAAAATTGTCGGCCTACGAATGTGGTTTC
GATCCTTTCGGTGATGCCAGAAGTCGTTTCGATATACGATTTTATCTTGTTTCAATTTTATTTATTATCCTTGAT
CCGGAAGTAACCTTTTTCTTTCCTTGGGCAGTACCTCTCAACAAGATTGATCCGTTTGGATCTTGGTCCATGATG
GCCTTTTTATTGATTTTAACGATAGGATTTCTCTATGAATGGAAAAGGGGTGCTTCGGATCGGGAGTAA

xxxxx – reverse complement sequence of the nad3-r UMP
xxxxx – homology to TK81-O mtDNA
xxxxx – homology to a stretch overlapping the 5' part of exon 1 from plant *nad5* genes
xxxxx – homology to a stretch preceding plant *nad3* ORFs
XXXXX – *nad3* ORF

Fig. 20. Sequence features of the S_nad3-r vectorette marker

Vectorette marker S_rrn26-f1

Almost the whole sequence (without the terminal custom primer sequence tgagtgcgc) showed 100% homology (identity) with a fragment of TK81-MS mtDNA located between *rps3* and *orf496* (acc. no. BA000024, nucleotides 344532–345158). A part of this sequence, the first 627 bp, was also present in another location – between *rps3* (another copy) and *orf184* (nucleotides 344532–345158). This 627 bp sequence, in only a slightly modified version, was also present in TK81-O – between *orf246* and *nad4* exon 3 (acc. no. BA000009, nucleotides 274028–274647). The left part of the marker sequence – approximately 900 bp (nucleotides 16–922) – was highly similar (96%) to a fragment of *Brassica* sp. mtDNA located ca. 1.7 kb downstream of the *trnfM* gene. The left terminal segment of this part (roughly, nucleotides 15–177) was even more conservative – its homologues were found in diverse mtDNA locations of both di- and monocotyledonous species. Also, the right terminal segment showed additional homologies – their extent was variable and all of them represented dicotyledonous species (not shown). Moreover, close to the custom primer terminus, there was a stretch of 111 bp (nucleotides 945–1054) showing homology to a fragment of mtDNA from the *Poaceae* species in which it was often located ca. 1.4 kb downstream of *nad2* exon 2. The fragment displaying this *Poaceae*

atttcattcaaagtaattcaaagtacaggaactgctttaacaaaagggctcgaactacagtgatctgagaccttc
ttcccttaatctgggaggtaggttatgtaggcgacagcccttaatttaataggttgtttcgataccaaagggcaa
aataatcgatgagtagtatgcctggtaagaaccttctcaagtagagaaagagaatctcatttttttataagaaaga
tgctaaggccgatattcctgacgttaagaatctcctcttatgcttcgccaacacccaataaggaccacaaagtgc
attactaatgaccactggagaagatcccttgcctctgctccttgatagagccgctcaaggagttgcttcttttat
cgctttactattattgcttctttgccaatcgaaagaggaagagattgccatactgttcaggatacattgtccatg
tctcttttctctgaaggctgttcctattcatattcaaattgttagatagcttctatctttcagtttatatctttg
gttctcgaggttctcaagtttcaatccagatgtaatgtattgtatttcaccctcgcgataagaactatatgggta
aattgcttgacgatctatccaaaaggagccgtgtccatgcggttttcctctctcgcatagagccaaccttactct
atgtgctgtcctgggtgggctaccatccctttccattccactactttacatgttagctcgtctctgccttcgtct
ttgtgtctgccgattcaattgatgctgctcgtccgtcaggatggctcgacgtgggttcctgttcttattcctagt
tccaccttcctgacatagtttccgtctcttgaataccaactttctgtaggggtggattgagcggtggcaacgtcc
gaaaggtcttattttgtagtagcggaaagggacatcttgaacgcaggggggaaggagacccctttctgagcaagct
aaagagcccttcctttagatccctgctttccggcaaggcactcaatgaaaccatccctttgagggaatttcgct
ccaagaagaagttcttccgtcagattgaaagatatggctttgtcccgccgttagagtgagtgcgc

xxxxx – reverse complement sequence of the rrn26-f1 UMP
xxxxx – homology to TK81-O mtDNA
xxxxx – homology to a plant mtDNA stretch of unconserved location
xxxxx – homology to a segment of *Brassica* sp. mtDNA
xxxxx – homology to a stretch from mtDNA of the *Fabaceae* species
xxxxx – homology to a segment of mtDNA from the *Poaceae* species located often ca. 1.4 kb downstream of *nad2* exon 2

Fig. 21. Sequence features of the S_rrn26-f1 vectorette marker

homology overlapped the short stretch – nucleotides 918–954/959 – with single-copy homologies in mtDNA of *Vigna* sp. and *Lotus japonicus*, both representing the family *Fabaceae*. The genomic location of this homologous sequence in *Vigna* sp. was different from that of *Lotus japonicus* (Fig. 21).

4.4. Mapping NGS reads to the identified cytoplasmic markers

The Illumina sequence reads generated for the NS 031 and NO 031 mitochondrial DNAs (mtDNAs) were mapped to the sequenced RAPD and vectorette markers which were described in the chapter *Sequence analysis of the identified RAPD and vectorette markers*. The reads from NS 031 were mapped both to the markers for the S- and N-cytoplasm, as were the reads from NO 031. The results of mapping are shown in figures 22–25, which illustrate how extensively the marker sequences, nucleotide by nucleotide, are covered by the Illumina reads. It should be noted that in the following description the given values of coverage and sequence length/position were read from the figures and therefore, their accuracy is limited by the resolution of these figures. This does not apply to the percentage of a covered marker sequence which was calculated in silico assuming that the minimal significant coverage was 5x.

Below, the main features of the coverage profiles obtained upon mapping with the NS 031 reads are described (Figs. 22 and 23).

Marker N_A-08 was covered in its left part (the marker sequences are presented in the chapter *Sequence analysis of the identified RAPD and vectorette markers*) comprising 43.9% of the entire sequence. It was here that the profile displayed two major features: the leftmost halved peak which was followed by a steep-sloped plateau – both separated by a crevice located over nucleotide position 30 and reaching as deep as to 40x coverage. The right slope of the plateau declined down to a coverage of 1x over the last 40 nt of the covered region. The maximum coverage – 100x – was reached in the left part of the plateau, whereas the height of the halved peak exceeded 80x – this height was reached on the left terminus of the marker sequence.

Although most of the N_B-05 marker was covered by the NS 031 reads (85.8%), for the majority of this sequence the coverage was comparatively low and did not exceed 50x. There were only two major elevations in this coverage profile – they were centered around nucleotide positions 320 and 2030. The former had the form of a peak reaching a coverage of over 550x, the latter was hill-shaped and approx. three times lower. At their base the elevations were 140 and 320 nt wide, respectively.

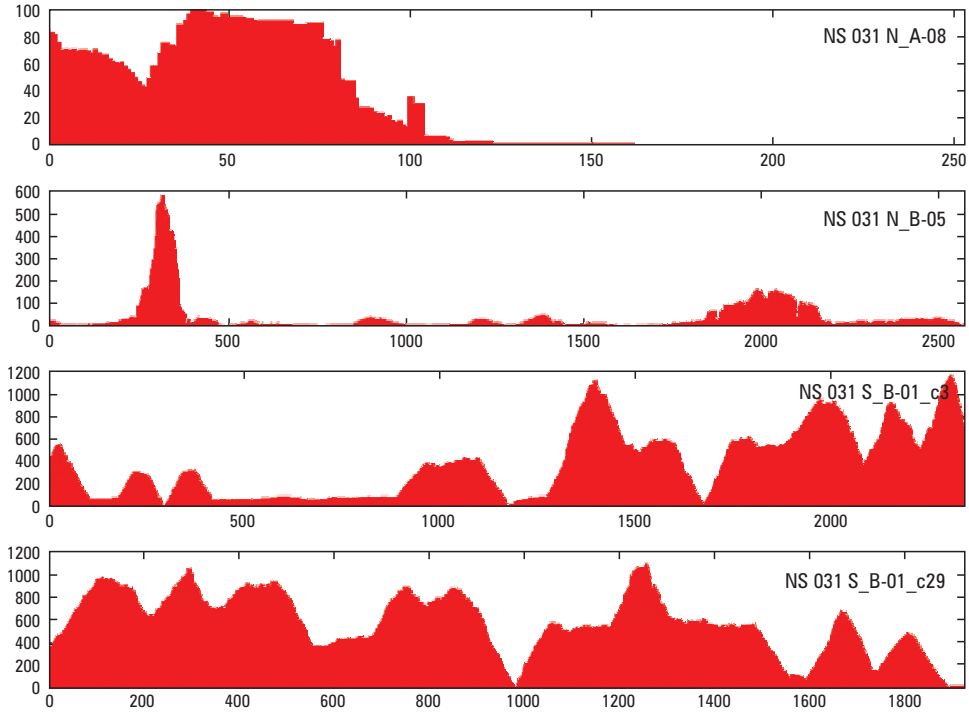


Fig. 22. Coverage of the RAPD markers with the Illumina sequence reads obtained for the NS 031 mitochondrial DNA

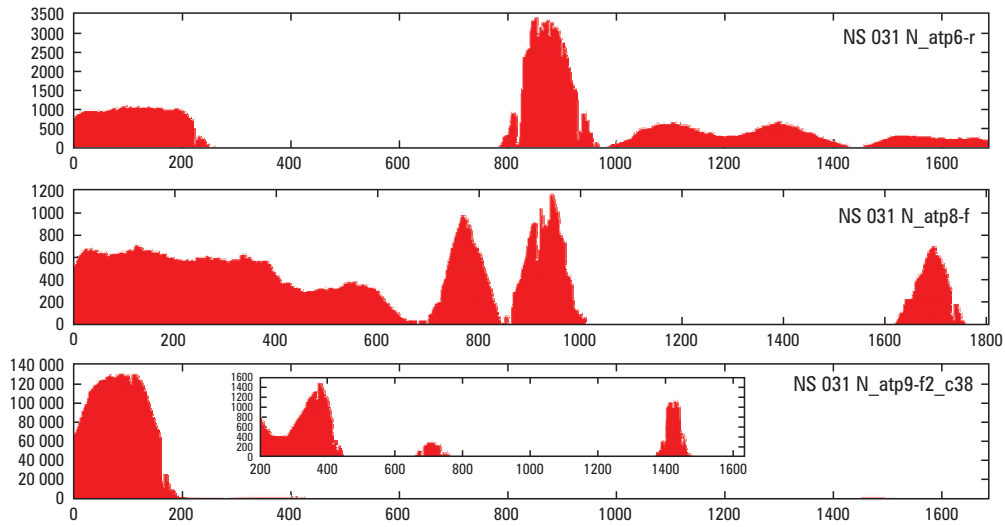


Fig. 23. Coverage of the vectorette markers with the Illumina sequence reads obtained for the NS 031 mitochondrial DNA. For the N_atp9-f2_c38 marker the coverage was also shown after removal of the left terminal 200 nt which contain the pEV1-like macrosatellite

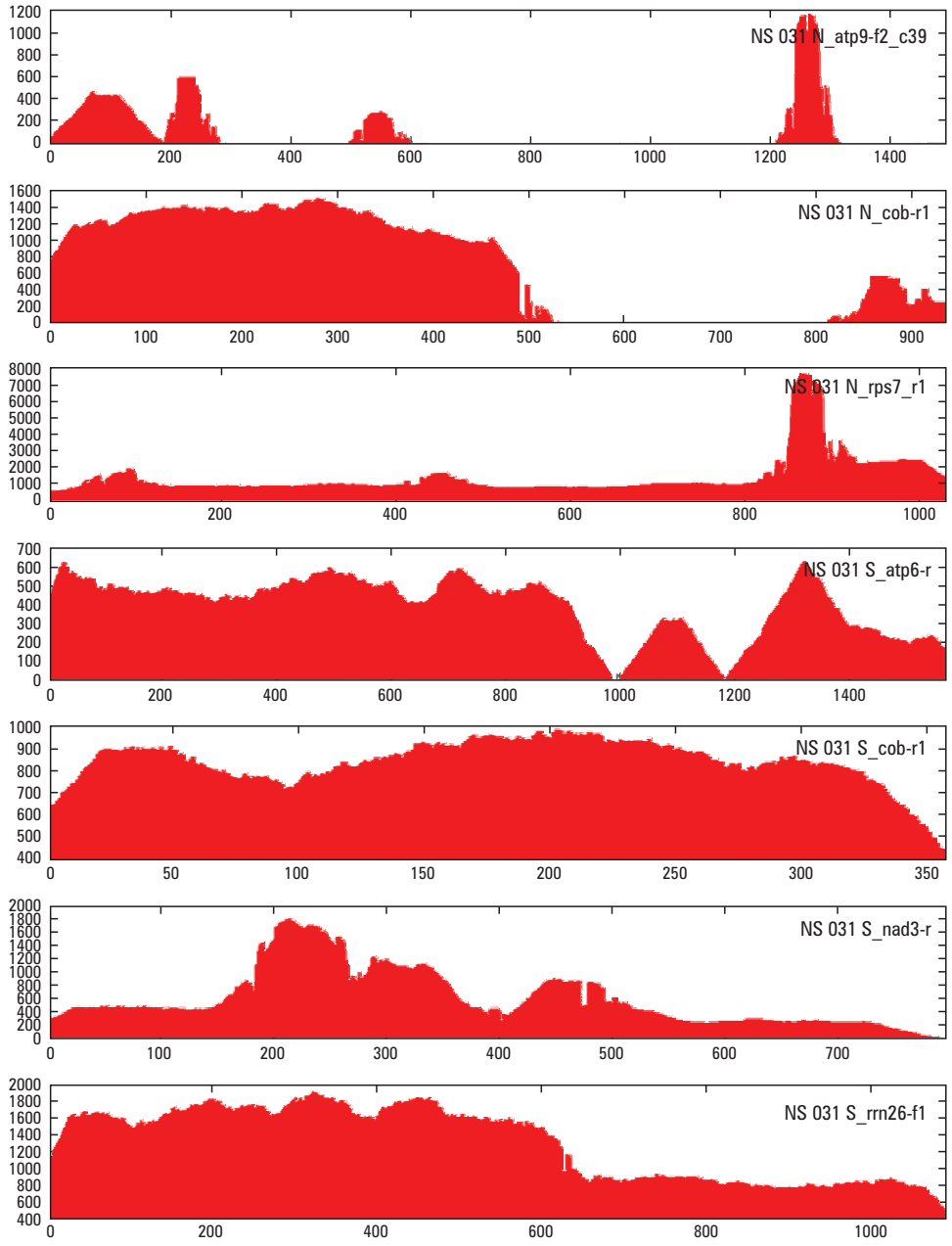

NS 031 N_atp9-f2_c39
NS 031 N_cob-r1
NS 031 N_rps7_r1
NS 031 S_atp6-r
NS 031 S_cob-r1
NS 031 S_nad3-r
NS 031 S_rrn26-f1

Fig. 23. cont.

Despite the fact that almost the whole sequence (99.9%) of the S_B-01_c3 marker was covered by the NS 031 reads the coverage profile was nonuniform. It was formed by a series of ten major peaks, which in most cases were separated by V-shaped crevices of variable depth. There were only two regions of an evened coverage oscillating around 70x. The first was 80 nt long and centered around nucleotide position 135, the second extended between nucleotide positions 420 and 890. The coverage ranged from 20x in the deepest leftmost crevice to almost 1200x reached by the highest rightmost peak. At the base the peaks were 110–270 nt wide.

The NS 031 reads almost completely (99.8%) covered marker S_B-01_c29 in the form of two elevations occupying 85% of the sequence and two sibling peaks seated on 290 nt at the right terminus of the sequence. The peaks and elevations were separated by V-shaped crevices with the deepest reaching as low as to 20x coverage. The left elevation occupied ca. 950 nt extending from the left terminus. It was crowned by a range of five peaks, with the highest reaching over 1000x coverage. The second elevation occupied 600 nt between the first one and the sibling peaks. There was a single peak which emerged from this internal elevation and reached over 1100x coverage. At their base the width of all peaks was comparable – ranging from 60 to 160 nt.

Marker N_atp6-r was covered over 68.9% of its length. The coverage profile consisted of a plateau extending over the 240 left terminal nucleotides followed by 520 nt of zero coverage, a peak built on 200 nt and a 650 nt wavy segment where coverage reached from 0 to 800x. The plateau and peak reached coverage of 1000 and 3300x, respectively.

Marker N_atp8-f was covered over 62.4% of its sequence. The coverage profile had the form of a plateau over the leftmost 700 nt followed by two neighboring shark tooth-shaped peaks, 600 nt region of zero coverage, another shark tooth-shaped peak and uncovered 30 rightmost nucleotides. Each of the three mentioned peaks spanned roughly 160 nt at their bases. The mean coverage of the plateau region was 400x. The peaks reached coverage level of 600–1000x for the rightmost and middle, respectively. The plateau region and the two adjacent peaks were separated by circa 20 nt stretches where coverage oscillated from 0 to 50x.

Marker N_atp9-f2_c38 was covered over 43.5% of its length. The left terminal 200 nt formed a peak of enormously large coverage – reaching a value of 120000x. The remaining portion of the profile showed the presence of three much lower peaks. The first was located on an extension of the super-high peak, spanning 200 nt at its base and reaching coverage of 1400x. It was separated from the next peak by 200 nt of zero coverage. Of the other two lower peaks, one was placed approximately in the central part of the sequence, the other – close to the right terminus. They reached a coverage level of 400 and 1200x, respectively. They had a similar width at their base – around 100 nt. These peaks were separated by 600 nt of zero or very scarce coverage. Also, the right terminal 160 nt were uncovered.

The sequence of marker N_atp9-f2_c39 was covered in 37.7%. The coverage profile was very similar to the one produced for marker N_atp9-f2_c38 except it lacked the super-high left terminal peak and its associated lower peak was substituted by two other neighboring peaks. Of these two peaks, the left terminal had 200 nt at the base and reached a coverage of 500x, the values for the right adjacent peak reached 100 nt and 600x, respectively.

Marker N_cob-r1 was covered in 70.4%. The covered regions were located terminally – 530 nt from the left side and 120 nt from the right side. Their coverage profile was plateau-shaped with irregular diminishing oriented towards the 270 nt uncovered region located in between. The mean coverage of these regions was 1200 and 400x for the left and right, respectively.

Marker N_rps7-r1 was covered over its entire length. The 800 nt extending from the left terminus were covered rather evenly – on average 1100x – with only two minor elevations around 100th and 450th nucleotide where the coverage reached 2000 and 1500x, respectively. The coverage of the rightmost 100 nt was also regular but two times higher than for the evenly covered portions of the left segment. Both those plain regions were separated by 100 nt showing a peak of coverage reaching a level of 7500x.

The S_atp6-r marker was covered over 99.2% of its length. The majority of the sequence showed coverage in the range of 200–600x. However, in two regions the profile exhibited the presence of V-shaped crevices at the very bottom of which the coverage dropped down to zero. These minima were located at distances of 1000 and 1180 nt from the left terminus.

The whole sequence of the S_cob-r1 marker was covered with the NS 031 reads. The profile had a form of soft waves where coverage reached from 450 to 1000x.

Marker S_nad3-r was covered over its entire length. The left terminal 160 nt and right terminal 270 nt were covered rather evenly reaching a level of 500 and 300x, respectively. In between these plains the profile contained a region of elevated but irregular coverage. The highest peak in this region reached a level of 1800x, while in the deepest rift the coverage was comparable to the maximal level reached in the right terminal region.

Solid coverage was also observed for the S_rrn26-f1 marker. Its profile was clearly bipartite – the segments left and right, although both being quite regular (smooth), displayed different coverage levels. They oscillated around 1700 and 800x for the left segment and the right segment, respectively.

As was mentioned before, the same set of markers was subjected to mapping with the NO 031 reads. The resulting coverage profiles are described below (Figs. 24 and 25).

The NO 031 reads covered the sequence of the N_A-08 marker over its entire length. The coverage profile was plateau-shaped with the minima on the sequence termini – 380 and 270x for the left and right terminus, respectively. The majority of the sequence was covered more than 1200x with a maximum of 1400x.

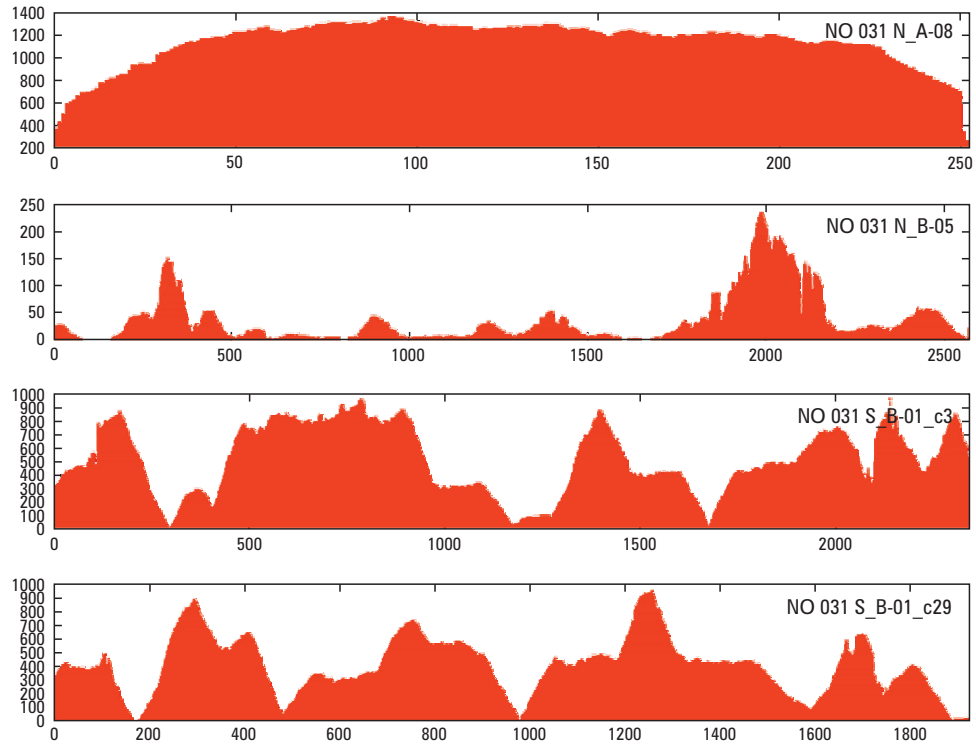


Fig. 24. Coverage of the RAPD markers with the Illumina sequence reads obtained for the NO 031 mitochondrial DNA

When the NO 031 reads were mapped to the N_B-05 marker, its sequence was covered in 86.4% and the resulting coverage profile resembled the one obtained with the NS 031 reads. The major difference was related to the height of the major elevations. Upon the use of the NO 031 reads the left elevation was lower than the right – the difference was approx. 1.5-fold.

For marker S_B-01_c3, the coverage profile formed by the NO 031 reads also resembled that obtained with the NS 031 reads. The major difference was that the two regions of evened ca. 70x coverage were substituted with elevations reaching a level of approx. 800x. Almost the entire marker sequence (99.8%) showed coverage exceeding the threshold of 4x.

In the case of marker S_B-01_c29, for the majority of its sequence the coverage profile looked very similar to the one generated with the reads of NS 031. The major difference appeared over the 550 left terminal nucleotides. Here, the coverage profile had a form of two elevations separated by a V-shaped crevice reaching zero coverage. In their upper parts both these elevations were split into two sibling peaks. The peaks of the left terminal elevation reached a coverage of 400x, with the

peaks of the second elevation roughly two times higher. Altogether the marker sequence was covered in 99.2%.

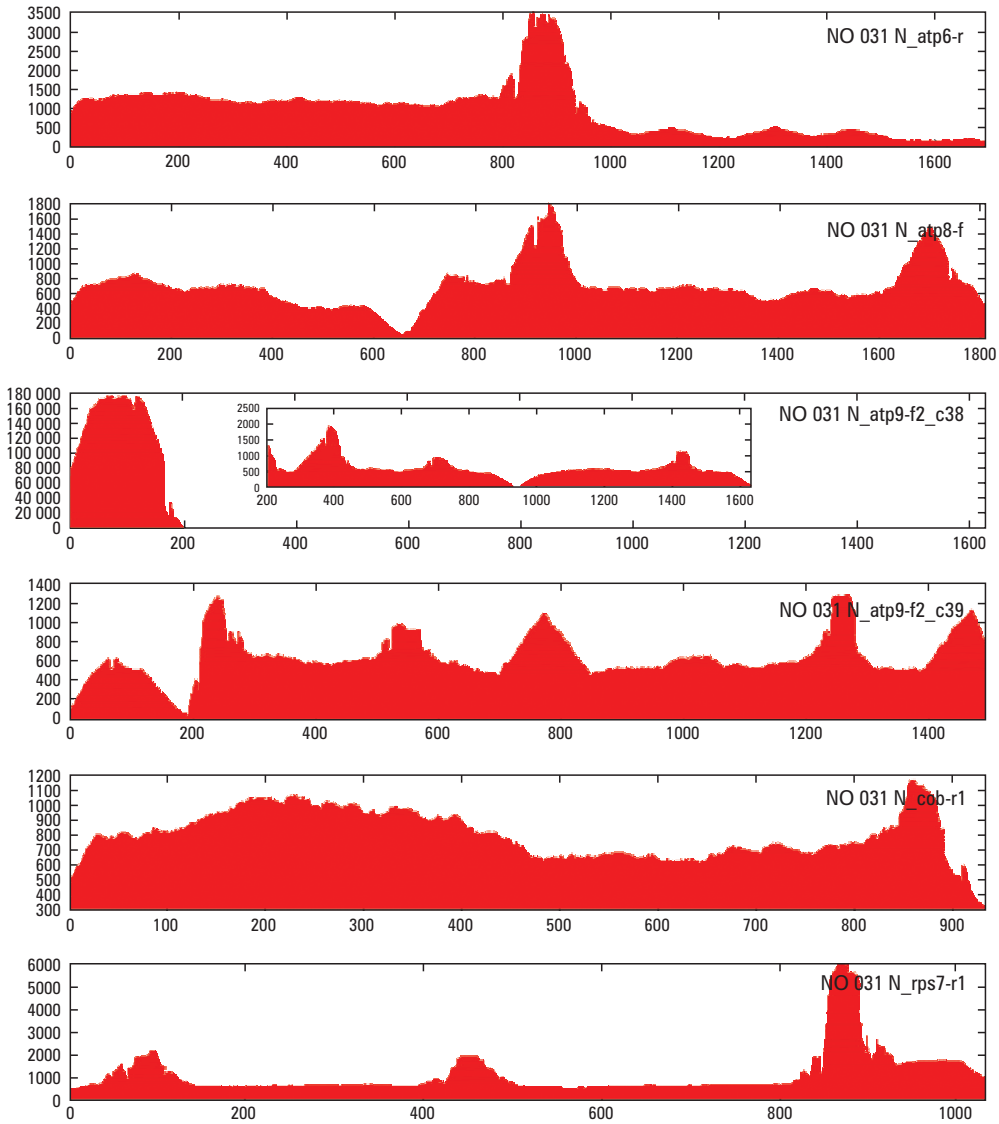


Fig. 25. Coverage of the vectorette markers with the Illumina sequence reads obtained for the NO 031 mitochondrial DNA. For the N_atp9-f2_c38 marker the coverage was also shown after removal of the left terminal 200 nt which contain the pEV1-like macrosatellite

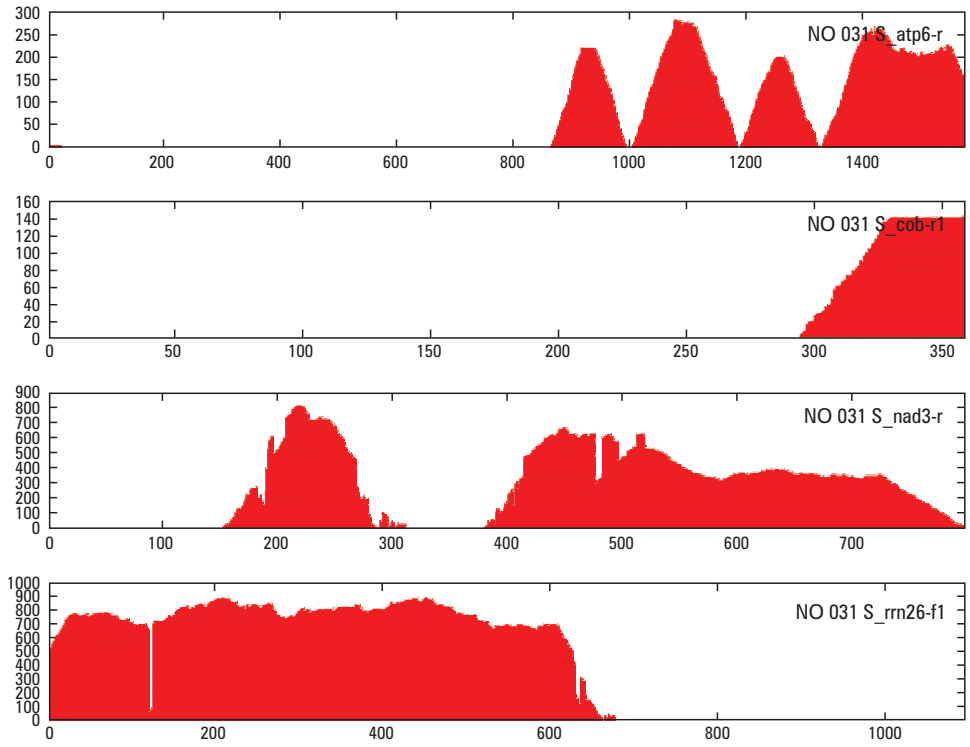


Fig. 25. cont.

For the N_atp6-r marker mapping with the NO 031 reads resulted in a hundred percent coverage. In the left segment of the profile, extending over 800 nt, the coverage was very uniform oscillating around 1200x. In the right segment of the sequence, which was of similar length, the coverage took the form of a wavy line with the minimum and maximum at 200 and 500x, respectively. These two segments were separated by a peak of coverage reaching a level of 3500x. At the footing, together with its tiny pre- and after peaks, the peak spanned the length of 180 nt.

The whole sequence of the N_atp8-f marker showed coverage with the NO 031 reads. Over the majority of the sequence coverage was within the range of 400–820x. It dropped down to as low as 50x in the region around the 650th nucleotide from the left terminus. In addition, the coverage profile displayed the presence of two peaks – one was located approximately in the center of the marker sequence, the other at the right terminus. Coverage reached a level of 1800 and 1400x for the former and latter peak, respectively. At the base, which was on a level of circa 600x, the central peak spanned 150 nt and the terminal peak – 180 nt.

Marker N_atp9-f2_c38 was covered on almost its entire sequence – only one of the 1631 nt remained uncovered. Similarly, as upon mapping with the NS 031 reads,

the left terminal 200 nt formed a peak of enormously large coverage – this time reaching a value of 175000x. The majority of the remaining sequence coverage was on a level of 500x. The profile also showed the presence of three elevations around nucleotides 380, 700 and 1420 where coverage reached a level of 2000, 1000 and 1200x, respectively. At the footing the first elevation spanned a length of 180 nt, the latter two – 100 nt. Between nucleotides 850 and 1000, the profile took the form of a symmetrical V-shaped valley with zero coverage in the lowest point at the approximate nucleotide position of 930.

Mapping with the NO 031 reads resulted in complete coverage of marker N_atp9-f2_c39. Over the left terminal 200 nt, the profile had the shape of a soft-sloped elevation reaching a level of 650x. This elevation emerged from a coverage level of 50x. The remaining part of the profile had the form of a plateau where coverage oscillated around 600x and from which five peaks emerged. The highest and lowest peak reached a coverage level of 1300 and 1000x, respectively. The first, second and fourth (when counted from left to right) peaks were shaped very similar to the respective elevations from the N_atp9-f2_c38 profile. The remaining two peaks had a shark-tooth shape and appeared in locations corresponding to the declines in the N_atp9-f2_c38 profile.

The N_cob-r1 marker was also covered over its entire length. The left segment of the profile – spanning 470 nt – had the form of a softly-sloped plateau. Moreover, there was a peak of coverage at the right terminus of the sequence – reaching a level of 1150x. The right slope of this peak declined to 300 nt at the very right terminus of the marker sequence. In between the plateau and the peak there were 380 nt of lower coverage – although its value did not drop below 600x.

The N_rps7-r1 marker was also entirely covered. The coverage profile resembled very much the one produced for this marker with the NS 031 reads – minor differences were only visible upon comparison of the absolute coverage values.

The S_atp6-r marker was covered over 45.6% of its length. The left segment of the sequence – approximately 870 nt – showed no, or very scarce (at most 5x), coverage. The remaining part of the coverage profile was very characteristic – it had the form of three shark-tooth-shaped peaks followed by a plateau at the right terminus. The peaks and the plateau reached a level of 200–280x and were separated by symmetrical V-shaped valleys where in the lowest points coverage dropped down to zero. At their footing the peaks spanned stretches of 130–180 nt, the right terminal plateau extended over 270 nt.

The vast majority (83.3%) of marker S_cob-r1 was not covered. Such regions extended from the left terminus up to the approximate nucleotide position 300. From this point, coverage was raised and over 30 nt reached a level of 150x which was maintained up to the right terminus.

Altogether, 24.2% of the sequence remained uncovered in the S_nad3-r marker. There were three regions of coverage. The left terminal region extended over

40 nt where coverage gradually diminished from 10 to 2x for the leftmost and rightmost nucleotide, respectively. The second region took the form of an irregularly-sloped peak reaching a level of 850x. The peak spanned the region between nucleotide positions 150 and 310. It was surrounded by two regions of zero coverage extending up to the left and right terminal (see below) regions of coverage. The third region of coverage extended from nucleotide 370 up to the right terminus. In the central part of this region, coverage ranged from 300 to 700x, while in the peripheral parts, coverage gradually slipped down to 2 and 25x for the leftmost and rightmost nucleotide.

Within the right segment of its sequence marker S_rrn26-f1 was either uncovered or covered very weakly – barely reaching the threshold of 5x. This segment corresponded to the right terminal 420 nt, the uncovered stretches amounted to 35.5% of the total sequence. The remaining part of the profile was plateau-shaped with coverage oscillating around 800x. After the left terminal of 120 nt, and over the next 5 nt, the plateau was incised to a level of 70x. At the peripheries of the plateau, coverage slipped down – to 520x at the left terminus and to zero on the right side. The line of the right plateau slope was steeper and showed the presence of incisions.

When the NS 031 reads were mapped against the S-cytoplasmic vectorette markers the whole marker sequence was usually covered. The same picture was obtained for one N-cytoplasmic marker – N_rps7-r1. Contrary to that, the remaining N-cytoplasmic vectorette markers showed patchy distribution of the mapped NS 031 reads – some regions of these markers were covered and some were not. This distribution pattern was largely reversed when the NO 031 reads were used for mapping. In this case all S-cytoplasmic vectorette markers showed the presence of uncovered regions, while the markers of the N-cytoplasm were covered completely or (in one case) almost completely.

For the purpose of comparison, the mapping of the sequencing reads was also applied to three standard mitochondrial open reading frames (ORFs) – *atp9*, *nad3* and *nad6*. In this case, regardless of the reads source – NS 031 or NO 031 – the entire ORFs were extensively covered (Figs. 26 and 27). Moreover, for *nad3* and *nad9* the patterns of coverage generated by the NS 031 and NO 031 reads were very similar. Different coverage profiles were obtained only for the *atp9* ORF. In the NS 031 profile the 5' region – approximately 70 terminal nucleotides – showed markedly (approx. two-fold) higher coverage than the remaining portion of the *atp9* ORF. The NO 031 reads covered this ORF rather evenly – only the region around the 70th nucleotide showed underrepresentation – coverage dropped here from approx. 1200 to as low as 750x.

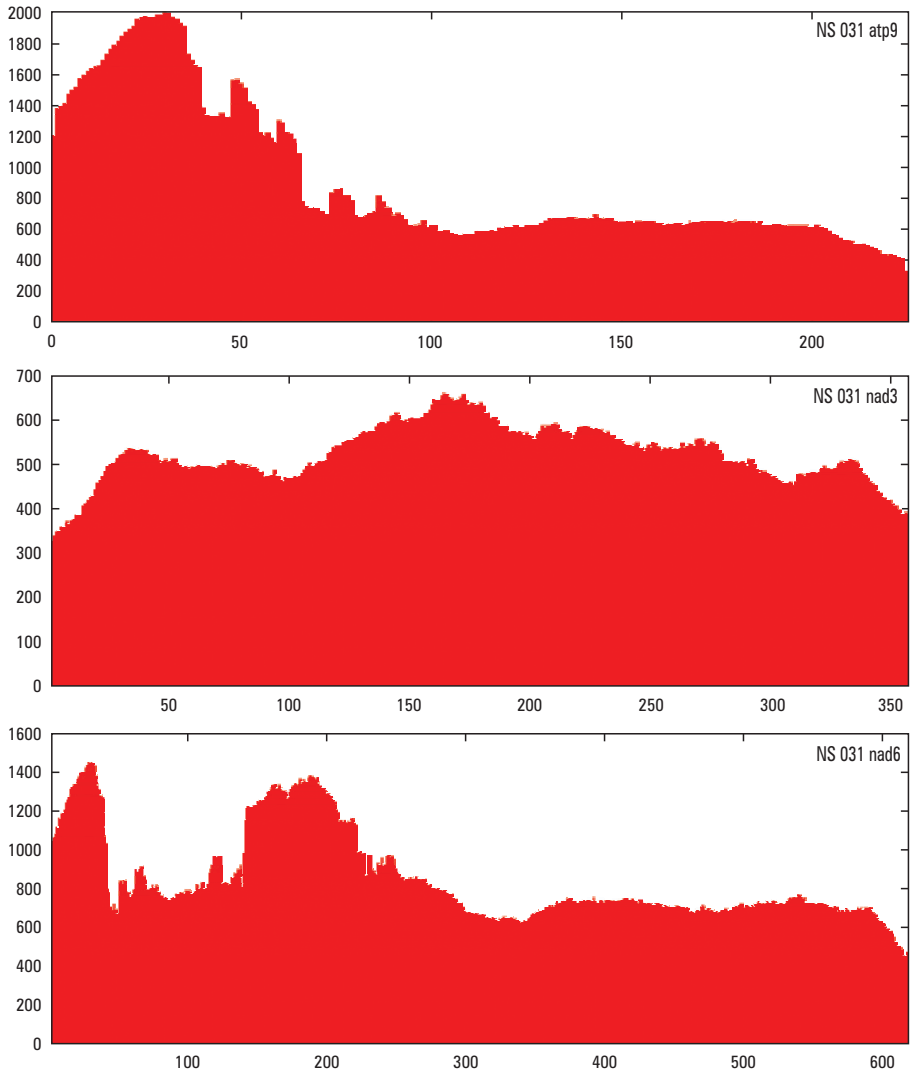


Fig. 26. Coverage of the selected sugar beet mitochondrial ORFs – *atp9*, *nad3* and *nad6* – with the Illumina sequence reads obtained for the NS 031 mitochondrial DNA

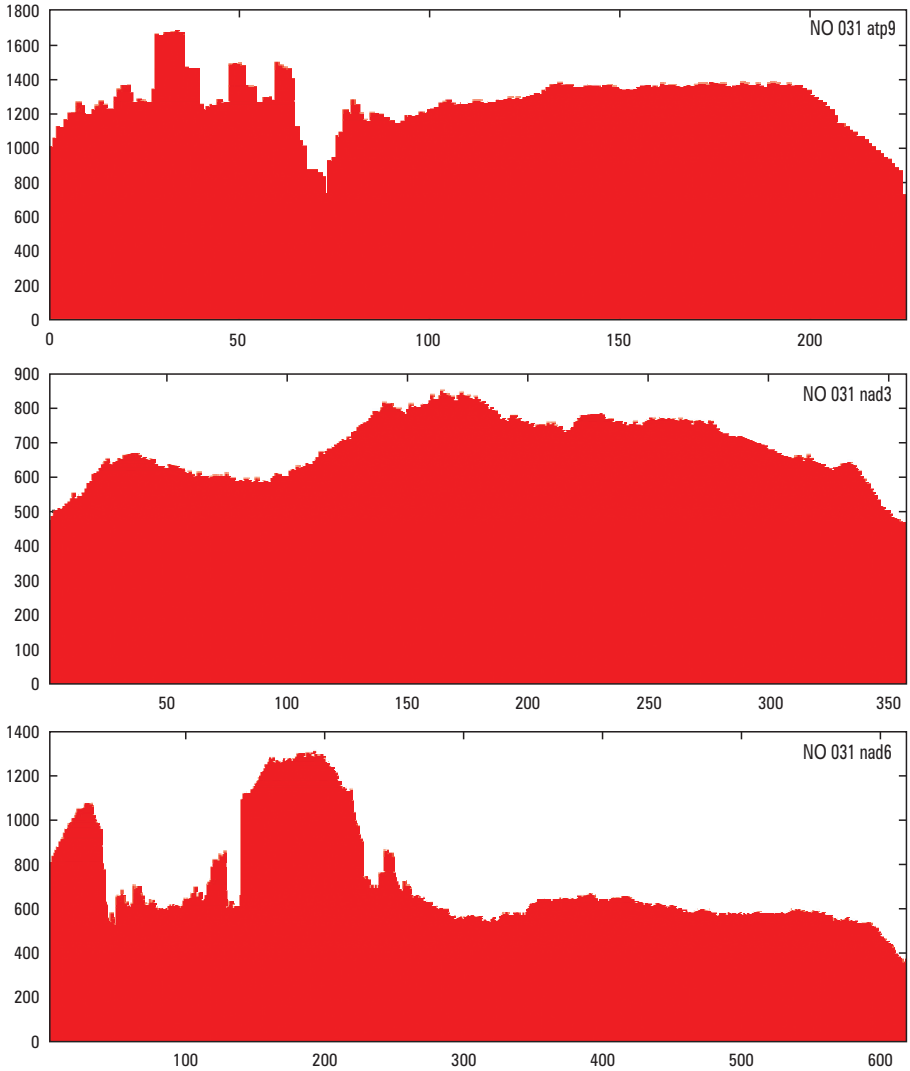


Fig. 27. Coverage of the selected sugar beet mitochondrial ORFs – *atp9*, *nad3* and *nad6* – with the Illumina sequence reads obtained for the NO 031 mitochondrial DNA

4.5. Elaboration of the plasmotype-specific SCAR markers

Sequence data obtained for the identified RAPD and vectorette markers allowed for the design of PCR primers which could be used for the purpose of cytoplasm identification. For each of the sequenced RAPD markers two such primers were designed – they were anchored in the terminal regions, one for each terminus, in opposite

(inward-facing) orientations. Both these primers were included in the amplification reaction. For each of the sequenced vectorette markers one such primer was designed – it was anchored slightly below the sequence of the vectorette unit and oriented towards the terminus bearing the sequence of the respective universal mitochondrial primer (UMP). In the PCR a given vectorette-derived primer was used together with the respective UMP. The diagnostic value of the RAPD- and vectorette-derived primers was tested on two sets of accessions. The first set consisted of 22 sugar beet breeding stocks – 11 male-sterile (MS) lines and 11 corresponding male-fertile maintainers (Owen type/type-O lines). In this set each line was represented by a single plant. The second set consisted of ten table beet breeding stocks – five MS lines and their respective maintainers. Each of the table beet lines was represented by two plants.

Table 6 shows the theoretical sizes and specificity of the expected amplification products. The product size was deduced on the basis of the sequence data (refer to chapter *Sequence analysis of the identified RAPD and vectorette markers*), the anticipated specificity of the products reflects specificity of the respective source markers. In the following text plasmotype-dependent distribution refers to the situation when a given PCR product i/ is present in at least two lines carrying the same cytoplasm and is absent or markedly more weakly amplified in all lines with the opposite cytoplasm type, ii/ is absent in at least two lines carrying the same cytoplasm and is present in the remaining lines.

Table 6. Theoretical sizes and specificity of the PCR products expected for the RAPD- and vectorette-derived primers

Primer pair	Source marker	Expected product size [bp]	Expected specificity (cytoplasm type)
O_A08_scar-f/O_A08_scar-r	N_A-08	212	N
O_B05_scar-f/O_B05_scar-r	N_B-05	1692	N
S_B01_scar-f/S_B01_scar-r	S_B-01_c3 and S_B-01_c29	2100 and 1686	S
atp6-r/O_atp6-r_scar	N_atp6-r	1704	N
atp8-f/O_atp8-f_scar	N_atp8-f	1822	N
atp9-f2/O_atp9-f2_38_scar	N_atp9-f2_c38	1632	N
atp9-f2/O_atp9-f2_39_scar	N_atp9-f2_c39	1507	N
cob-r1/O_cob-r1_scar	N_cob-r1	928	N
rps7-r1/O_rps7-r1_scar	N_rps7-r1	1034	N
atp6-r/S_atp6-r_scar	S_atp6-r	1552	S
cob-r1/S_cob-r1_scar	S_cob-r1	381	S
nad3-r/S_nad3-r_scar	S_nad3-r	767	S
rrn26-f1/S_rrn26-f1_scar	S_rrn26-f1	1104	S

4.5.1. Examination of the sugar beet breeding stocks

The analysis was performed for the following accessions: NS 031, NS 032, NS 033, NS 041, NS 042, NS 043, NS 044, S 00188, S 00500, S 001299, S 001353, NO 031, NO 032, NO 033, NO 041, NO 042, NO 043, NO 044, O 00188, O 00500, O 001299 and O 001353 (Figs. 28, 29, 30).

Primers O_A08_scar-f and O_A08_scar-r yielded a product with the expected size of 0.21 kb and present in all analyzed accessions. A generally higher amplification efficiency was obtained for the type-O lines – the only exceptions were NO 033, NO 041, NO 042 for which this efficiency was comparable or smaller than the one observed in the MS lines.

The use of primers O_B05_scar-f and O_B05_scar-r led to the amplification of products estimated at 1.15, 1.7 and 2.1 kb. For a given plant, efficiency of their amplification was size-dependent with the larger products being more abundant. These products were clearly visible but only in the lanes of five maintainers: NO 031, NO 032, NO 043 (only the larger fragments), NO 044 and O 001299. For the remaining O-type lines no amplification was observed. As far as the MS lines

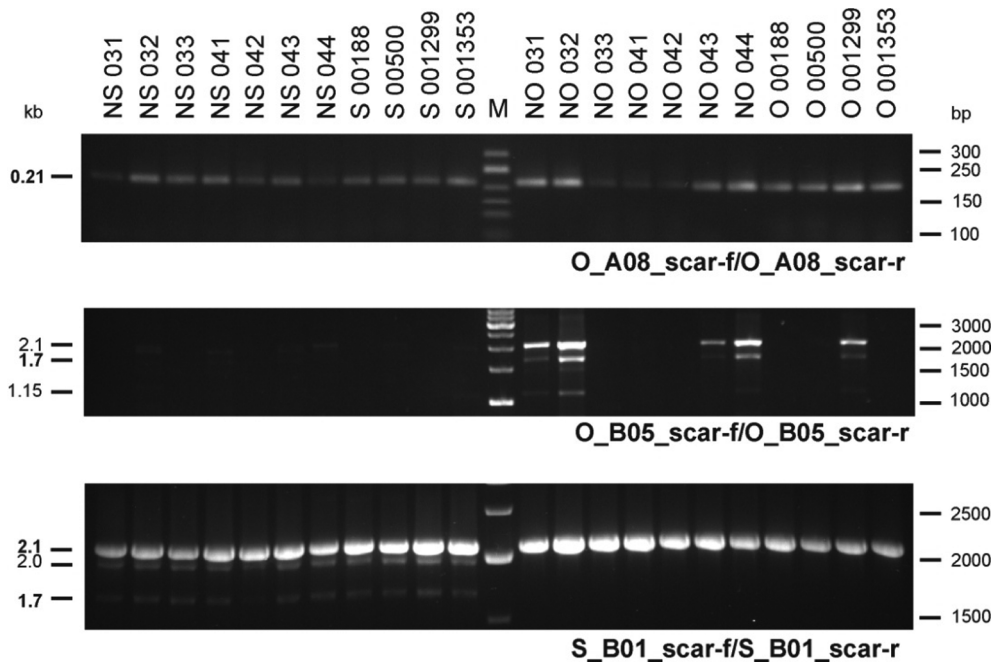


Fig. 28. PCR products generated for the sugar beet male-sterile lines and their maintainers with primer pairs O_A08_scar-f/O_A08_scar-r, O_B05_scar-f/O_B05_scar-r and S_B01_scar-f/S_B01_scar-r. Both primers of each pair were anchored within the corresponding RAPD marker. The sizes of the expected products are shown in bold on the left

are concerned PCR products were either not observed or barely visible – mostly with a size of approx. 2 kb.

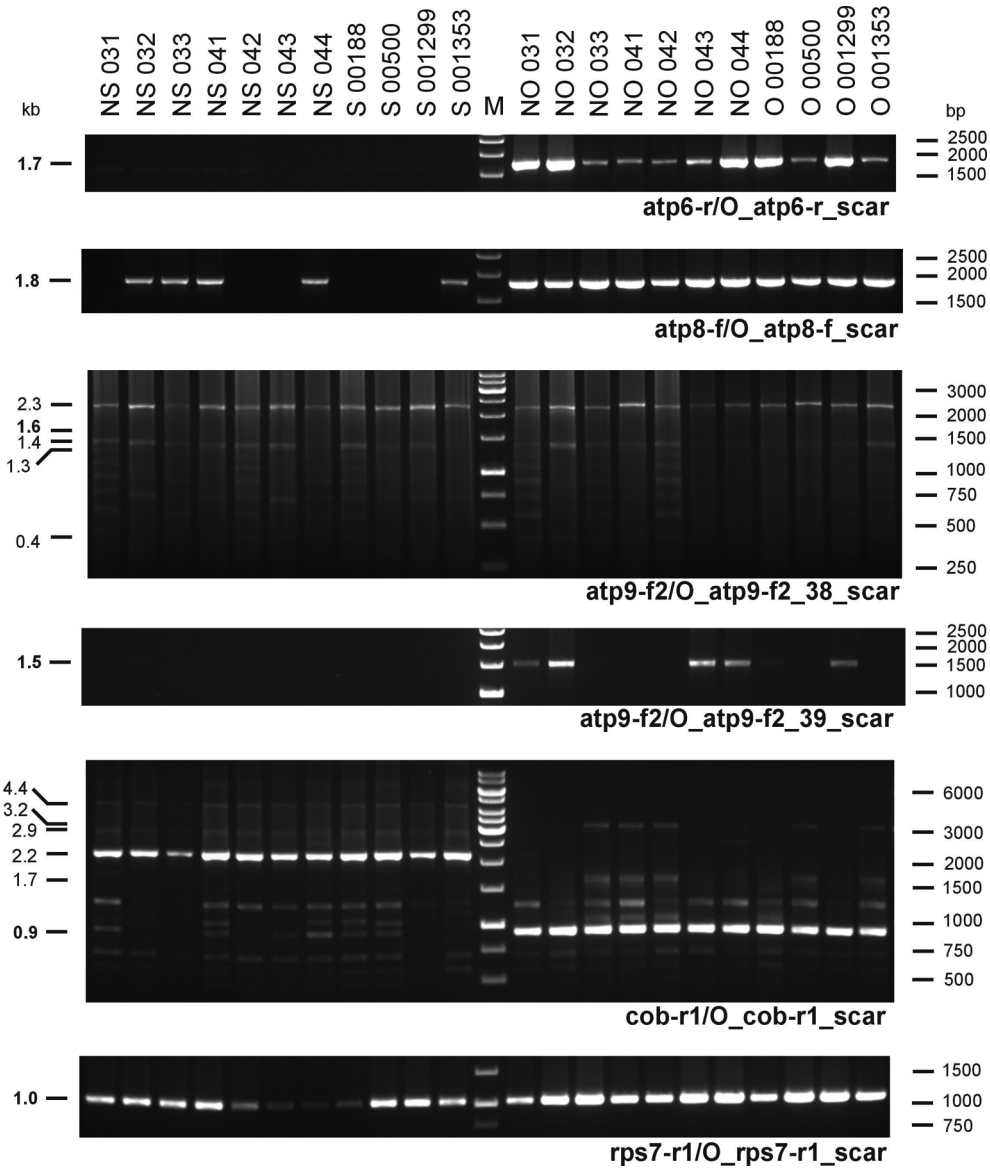


Fig. 29. PCR products generated for the sugar beet male-sterile lines and their maintainers with primer pairs atp6-r/O_atp6-r_scar to rps-r1/O_rps7-r1_scar. The first primer of each pair was a universal mitochondrial primer (UMP), the second was anchored within the corresponding vectorette marker. The sizes of the expected products are shown in bold on the left

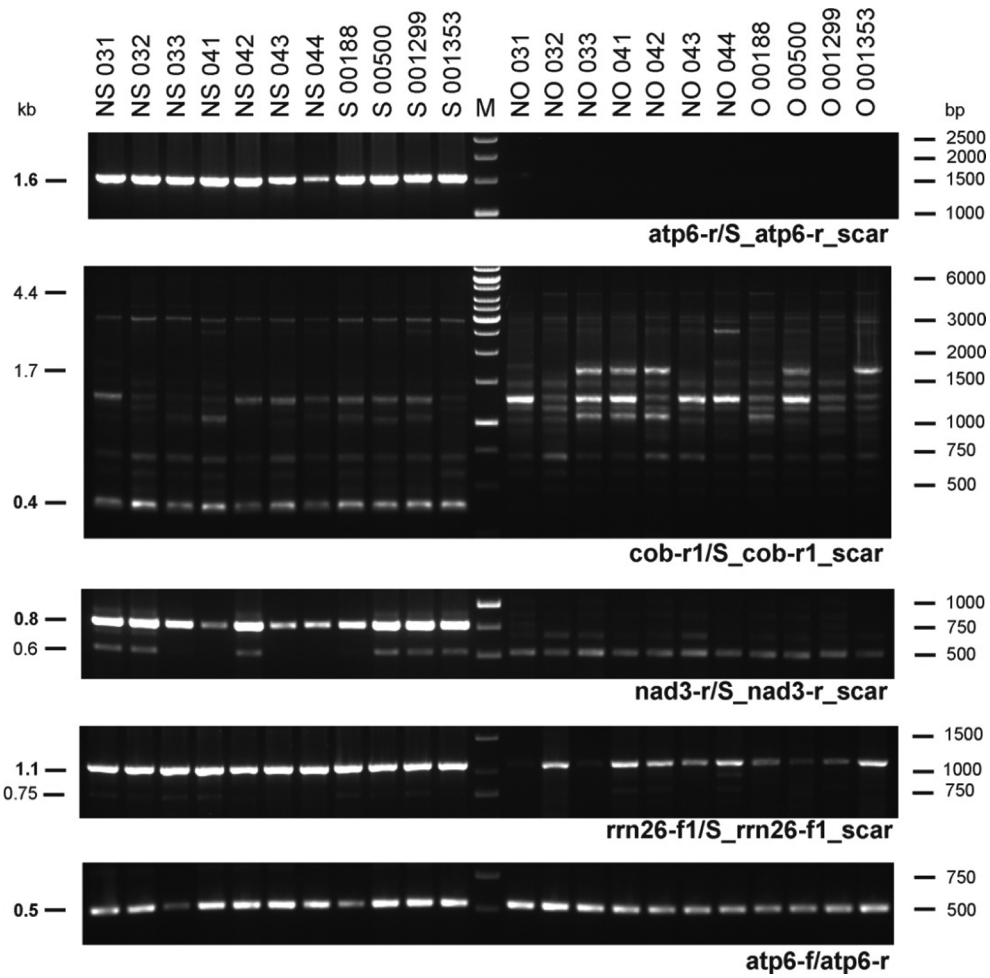


Fig. 30. PCR products generated for the sugar beet male-sterile lines and their maintainers with primer pairs atp6-r/S_atp6-r_scar to rrn26-f1/S_rrn26-f1_scar and with the control primers (atp6-f, atp6-r). Both control primers were universal mitochondrial primers (UMPs). In the remaining primer pairs the first primer was a universal mitochondrial primer (UMP) and the second was anchored within the corresponding vectorette marker. The sizes of the expected products are shown in bold on the left

The strongest DNA fragment, amplified with primers S_B01_scar-f and S_B01_scar-r had a size of 2.1 kb and was present in all analyzed breeding stocks. It was the only fragment generated for the male-fertile lines. Each of the male-sterile lines exhibited two additional amplification products of 1.7 and 2.0 kb. The efficiency of their amplification was much weaker than that of the 2.1 kb fragment.

Primers atp6-r and O_atp6-r_scar produced a single DNA fragment of the expected size – 1.7 kb. Its amplification was observed in all analyzed maintainer lines, but for lines NO 031, NO 032, NO 044, O 00188 and O 001299 the efficiency of the amplification was distinctly higher than for the remaining O-type lines. In the MS lines this DNA fragment was either not amplified or barely visible.

For the majority of the analyzed accessions primers atp8-f and O_atp8-f_scar yielded a single amplification product of the expected size – 1.8 kb. The product was generated for all male-fertile lines. It was also obtained for five male-sterile lines – NS 032, NS 033, NS 041, NS 044 and S 001353 – however, in this case amplification was weaker. The remaining six MS lines did not yield any amplification products with this primer pair.

After the use of primers atp9-f2 and O_atp9-f2_38_scar the expected amplification product (1.6 kb) was not obtained. Instead, weakly amplified products of 2.3 and 1.4 kb were observed – the former in all analyzed accessions, the latter in the majority of them. Moreover, for some samples a ladder of weakly amplified fragments ranging from 1.3 to 0.4 kb was produced. The resulting amplification profiles did not show the plasmotype-specific distribution.

PCR with primers atp9-f2 and O_atp9-f2_39_scar either did not yield any product or gave a product of the expected 1.5 kb. This product was obtained for only the following six maintainer lines – NO 031, NO 032, NO 043, NO 044, O 00188 (barely visible) and O 001299.

According to the sequence data, primers cob-r1 and O_cob-r1_scar should yield the N-plasmotype-specific product of 0.9 kb. Indeed, this was the major product obtained for the maintainer lines. It was also visible in the profiles of six male-sterile lines (NS 031, NS 041, NS 043, NS 044, S 00188, S 00500) – albeit, in this case the respective bands were much weaker. The major DNA fragment produced for the MS lines had a length of 2.2 kb – this product was absent in the profiles of the maintainer lines. Both MS and O-type lines exhibited the presence of a few additional, poorly amplified PCR products. Four of them showed plasmotype-dependent distribution – the products of 2.9 and 4.4 kb were specific for the MS lines while the products of 1.7 and 3.2 kb were obtained only for five O lines – NO 033, NO 041, NO 042, O 00500 and O 001353.

Primer pair rps-r1/O_rps7-r1_scar generated a single product of the expected 1 kb. Its amplification was observed in both the male-fertile and male-sterile lines. Generally, in the latter, the efficiency of the amplification was weaker.

The PCR product amplified with primers atp6-r and S_atp6-r_scar showed strictly plasmotype-dependent distribution – it was obtained for all eleven analyzed male-sterile lines, but missing in the male-fertile lines. The length of this DNA fragment – 1.6 kb – was accordant with the sequence data.

Both for the male-fertile and male-sterile lines primers cob-r1 and S_cob-r1_scar yielded multi-band profiles. Three of the PCR products displayed plasmotype-de-

pendent distribution. Each of the MS lines showed the presence of the expected 0.4 kb DNA fragment – it was absent in the profiles of the type-O lines. The product of 1.7 kb was obtained only for five maintainer lines – NO 033, NO 041, NO 042, O 00500 and O 001353. The 4.4 kb product was weakly amplified in most of the maintainer lines, but not in the MS lines.

According to the sequence data, primers nad3-r and S_nad3-r_scar should yield a product of 0.8 kb. Efficient amplification of this DNA fragment was observed in all analyzed male-sterile lines. In the male-fertile lines this product was either absent or amplified very weakly. Yet another DNA fragment obtained with this primer pair showed plasmotype-dependent distribution. The following MS lines: NS 033, NS 041, NS 043, NS 044 and S 00188, were the only ones for which the 0.6 kb product was not amplified.

With primers rrn26-f1 and S_rrn26-f1_scar a product of the expected size – 1.1 kb – was obtained for all of the analyzed breeding stocks. It was strongly amplified in the MS lines while in the maintainers its amplification efficiency varied from barely traceable to slightly weaker than that of the male-sterile lines. In addition to the 1.1 kb product, two weakly amplified smaller fragments were obtained but their distribution was not plasmotype-dependent.

The control primers – atp6-f and atp6-r – for all analyzed accessions yielded an expected fragment of 0.5 kb.

4.5.2. Examination of the table (red) beet breeding stocks

In this analysis the following accessions were included: AR79A, 279A, W411A, 218A, 375A, AR79B, 279B, W411B, 218B and 375B (Figs. 31, 32, 33).

With the use of primers O_A08_scar-f and O_A08_scar-r, a product of the expected 0.21 kb was generated for three male-sterile and three male-fertile lines – these were W411A, 218A, 375A, W411B, 218B and 375B. The amplification efficiency did not show any legible consistency.

The sequence data indicate that primers O_B05_scar-f and O_B05_scar-r should yield a product of 1.7 kb. Within the analyzed set of table beet accessions, the amplification of this product was not observed. Instead, for some samples, these primers produced a weak product of 2.1 kb. It was visible in single plants from each analyzed accession except 218A.

Altogether, four PCR products were obtained with primers S_B01_scar-f and S_B01_scar-r. Two were monomorphic in the analyzed set of accessions – their sizes were estimated at 2.1 and 4.5 kb. The other two – of 1.7 and 2.0 kb – were observed in all male-sterile but not in the maintainer lines. With respect to amplification efficiency the 2.1 kb product markedly dominated over the remaining three.

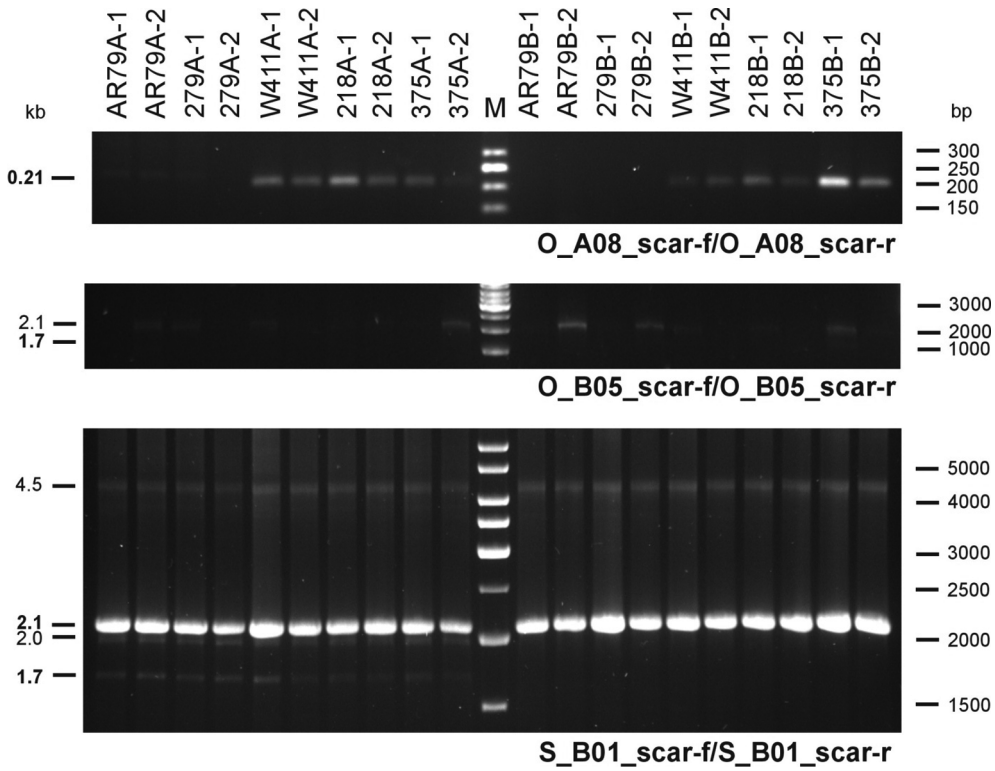


Fig. 31. PCR products generated for the table beet male-sterile lines and their maintainers with primer pairs O_A08_scar-f/O_A08_scar-r, O_B05_scar-f/O_B05_scar-r and S_B01_scar-f/S_B01_scar-r. Both primers of each pair were anchored within the corresponding RAPD marker. The sizes of the expected products are shown in bold on the left

PCR with primers atp6-r and O_atp6-r_scar either did not yield any DNA fragment or gave a product of the expected 1.7 kb. The former result was obtained for the MS lines, the latter for the maintainers.

Primers atp8-f and O_atp8-f_scar yielded a PCR product of the expected size and specificity – it was 1.8 kb long and appeared only in the five analyzed O-type lines.

Primer pair atp9-f2/O_atp9-f2_38_scar did not amplify the expected 1.6 kb product. Instead, weak amplification of the 2.2 kb product was observed in all analyzed samples.

No amplification was observed after the use of primers atp9-f2 and O_atp9-f2_39_scar (gel picture not shown).

Primers cob-r1 and O_cob-r1_scar produced multi-band profiles. A product of the expected size – 0.9 kb – was obtained for all analyzed male-fertile lines. It was also observed in the two male-sterile lines – 218A and 375A – however, in this case

the efficiency of the amplification was much weaker. All MS lines exhibited the presence of a 2.2 kb DNA fragment which was absent in the maintainers. The remaining DNA fragments were amplified less efficiently and four of them showed plasmotype-dependent distribution. Products of 1.7 and 3.2 kb were observed in all analyzed O-type lines while these fragments were missing in the MS lines. Reversed distribution/specificity was noted for products of 3 and 4.4 kb.

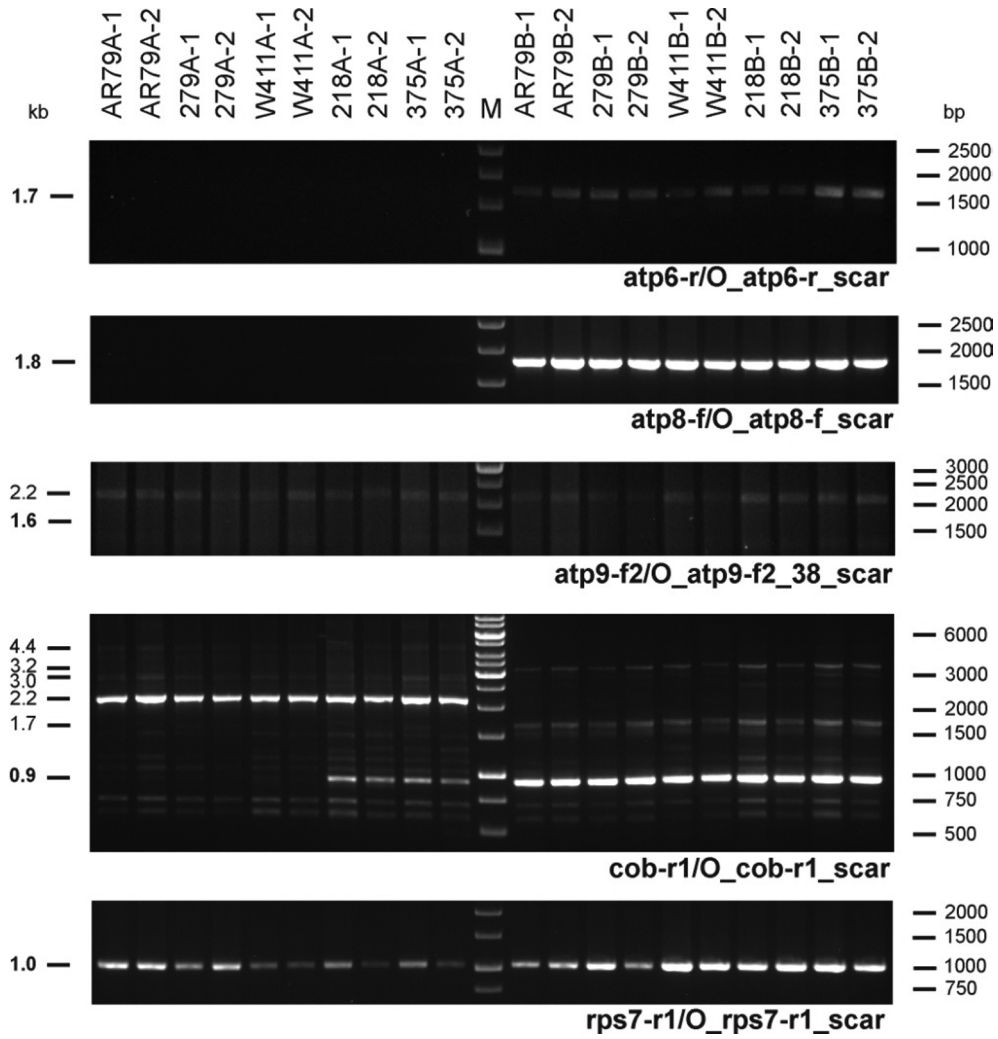


Fig. 32. PCR products generated for the table beet male-sterile lines and their maintainers with primer pairs atp6-r/O_atp6-r_scar to rps-r1/O_rps7-r1_scar. The first primer of each pair was a universal mitochondrial primer (UMP), the second was anchored within the corresponding vectorette marker. The sizes of the expected products are shown in bold on the left

A single monomorphic product was obtained with the use of primers rps-r1 and O_rps7-r1_scar. Its size – 1 kb – was accordant with the sequence data. A generally stronger amplification of this product was observed in the O-type lines.

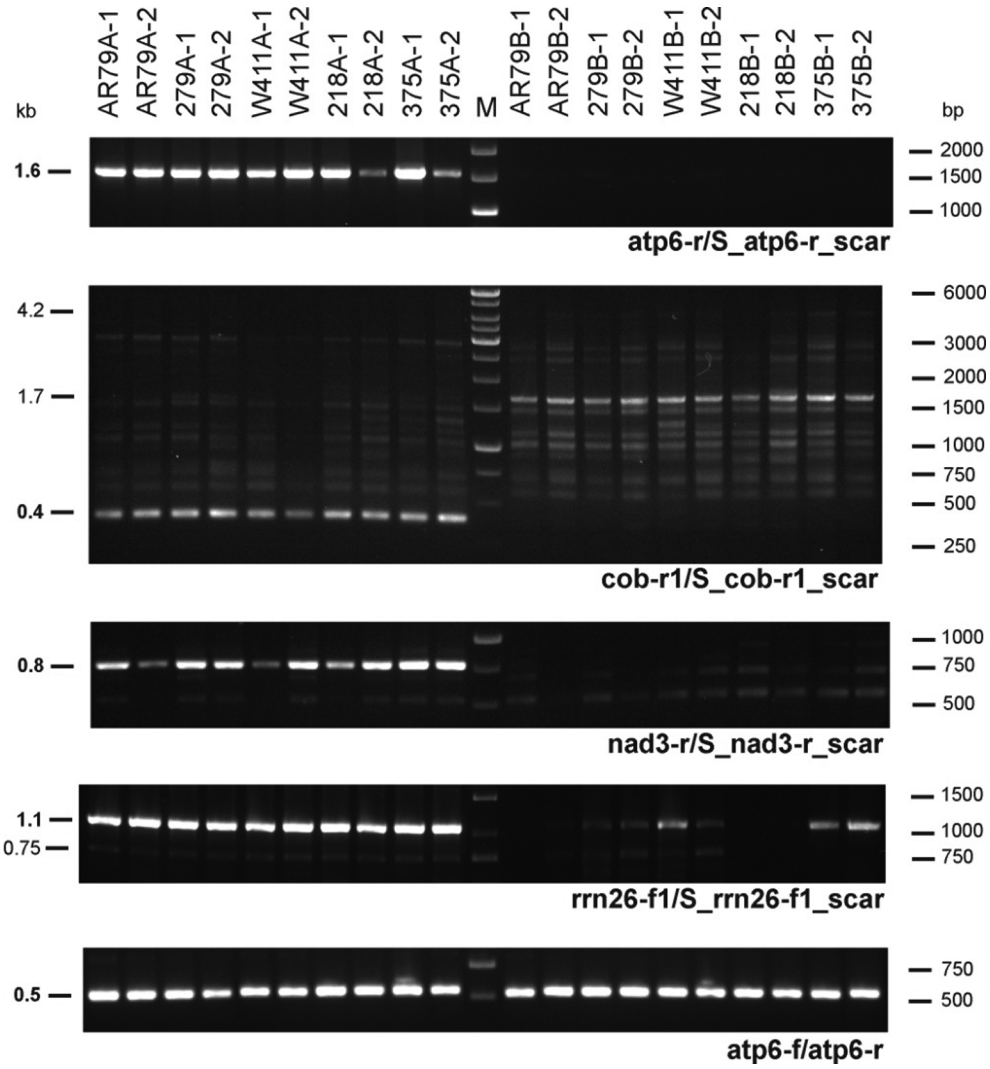


Fig. 33. PCR products generated for the table beet male-sterile lines and their maintainers with primer pairs atp6-r/S_atp6-r_scar to rrn26-f1/S_rrn26-f1_scar and with the control primers (atp6-f, atp6-r). Both control primers were universal mitochondrial primers (UMPs). In the remaining primer pairs the first primer was a universal mitochondrial primer (UMP) and the second was anchored within the corresponding vectorette marker. The sizes of the expected products are shown in bold on the left

PCR with primers atp6-r and S_atp6-r_scar yielded a PCR product of the expected size and distribution – it was 1.6 kb long and strictly specific to the male-sterile lines.

Amplification with primers cob-r1 and S_cob-r1_scar resulted in multi-band profiles for both the male-sterile and male-fertile lines. The major product present in all MS lines had an expected size of 0.4 kb – this DNA fragment was missing in the profiles of the O-type lines. Instead, each of these lines exhibited the presence of a 1.7 kb product. The PCR profiles of the MS lines were devoid of this 1.7 kb product. Moreover, most of the O-type plants, but not the MS plants, exhibited very weak amplification of the 4.2 kb fragment. The remaining products generated with this primer pair did not show the plasmotype-dependent distribution.

According to expectations, primers nad3-r and S_nad3-r_scar generated an MS-specific product of 0.8 kb. Another two DNA fragments obtained with this primer pair were much more weakly amplified and did not show the plasmotype-specific distribution.

The major PCR product generated with primers rrn26-f1 and S_rrn26-f1_scar had an expected size of 1.1 kb. It was present in all analyzed MS lines as well as in AR79B (only in one plant), 279B, W411B and 375B. For the maintainers, the efficiency of its amplification was variable e.g. in the above mentioned AR79B plant, the respective band was barely visible while in 375B it was pronounced but still markedly weaker than the one from the MS lines. In the MS lines as well as in 279B and W411B yet another product was visible – it was a very weakly amplified fragment of 0.75 kb.

The control reaction with the *atp6* primers (atp6-f and atp6-r) yielded an expected fragment of 0.5 kb for all analyzed samples.

Altogether, out of the thirteen tested primer pairs, the plasmotype-dependent product distribution was found for ten and eight of them for sugar beet and table (red) beet, respectively (Table 7, barely visible products were not taken into account). In both species, primer pairs O_A08_scar-f/O_A08_scar-r, rps7-r1/O_rps7-r1_scar and atp9-f2/O_atp9-f2_38_scar did not yield products of diagnostic value. In the case of table beet, this was also true for primers O_B05_scar-f/O_B05_scar-r and atp9-f2/O_atp9-f2_39_scar. In addition to products of the expected size, the plasmotype-dependent distribution was observed for some additional DNA fragments – eleven in sugar beet and eight in table beet (all of the latter were represented among the former). The majority were generated using the *cob*-related primer pairs (cob-r1/O_cob-r1_scar and cob-r1/S_cob-r1_scar) and exhibited reduced amplification efficiency. In the case of table beet, the products with plasmotype-dependent distribution usually did not show the intra-line (qualitative) polymorphism. The only exception was noted for the 4.2 kb fragment which was very weakly amplified in the maintainer lines with primers cob-r1 and S_cob-r1_scar.

Table 7. Summary of the PCR products showing plasmotype-dependent distribution among the tested sugar beet and table (red) beet breeding stocks. The sizes of the expected products are written in bold

Primer pair	Product size [kb]	Presence in	
		S-cytoplasm	N-cytoplasm
Sugar beet			
O_B05_scar-f/O_B05_scar-r	2.1	–	–/+/+
	1.7	–	–/+/+
	1.15	–	–/+
S_B01_scar-f/S_B01_scar-r	2.0	+	–
	1.7	+	–
atp6-r/O_atp6-r_scar	**1.7**	–	+/+
atp8-f/O_atp8-f_scar	**1.8**	– /+	+
atp9-f2/O_atp9-f2_39_scar	**1.5**	–	–/+
cob-r1/O_cob-r1_scar	4.4	+	–
	3.2	–	–/+
	2.9	+	–
	2.2	+	–
	1.7	–	–/+
	0.9	–/+	+
atp6-r/S_atp6-r_scar	**1.6**	+	–
cob-r1/S_cob-r1_scar	4.4	–	–/+
	1.7	–	–/+
	0.4	+	–
nad3-r/S_nad3-r_scar	**0.8**	+	–
	0.6	–/+	+
rrn26-f1/S_rrn26-f1_scar	**1.1**	+	–/+
Table beet			
S_B01_scar-f/S_B01_scar-r	2.0	+	–
	1.7	+	–
atp6-r/O_atp6-r_scar	**1.7**	–	+
atp8-f/O_atp8-f_scar	**1.8**	–	+
cob-r1/O_cob-r1_scar	4.4	+	–
	3.2	–	+
	3.0	+	–
	2.2	+	–
	1.7	–	+
	0.9	–/+	+

Table 7. cont.

Primer pair	Product size [kb]	Presence in	
		S-cytoplasm	N-cytoplasm
atp6-r/S_atp6-r_scar	**1.6**	+	–
cob-r1/S_cob-r1_scar	4.2	–	–/+
	1.7	–	+
	0.4	+	–
nad3-r/S_nad3-r_scar	**0.8**	+	–
rrn26-f1/S_rrn26-f1_scar	**1.1**	+	–/+

+/+ a pronounced difference in the efficiency of the amplification

4.6. Searching for the plasmotype-specific mitochondrial proteins

In the experiments the mitochondrial protein preparations were subjected to isoelectric focusing followed by SDS-PAGE electrophoresis. In a single experiment two plants were analyzed – one from a male-sterile (MS) line and one from the respective maintainer (O-type) line. For one pair of lines either six or four such experiments were performed – in each different plants were tested. The two-dimensional protein separations obtained in a given experiment were examined with respect to the presence of protein spots which either qualitatively or quantitatively differentiated the male-sterile from the male-fertile plant. Quantitative differences were considered significant if the observed spot intensities (% Vol values calculated by Melanie software) of a given protein differed at least twofold. A difference (either qualitative or quantitative) was considered repeatable if it was observed in at least half of the experiments performed for a given pair of lines.

4.6.1. Sugar beet – lines FMS 121 and FO 121

Six experiments were performed for these lines – in each experiment one plant from FMS 121 and one plant from FO 121 were analyzed. On average, silver staining enabled the visualization of 138 and 135 proteins per plant for FMS 121 and FO 121, respectively. Altogether, five proteins differentiated FMS 121 and FO 121 (Figs. 34 and 35, Table 8). In four experiments protein no. 1 showed a greater accumulation in the male-sterile line (FMS 121). In at least three of the experiments performed, proteins no. 2, 3 and 4 were more abundant or present exclusively in the male-fertile line (FO 121). Protein no. 5 in three of the experiments appeared only in FMS 121.

Table 8. Mitochondrial proteins which differentiated lines FMS 121 and FO 121 in the IEF/SDS-PAGE separations

Protein no.	Molecular weight [kDa]	Isoelectric point (pI)	Type of differentiation	Repeatability of the difference
1	38	6.3	FMS 121 > FO 121	4/6
2	39	7.6	FMS 121 < FO 121	5/6
3	70	5.3	FMS 121 < FO 121	3/6
4	72	5.9	FO 121, FMS 121 < FO 121	3/6
5	74	5.9	FMS 121	3/6

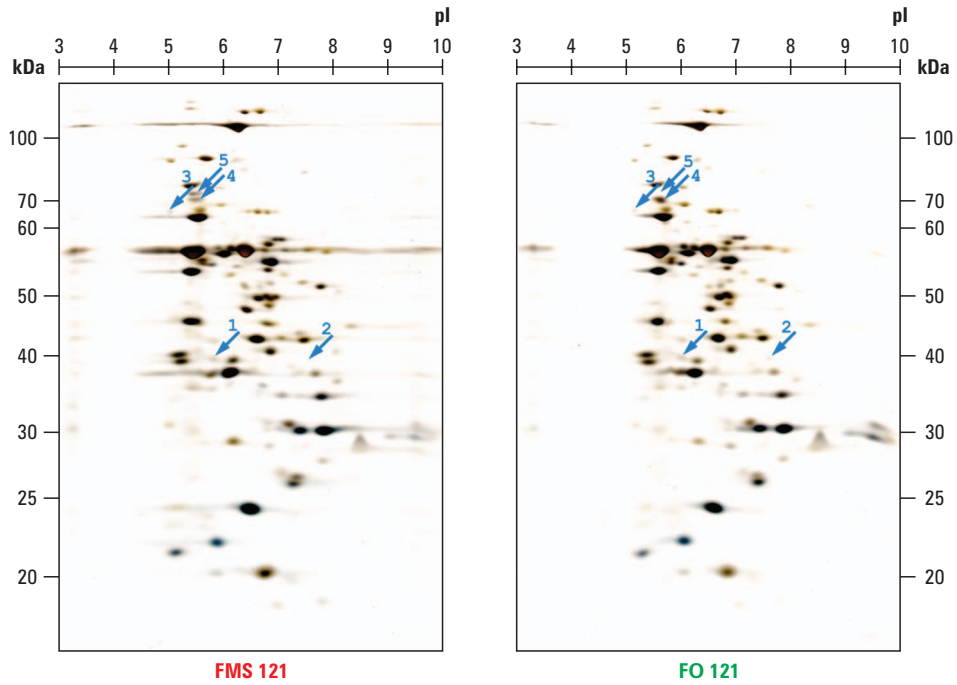


Fig. 34. IEF/SDS-PAGE separation of mitochondrial proteins from single sugar beet plants representing lines FMS 121 and FO 121 (experiment no. 1). Arrows point to the locations of differentiating proteins

Moreover, an interesting correlation was observed in the distribution of proteins 4 and 5 – their sizes were estimated at 72 and 74 kDa, respectively and they both had a pI of 5.9. Four out of the six analyzed FMS 121 plants were characterized by the presence of both these proteins. The remaining two FMS 121 plants had only one – either that of 72 or 74 kDa. In FO 121 the proportion was reversed – only

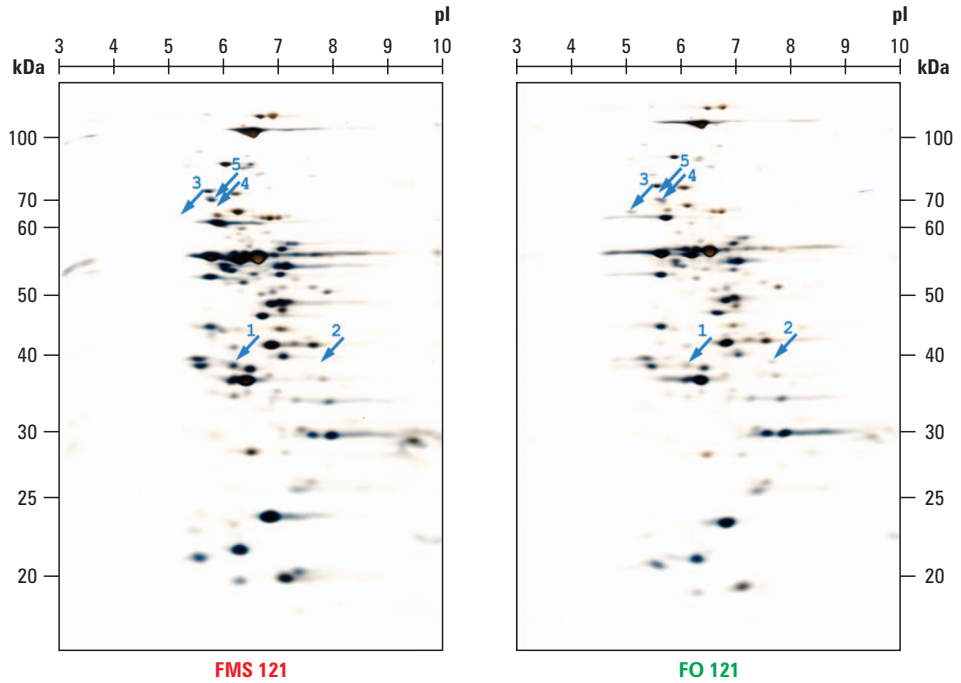


Fig. 35. IEF/SDS-PAGE separation of mitochondrial proteins from single sugar beet plants representing lines FMS 121 and FO 121 (experiment no. 4). Arrows point to the locations of differentiating proteins

Table 9. Distribution of proteins 4 and 5 among the analyzed plants from lines FMS 121 and FO 121

Protein no. (mol. weight)		No. of plants
4 (72 kDa)	5 (74 kDa)	
FMS 121		
+	+	4
+	–	1
–	+	1
FO 121		
+	+	1
+	–	3
–	+	2

\+ present
– absent
+/+ difference in accumulation

a single plant showed the presence of both proteins while in the remaining five plants only one was observed: the 72 kDa protein in three plants and the 74 kDa protein in two plants. Generally, and regardless of the analyzed line, when both proteins were present their accumulation was weaker than the accumulation associated with their single occurrence (Figs. 34 and 35, Table 9).

4.6.2. Sugar beet – lines FMS 1218 and FO 1218

Single plants from both lines were studied in each of the four performed experiments. The average number of detected proteins reached 145 for FMS 1218 and 146 for FO 1218. According to the assumed criteria, eleven proteins showed accumulation differences between the analyzed lines (Figs. 36 and 37, Table 10). Among them were three proteins which also differentiated FMS 121 and FO 121. However, it was only protein 3, for which the type of differentiation was similar to that observed for FMS 121 and FO 121 – in half of the performed experiments its accumulation was stronger in the O-type line. For proteins 4 and 5 the observed type of differentiation was opposite to that observed for lines FMS 121 and FO 121. Protein 4 showed an association with the MS line – in two experiments it was found exclusively in FMS 1218. In turn, protein 5 in two experiments showed a stronger accumulation in FO 1218.

Table 10. Mitochondrial proteins which differentiated lines FMS 1218 and FO 1218 in the IEF/SDS-PAGE separations

Protein no.	Molecular weight [kDa]	Isoelectric point (pI)	Type of differentiation	Repeatability of the difference
3	70	5.3	FMS 1218 < FO 1218	2/4
4	72	5.9	FMS 1218	2/4
5	74	5.9	FMS 1218 < FO 1218	2/4
6	104	6	FMS 1218 < FO 1218	3/4
7	84	6.7	FMS 1218 < FO 1218	3/4
8	78	5.2	FMS 1218 < FO 1218	3/4
9	18	7	FMS 1218	3/4
10	55.1	7	FMS 1218	3/4
11	51.8	7.9	FMS 1218	3/4
12	55.2	7.3	FO 1218, FMS 1218 < FO 1218	3/4
13	50.9	7.8	FO 1218, FMS 1218 < FO 1218	4/4

From among the remaining differentiating proteins, five – numbered 6, 7, 8, 12 and 13 – were also preferentially accumulated in the maintainer line. Depending on the protein, this was shown in three or four performed experiments. In most cases these proteins were more abundant in FO 1218. In one experiment proteins 12 and 13 were detected only in the FO 1218 plant. In the previous experimental series (performed for FMS 121 and FO 121) proteins 6, 7 and 8 did not show apparent regularities in their distribution, while proteins 12 and 13 were not observed at all.

Finally, three differentiating proteins – 9, 10 and 11 – showed association with the male-sterile line. In three performed experiments these proteins were found exclusively in FMS 1218. In single experiments proteins 9, 10 and 11 were not detected – either in FMS 1218 or in FO 1218. Interestingly, in the analyses of FMS 121 and FO 121, protein 9 was also observed exclusively in the MS plants – but only in two out of six experiments and therefore, it was not counted as differentiating. Protein 10 was detected both in FMS 121 and FO 121 – however, its accumulation in these lines did not show apparent regularities. In the FMS 121/FO 121 experiments, protein 11 was either not detected, was accumulated comparably in these accessions or was stronger in the MS line – but with the latter status observed in only one experiment.

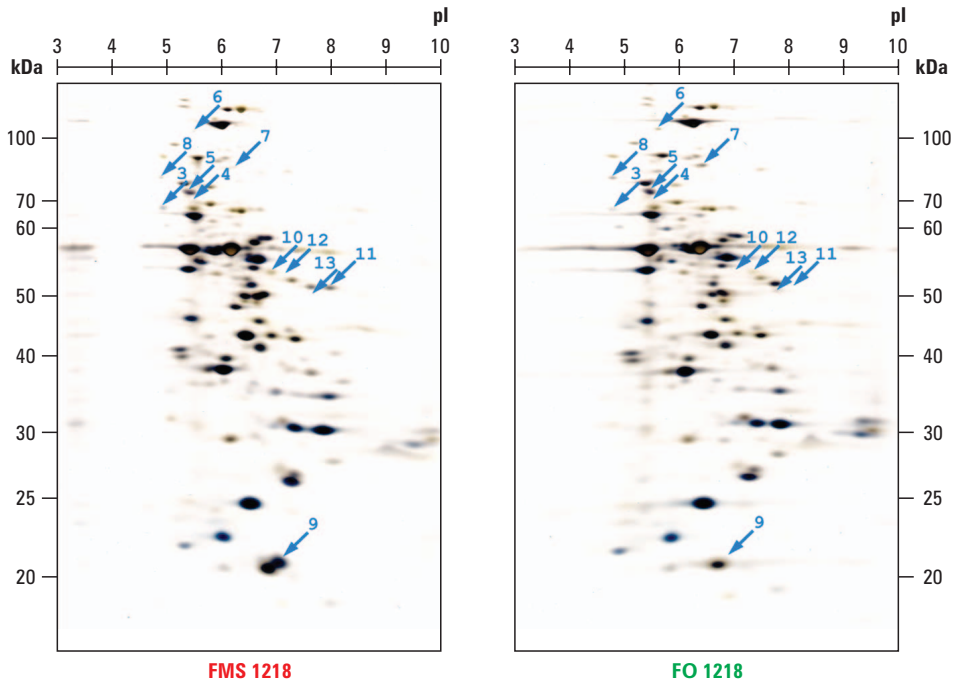


Fig. 36. IEF/SDS-PAGE separation of mitochondrial proteins from single sugar beet plants representing lines FMS 1218 and FO 1218 (experiment no. 1). Arrows point to the locations of differentiating proteins

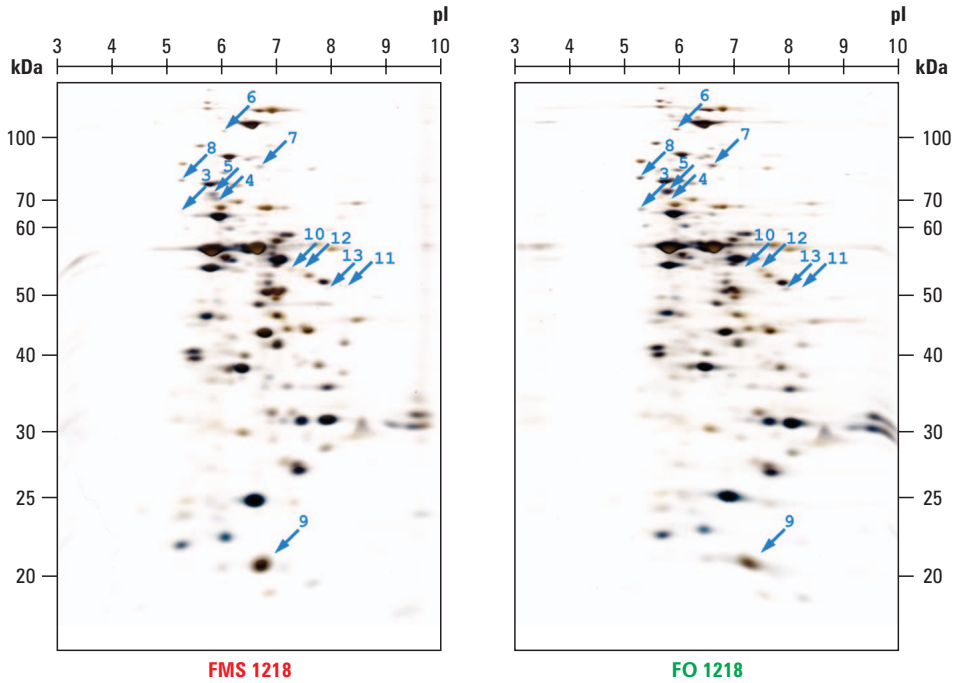


Fig. 37. IEF/SDS-PAGE separation of mitochondrial proteins from single sugar beet plants representing lines FMS 1218 and FO 1218 (experiment no. 3). Arrows point to the locations of differentiating proteins

Table 11. Distribution of proteins 4 and 5 among the analyzed plants from lines FMS 1218 and FO 1218

Protein no. (mol. weight)		No. of plants
4 (72 kDa)	5 (74 kDa)	
FMS 1218		
+	+	2
–	+	2
FO 1218		
–	+	4

\+ present
– absent
+/+ difference in accumulation

In accordance with the FMS 121/FMO 121 experiments the correlation in the distribution of proteins 4 and 5 was also observed for FMS 1218 and FO 1218. In two of the four analyzed FMS 1218 plants both proteins were detected. The remain-

ing FMS 1218 plants exhibited only the presence of protein 5. In all analyzed FO 1218 plants only one of these proteins was observed – protein 5. Similarly, as for the FMS 121/FO 121 pair, when both proteins were present their accumulation was weaker than the accumulation associated with their single occurrence (Figs. 36 and 37, Table 11).

4.6.3. Table beet – lines Re 1A and Re 1B

The analysis was performed for six plants from each line. Silver staining visualized an average of 115 proteins in a single plant. Three proteins displayed repeatable accumulation differences between Re 1A and Re 1B – these were proteins 2, 6 and 13 which already appeared differentiating in the set of sugar beet accessions (see above) (Figs. 38 and 39, Table 12). In three of the performed experiments, proteins 2 and 6 had a stronger accumulation in the maintainer line. Protein 13 showed association with the MS line – in two experiments it was detected only in Re 1A, in another two experiments its accumulation in Re 1A was stronger. Therefore, the differentiating character of proteins 2 and 6 was accordant with that observed in

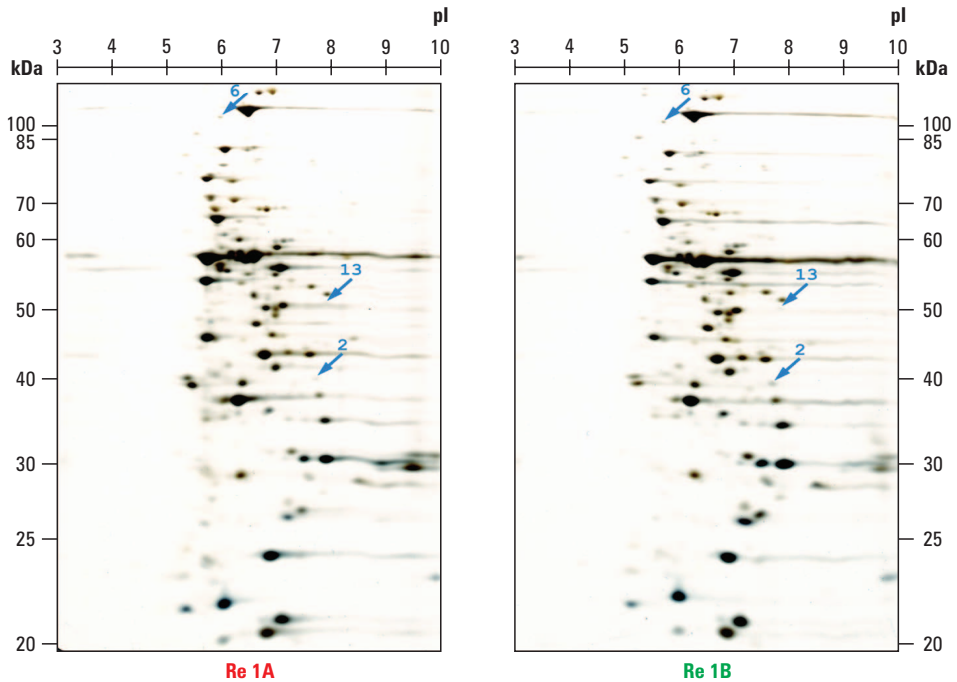


Fig. 38. IEF/SDS-PAGE separation of mitochondrial proteins from single table beet plants representing lines Re 1A and Re 1B (experiment no. 3). Arrows point to the locations of differentiating proteins

sugar beet. This was not the case for protein 13 which in one experimental series (FMS 1218/FO 1218) showed an association with the maintainer line.

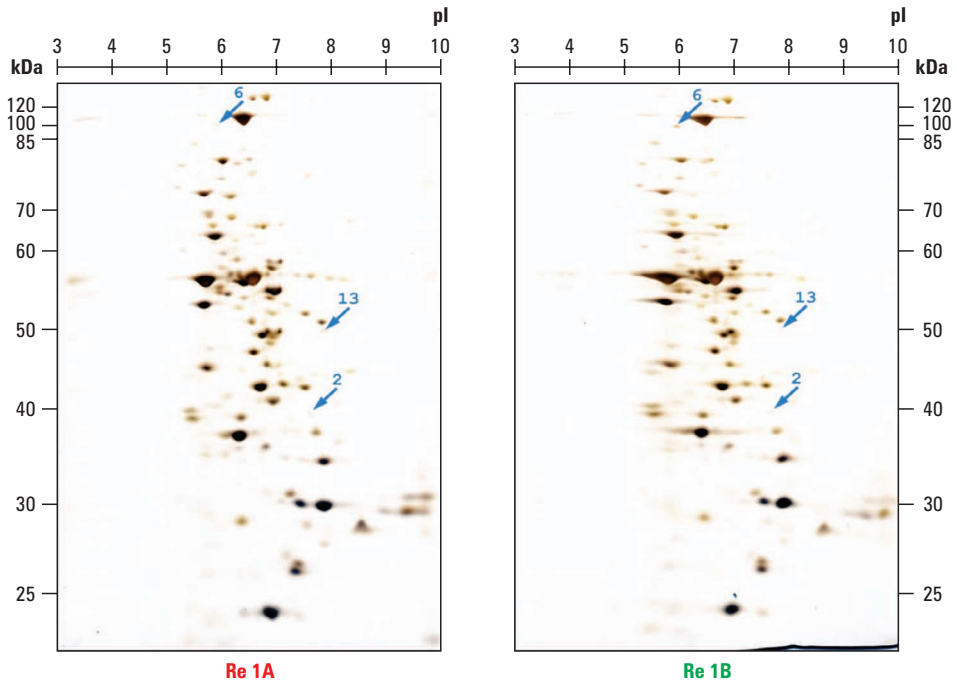


Fig. 39. IEF/SDS-PAGE separation of mitochondrial proteins from single table beet plants representing lines Re 1A and Re 1B (experiment no. 4). Arrows point to the locations of differentiating proteins

Table 12. Mitochondrial proteins which differentiated lines Re 1A and Re 1B in the IEF/SDS-PAGE separations

Protein no.	Molecular weight [kDa]	Isoelectric point (pI)	Type of differentiation	Repeatability of the difference
2	39	7.6	Re 1A < Re 1B	3/6
6	104	6	Re 1A < Re 1B	3/6
13	50.9	7.8	Re 1A, Re 1A > Re 1B	4/6

Taking into account the correlation in the distribution of sugar beet proteins 4 and 5, their accumulation was also looked at in table beet. In all of the performed experiments, Re 1A and Re 1B displayed an identical electrophoretic pattern of

these proteins – they either migrated in a very close proximity (in three experiments) or apparently co-migrated forming a single spot (in the other three experiments). In the former case their molecular weight was estimated at 68 and 69 kDa, in the latter – at 70 kDa (Figs. 38 and 39).

4.6.4. Table beet – lines Re 3A and Re 3B

The analysis was also performed for six plants from each line. The average number of visualized proteins (per plant) was 140 for Re 3A and 142 for Re 3B. Only two proteins which repeatedly differentiated Re 3A and Re 3B were found. These were proteins 3 and 6 which had already appeared differentiating in the set of the analyzed sugar beet accessions (see above) (Figs. 40 and 41, Table 13). Both proteins showed their differentiating character in three experiments. In these experiments protein 3 was qualitatively specific for the maintainer line, while protein 6 showed a stronger accumulation in the MS line. The behavior of protein 3 is similar to that observed in both sugar beet MS/O-type pairs (in the Re 1A/Re 1B experiments this protein did not show up as differentiating). In turn the behavior of protein 6 is contrary to

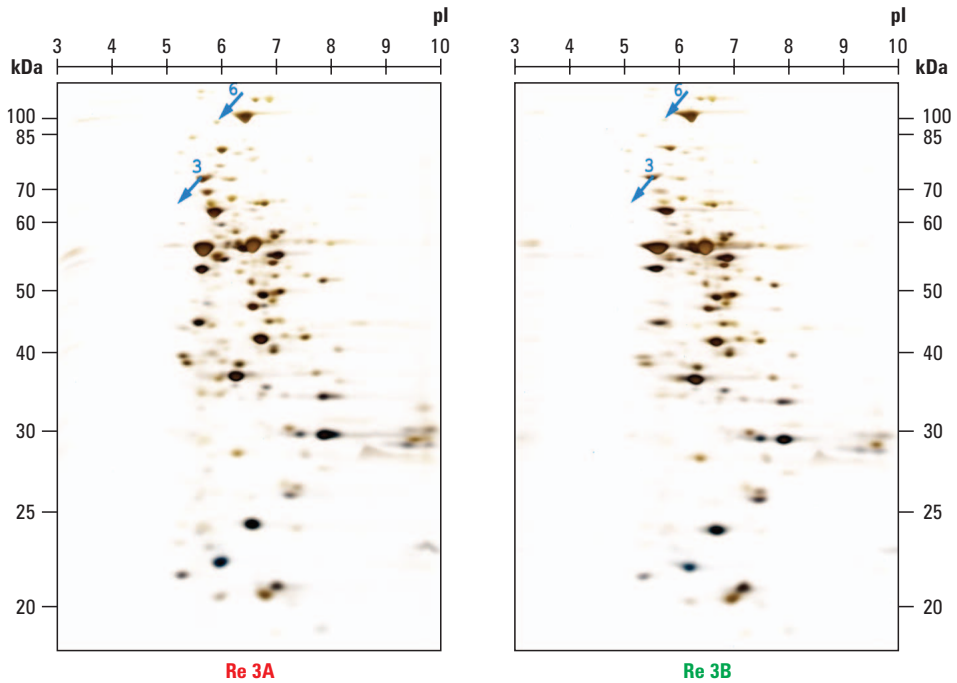


Fig. 40. IEF/SDS-PAGE separation of mitochondrial proteins from single table beet plants representing lines Re 3A and Re 3B (experiment no. 2). Arrows point to the locations of differentiating proteins

that seen in pairs FMS 1218/FO 1218 and Re 1A/Re 1B (in the FMS 121/FO 121 analyses protein 6 was not recognized as polymorphic).

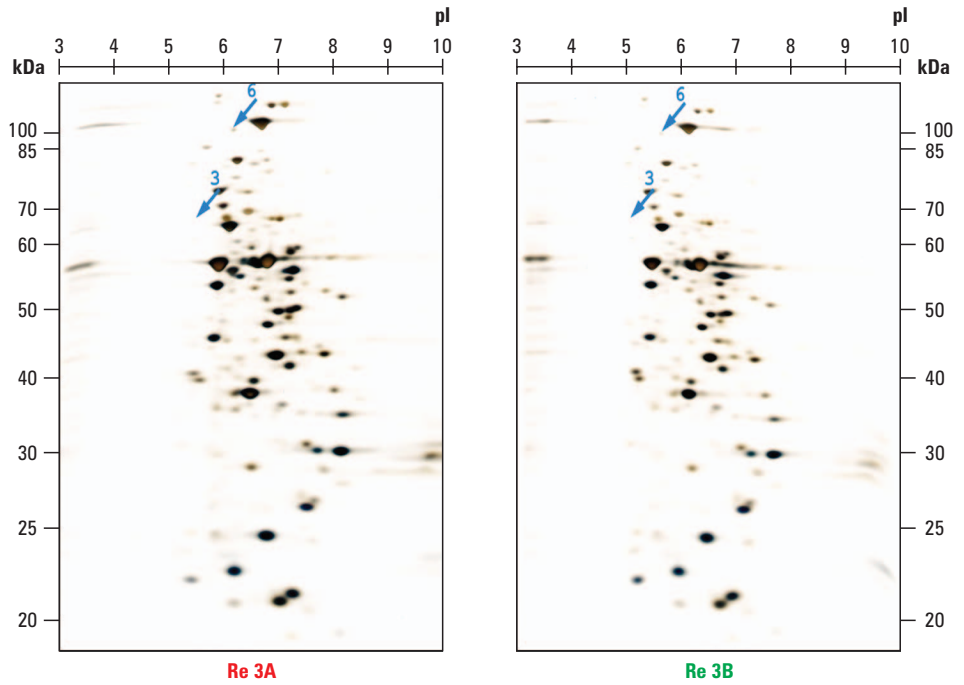


Fig. 41. IEF/SDS-PAGE separation of mitochondrial proteins from single table beet plants representing lines Re 3A and Re 3B (experiment no. 4). Arrows point to the locations of differentiating proteins

Table 13. Mitochondrial proteins which differentiated lines Re 3A and Re 3B in the IEF/SDS-PAGE separations

Protein no.	Molecular weight [kDa]	Isoelectric point (pI)	Type of differentiation	Repeatability of the difference
3	70	5.3	Re 1B	3/6
6	104	6	Re 1A > Re 1B	3/6

In all experiments the proteins corresponding to sugar beet proteins 4 and 5 apparently co-migrated forming a single spot on the level of 69 kDa (Figs. 40 and 41).

4.6.5. Protein identification

The proteins which differentiated the male-sterile and male-fertile accessions of sugar beet were subjected to LC-MS/MS analysis for the purpose of identification. This analysis was also performed for one non-differentiating protein – designated as protein 19 – which on the 2-DE separations neighbored or even overlapped protein 9. The reasonable mass spectrometry data were obtained for the following twelve proteins: 1, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 19 (Table 14).

Table 14. Results of protein identification

Protein no.	Origin	Accession no.	Protein name (species)	MOWSE score	No. of peptides matched	Sequence coverage [%]	Molecular weight [kDa]		Isoelectric point (pI)	
							Theoretical	Observed	Theoretical	Observed
1	FO 1218	gi|75282478	bifunctional L-3-cyanoalanine synthase/cysteine synthase, mitochondrial (*Spinacia oleracea*)	738	4	10	39.6	38	8.7	6.3
4	FO 121	gi|2493121	V-type proton ATPase catalytic subunit A (*Beta vulgaris*)	1044	26	39	68.8	72	5.1	5.9
5	FO 1218	gi|2493121	V-type proton ATPase catalytic subunit A (*Beta vulgaris*)	567	10	18	68.8	72	5.1	5.9
6	FO 1218	gi|4105131	ClpC protease (*Spinacia oleracea*)	2672	29	32	99.6	104	8.8	6.0
7	FO 1218	gi|674244562	transketolase (*Arabis alpina*)	413	6	8	79.8	84	5.85	6.7
8	FO 1218	gi|587840123	stromal 70 kDa heat shock-related protein (*Morus notabilis*)	1326	15	22	75.5	78	5.3	5.2
9	FMS 121	gi|255570875	CBS domain-containing protein CBSX3, mitochondrial-like (*Ricinus communis*)	476	7	18	22.9	18.0	9.2	7.0

Table 14. cont.

10	FMS 1218	gi|503716737	cysteine desulfurase (*Candidatus Midichloria mitochondrii*)	487	2	9	44.2	55.1	6.8	7.0
–//–	–//–	gi|661894213	cysteine desulfurase (*Coffea canephora*)	346	6	17	50.1	55.1	6.85	7.0
11	FMS 1218	gi|565376407	alanine-glyoxylate aminotransferase 2 homolog 1, mitochondrial-like (*Solanum tuberosum*)	421	4	9	52.0	51.8	8.3	7.9
12	FO 1218	gi|604346038*	cysteine desulfurase (*Erythranthe guttata*)	554	5	25	30.4	55.2	9.1	7.3
–//–	–//–	gi|294461769	cysteine desulfurase 1, mitochondrial-like (*Picea sitchensis*)	463	6	16	50.3	55.2	6.9	7.3
13	FO 1218	gi|460404718	3-ketoacyl-CoA thiolase 2, peroxisomal-like (*Solanum lycopersicum*)	857	6	18	49.6	50.9	8.0	7.8
19**	FMS 1218	gi|586679367	CBS domain-containing protein CBSX3, mitochondrial (*Amborella trichopoda*)	2646	7	18	27.8	17.1	8.65	6.8

* partial sequence

** protein neighboring/overlapping protein 9

5. Discussion

The present work shows the use of two alternatives for the identification of plasmotype-specific markers – RAPDs and vectorette PCR. Both techniques were introduced in 1990, the former – as yet another method to generate genetic markers [Williams et al. 1990], the latter – as a method to isolate terminal sequences of DNA cloned in yeast artificial chromosomes (YACs) [Riley et al. 1990]. Although in principle RAPDs were designed to track differences in the nuclear DNA, the use of short arbitrary primers did not exclude the concurrent amplification of organellar (mitochondrial and plastid) sequences, which resulted in the occasional use of RAPDs to search for CMS-associated mtDNA fragments [Nakajima et al. 1997, Ivanov et al. 2002]. On the contrary, there have been no reports concerning the application of vectorette PCR in a similar context. Therefore, the present work extends the traditional concept of vectorette PCR – as a means to isolate unknown sequences neighboring already known DNA regions – to a method of variation analysis based on organellar DNA. In this new context, vectorette PCR shows an important advantage over RAPDs – this is genome selectivity provided by the use of gene-specific primers, here represented by a set of universal mitochondrial primers (UMPs). Such selectivity was clearly evidenced by the fraction of plasmotype-specific markers within the pool of all polymorphic markers – in the case of vectorette PCR this fraction was ten times larger than for RAPDs (56.8% for vectorette PCR vs. 5.4% for RAPDs). This in turn influenced the pattern in which the beet accessions were differentiated with the use of markers of both types. The principal component analysis (PCA) revealed that only vectorette PCR markers grouped the analyzed accessions according to their plasmotypes (Figs. 1 and 7).

A bulk of interesting conclusions and generalizations can be drawn from the results of homology searches performed for the cloned and sequenced RAPD/vectorette markers. First of all, it was interesting to see whether the marker sequences were represented in the reference sugar beet mitochondrial genomes available for lines TK81-O and TK81-MS [Kubo et al. 2000, Satoh et al. 2004]. According to expectations, entire (undivided) sequences of the N- and S-cytoplasmic RAPD/

vectorette markers should be present in mtDNA of TK81-O and TK81-MS, respectively. This was confirmed for three markers of the normal cytoplasm – N_atp8-f, N_cob-r1 and N_rps7-r1; as well as for five markers of the sterilizing cytoplasm – S_atp6-r, S_cob-r1, S_nad3-r, S_rrn26-f1 and S_B-01_c29. However, entire (undivided) sequences of another three N-cytoplasmic markers were not found in TK81-O mtDNA. Of these markers: the whole N_atp6-r sequence was represented in the reference genome – but as two discontinuous blocks; for N_atp9-f2 homology with TK81-O mtDNA was restricted to five short blocks – therefore, only 14.8% of the marker sequence (version c39) was represented in the reference genome; N_A-08 did not show any similarity to TK81-O mtDNA. On the other hand entire continuous sequences of these three markers were represented in mtDNA of wild beet accessions – either in *Beta macrocarpa* (N_atp6-r) or in *Beta vulgaris maritima* (N_atp9-f2, N_A-08). This observation indicates that complete plant mitochondrial genome sequences deposited in the databases in fact may be devoid of some mtDNA regions. This may be related to poor clonability of such regions resulting in their underrepresentation within the sequenced genomic libraries. Alternatively, one can assume that such regions had been present in the common ancestor of wild and cultivated beets and that subsequently they were lost in some cultivated germplasm.

It needs to be emphasized that it is likely that the majority of point sequence differences observed between the RAPD/vectorette markers and the corresponding sequence blocks from the reference genomes may be attributed to copying errors introduced by *Taq* polymerase. The high error rate of this enzyme results from the lack of a 3'→5' exonuclease (proofreading) activity – it is estimated that after a 30-cycle PCR one error occurs every 300 bp [Brown 2010]. Even so, it is not a problem for the direct sequencing of PCR products – errors in a particular position are restricted only to a small fraction of molecules and due to this *dilution* the respective changes are not manifested in sequence chromatograms. This situation changes dramatically upon prior cloning of amplification products. The cloning process picks up a specific product molecule together with its particular errors. Thus, in a given plasmid clone, errors are fixed – they are present in all sequenced molecules, which is clearly reflected in sequence chromatograms. All marker sequences presented in this work were obtained with prior cloning of the respective PCR products – this option offered better quality of the resulting sequence chromatograms but, as a consequence, care must be taken when attention is directed at some specific nucleotide positions e.g. in processes such as ORF searching or SNP marker design. Obviously, it cannot be excluded that some of the observed point sequence differences reflect genuine polymorphisms differentiating the accessions for which the marker products were amplified and the accessions used as a source for the reference genomes.

Another possible type of PCR-borne sequence artifacts is the formation of chimeric products – either by template switching or through incomplete primer extension [Kanagawa 2003]. It is likely that this type of artifact is not represented

among the RAPD and vectorette markers reported in this study. Such a conclusion is justified by the fact that for all but one (N_B-05) of the markers the corresponding mtDNA sequence blocks were found – if not in the sugar beet reference accessions (TK81-O, TK81-MS), at least in wild beets – *Beta vulgaris maritima* and *Beta macrocarpa* (see above). However, the sequence analysis of two N_atp9-f2 markers revealed another possible source of artifacts, namely – the formation of composite ligation products made from the vectorette unit and more than one restriction fragment. Such restriction fragments, brought together in a ligation event, will simulate physical linkage of their sequences. In the N_atp9-f2_c38 case, artifact identification was relatively easy due to the different organellar origins of the component fragments. When both (all) fragments are of mitochondrial origin, chimerism can be easily overlooked – although its occurrence may be suspected whenever a restriction site for the enzyme used for vectorette library preparation is present and whenever this is accompanied by a failure in PCR verification of sequence arrangement.

Surprisingly, the sequence examination of the vectorette markers revealed that the majority of them were not anchored within the target coding sequences as was expected on the basis of custom primer (UMP) design. Out of nine sequenced markers only three – N_atp6-r, S_atp6-r and S_nad3-r – had the expected coding segments. The remaining six vectorette markers – N_atp8-f, N_cob-r1, N_rps7-r1, N_atp9-f2, S_cob-r1 and S_rrn26-f1 – were anchored within non-coding or probably non-coding sequences (N_atp8-f was anchored within an ORF of non-confirmed functionality). Moreover, for three of the latter markers – N_cob-r1, S_cob-r1 and S_rrn26-f1 – the homologous blocks from the corresponding reference genomes did not included 5' parts of the respective custom primers (UMPs). This implies that only partial primer-template complementarity, limited to the 3' primer termini, was sufficient to generate these three vectorette markers. Taken together, these observations indicate that the non-coding sequences of beet mtDNA are seeded with short (comparable to primer size) coding sequence homologies, abundant enough to be targeted by the vectorette PCR assay. The possible high incidence of such sequence stretches also becomes evident upon comparison of both studied N_atp9-f2 marker versions – they resulted from the use of alternative priming sites separated by only 4 bp and displaying homology to the same fragment of the *atp9* gene.

Generally, when Illumina sequence reads generated for the MS line were mapped to the S-cytoplasm-specific vectorette markers their continuous coverage was observed, while mapping to the N-cytoplasmic markers showed both covered and uncovered regions. Accordingly, when the sequence reads generated for the O-type line were used for mapping, the continuous coverage was observed for the N-cytoplasmic vectorette markers, while for the S-cytoplasmic markers a pattern of covered and uncovered regions was obtained. Two important implications may be drawn from these observations. First, the sequence read mapping confirms specificity of the analyzed markers since the continuous coverage proves physical existence

of a genomic fragment, while the fragmentary coverage excludes this possibility indicating that only some unlinked marker regions are represented in the sequenced mtDNA. Secondly, these data point at the possibility of a more generalized approach towards the identification of plasmotype-specific markers. This approach would involve NGS read mapping to assembled mitochondrial genomes – if such reads and a reference genome represented different mitotypes, the uncovered regions could be used as targets for the design of mitotype-specific SCAR markers. It seems that the proposed approach will become more and more feasible with growing accessibility and the decreasing costs of NGS technologies.

The present work offers a general approach towards the identification of plant plasmotype-specific markers. This approach is based on the use of vectorette PCR and universal mitochondrial primers (UMPs). Moreover, in conclusion of the NGS read mapping an alternative approach was proposed (see the previous paragraph) and is likely to be more comprehensive than the one based on vectorette PCR or any other method used so far. The availability of plasmotype markers can be very advantageous in the breeding practice. This is due to the fact that breeders need to know the plasmotypes of their breeding accessions and the fact that traditional cytoplasm determination, based on test-crossing and phenotypic evaluation of the resulting offspring, is both very laborious and time-consuming. According to Havey [1995] in the case of a biennial crop the cytoplasm identification needs four to eight years. With the availability of adequate PCR markers this period can be shortened to several hours. Such markers can be especially useful in the extraction of maintainers from populations of mixed plasmotype composition [Havey 1995, Stojałowski et al. 2006]. In turn, the selection of maintaining genotypes is one of the most crucial challenges in the course of hybrid breeding which to a large extent determines the time required to produce a new variety – in the case of sugar beet it is roughly 20–25 years (Adam Sitarski, KHBC Straszków, personal communication). Contamination of maintainers with S-cytoplasmic plants is likely to be associated with the carry-over of restoring alleles which will negatively influence the phenotypic stability of the respective CMS lines. The quality of maternal components can also be directly affected by an admixture of plants carrying the normal cytoplasm. Regardless of the causal factor, the lowered quality of maternal components will inevitably compromise the quality of the resulting hybrid since the seed harvested from a CMS line will be produced not only from cross-pollination (as it should be) but also from self- and sib-pollination. The above indicated negative effects can be minimized by the routine use of plasmotype-specific DNA markers. As an outcome of this study a toolkit of five mitochondrial SCAR markers was elaborated which enables a distinction between the normal and sterilizing cytoplasm of both sugar and table (red) beet. In addition to the accordant indications produced for both types of beet, the markers of this set were characterized by qualitative polymorphism and strong amplification of diagnostic products. These included one N-cytoplasm-specific marker – atp6-r/O_atp6-r_scar, and three

S-cytoplasmic markers – atp6-r/S_atp6-r_scar, cob-r1/S_cob-r1_scar and nad3-r/S_nad3-r_scar. Moreover, another SCAR – cob-r1/O_cob-r1_scar, although designed on the basis of the N-cytoplasm-specific vectorette marker, displayed sui generis co-dominance – one of its major products was specific for the N-cytoplasm (either qualitatively or quantitatively) and the other for the S-cytoplasm (qualitatively). In addition to these five SCAR markers, a further two – atp8-f/O_atp8-f_scar and rrn26-f1/S_rrn26-f1_scar – exhibited quantitative polymorphism. The former showed greater amplification for N-cytoplasmic accessions (in sugar beet, in table beet its polymorphism was accordant but qualitative), while in the case of the latter a higher amplification efficiency was observed for accessions carrying the S-cytoplasm. Taken altogether, the markers of this set offer the possibility of accurate cytoplasm identification within both the sugar beet and table beet germplasm. It also needs to be emphasized that one of the identified vectorette markers – S_atp6-r – contains a putative sterility determinant – the *preSatp6* sequence [Yamamoto et al. 2005]. This fact indicates the potential of the vectorette PCR approach for cloning CMS-associated factors in plants.

Within the male-fertile sugar beets a few markers distinguished two separate groups – the group with and the group without a given marker product (such a product was also absent in the MS lines). This observation applies to the following, predominantly efficiently amplified, markers:

- RAPD – N_A-08 (0.25 kb);
- vectorettes – N_atp6-r (1.65 kb), N_atp9-f2 (1.48 kb) and
- SCARs – O_B05_scar-f/O_B05_scar-r (2.1, 1.7 and 1.15 kb), atp6-r/O_atp6-r_scar (1.7 kb, quantitatively polymorphic), atp9-f2/O_atp9-f2_39_scar (1.5 kb).

Moreover, in the case of vectorette PCR with primer ccb203-f it was not the presence but the absence of the product (1.83 kb) which was plasmotype-specific and distinguished two male-fertile accessions from all those remaining. Although typically, the shared absence of marker bands is regarded as a weaker indicator of genetic similarity than the shared presence of bands [Mohammadi and Prasanna 2003, Bussell et al. 2005], distinguishing these two groups within the N-cytoplasmic germplasm is justified by the accordant indication of several markers – the more so because these markers represent three different experimental approaches: RAPD, vectorette PCR as well as SCAR. It is worth mentioning that whenever the above mentioned markers were analyzed in the same N-cytoplasmic accessions they exhibited accordant segregation (co-segregation). In turn, two vectorette markers indicated mitotypic differentiation within the S-cytoplasm. Lines MSPT 68 and NS 023 were the only ones for which primer nad7ex1-f generated a fragment of 2.05 kb and primer nad9-r did not generate the 1.15 kb product (data not shown). However, both these vectorette PCR products were amplified very weakly and therefore, the presence/absence polymorphism of these DNA fragments (among the S-cytoplasmic accessions) could have resulted from the fact that the respective target amplicons

were present in the template DNA in amounts close to the PCR detection limit generating the randomized output. Likewise, a few SCAR markers showed polymorphism among the S-cytoplasmic accessions (Table 7), albeit these were also minor (weakly amplified) products which, furthermore, did not co-segregate, indicating a low rank for the observed differentiation. These data agree with the commonly accepted view that CMS in beets represents a single source – Owen's cytoplasm [Mglinets and Veprev 1998]. Mitotype differentiation within the normal cytoplasm seems to be unexceptional for beets as it was observed e.g. in carrots [Szklarczyk et al. 2000] and onions [Havey 2000].

In addition to the search for mtDNA polymorphisms male-sterile and male-fertile beets were compared with respect to the IEF/SDS-PAGE-resolved pools of mitochondrial proteins. Although, there were altogether 13 differentiating proteins found, none of them displayed accordant distribution across the four studied genetic backgrounds. In some cases a given protein showed association with male sterility in one genetic background and with male fertility in the other. This indicates that the differentiating character of these proteins represents rather a genotype-specific effect than a true association with the plasmotype. In this context it is not surprising that all identified proteins are encoded in the nucleus and that six of them are not even mitochondrial, but plastid (proteins 6, 7 and 8), vacuolar (proteins 4 and 5) and peroxisomal (protein 13). The presence of non-mitochondrial proteins is likely to be a consequence of other organelle contaminations within the analyzed mitochondrial preparations. From among the identified proteins a clear association with male sterility was reported for CBS domain-containing protein CBSX3 – protein 9. According to the report of Yoo et al. [2011], an overexpression of CBSX1, a chloroplast homolog of CBSX3, caused severe male sterility in *Arabidopsis thaliana*. The sterilizing effect had a form of anther indehiscence attributed to a decreased hydrogen peroxide level causing defective lignin deposition and resulting in impaired secondary wall thickening in the anther endothecium. In sugar beet, protein 9 – when detected – was observed solely in the male sterile lines, which supported its association with the trait – the more so because Kinoshita [1971] had already indicated anther indehiscence as a component of male sterility in this plant. However, examination of the table (red) beet accessions did not back up this view – protein 9 appeared both in the MS and maintainer lines. On the other hand one needs to bear in mind that the mechanisms of CMS expression in table beet may not be identical to those of sugar beet. It is also worth mentioning that differential distribution of protein 9 in sugar beet is possibly not a consequence of its differential expression but rather it results from diverse post-translational modifications (PTMs). In the 2-DE separations this protein always appears in the context of a neighboring protein which has also been subjected to LC-MS/MS analysis and identified as CBSX3 (Table 14, protein 19). Therefore, it seems that it is CBSX3 PTM which differentiates male-sterile and male-fertile sugar beets, but not table beets.

References

Brown T.A. 2010. Gene cloning and DNA analysis: an introduction. Wiley-Blackwell.

Bussell J.D., Waycottb M., Chappill J.A. 2005. Arbitrarily amplified DNA markers as characters for phylogenetic inference. Perspect. Plant Ecol. Evol. Syst., 7, 3–26.

Chomczynski P., Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem., 162, 156–159.

Goldman I.L., Navazio J.P. 2007. Table beet. [In:] Vegetables I: Asteraceae, Brassicaceae, Chenopodicaceae, and Cucurbitaceae. Eds. J. Prohens, F. Nuez. Springer, New York, 219–238.

Hagihara E., Matsuhira H., Ueda M., Mikami T., Kubo T. 2005. Sugar beet BAC library construction and assembly of a contig spanning *Rf1*, a restorer-of-fertility gene for Owen cytoplasmic male sterility. Mol. Gen. Genomics, 274, 316–323.

Hammer Ø., Harper D.A.T., Ryan P.D. 2001. Past: paleontological statistics software package for education and data analysis. Palaeont. Electron., 4, 4.

Havey M.J. 1995. Identification of cytoplasms using the polymerase chain reaction to aid in the extraction of maintainer lines from open-pollinated populations of onion. Theor. Appl. Genet., 90, 263–268.

Havey M.J. 2000. Diversity among male-sterility-inducing and male-fertile cytoplasms of onion. Theor. Appl. Genet., 101, 778–782.

Henry K. 2010. Fodder beet. [In:] Root and tuber crops. Ed. J.E. Bradshaw. Springer, New York, 221–243.

Hjerdin-Panagopoulos A., Kraft T., Rading I.M., Tuvesson S., Nilsson N.-O. 2002. Three QTL regions for restoration of Owen CMS in sugar beet. Crop Sci., 42, 540–544.

Honma Y., Taguchi K., Hiyama H., Yui-Kurino R., Mikami T., Kubo T. 2014. Molecular mapping of *restorer-of-fertility 2* gene identified from a sugar beet (*Beta vulgaris* L. ssp. *vulgaris*) homozygous for the non-restoring *restorer-of-fertility 1* allele. Theor. Appl. Genet., 127, 2567–2574.

Ivanov M.K., Revenko A.S., Kabilov M.R., Dymshits G.M. 2002. The use of RAPD analysis of sugar-beet mitochondrial DNA in search of CMS genes. Doklady Biol. Sci., 387, 536–538.

Jungblut P.R., Seifert R. 1990. Analysis by high-resolution two-dimensional electrophoresis of differentiation-dependent alterations in cytosolic protein pattern of HL-60 leukemic cells. J. Biochem. Biophys. Methods, 21, 47–58.

Kanagawa T. 2003. Bias and artifacts in multitemplate polymerase chain reactions (PCR). J. Biosci. Bioeng., 96, 317–323.

Kaul M.L.H. 1988. Male sterility in higher plants. Springer, Berlin.

Kinoshita T. 1971. Genetical studies on the male sterility of sugar beets (*Beta vulgaris* L.) and its related species. J. Facul. Agr. Hokkaido Univ., 56, 435–565.

Kubo T., Nishizawa S., Sugawara A., Itchoda N., Estiati A., Mikami T. 2000. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys)(GCA). Nucleic Acids Res. 28, 2571–2576.

Lunter G., Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res., 21, 936–939.

Majewska-Sawka A., Rodriguez-Garcia M.I., Nakashima H., Jassen B. 1993. Ultrastructural expression of cytoplasmic male sterility in sugar beet (*Beta vulgaris* L.). Sex Plant Reprod., 6, 22–32.

Matsuhira H., Kagami H., Kurata M., Kitazaki K., Matsunaga M., Hamaguchi Y., Hagihara E., Ueda M., Harada M., Muramatsu A., Yui-Kurino R., Taguchi K., Tamagake H., Mikami T., Kubo T. 2012. Unusual and typical features of a novel restorer-of-fertility gene of sugar beet (*Beta vulgaris* L.). Genetics, 192, 1347–1358.

Mglinets A.V., Veprev S.G. 1998. Creation and study of S-cytoplasm collection from varied sources of sugar beet. Int. Crop Net. Ser., 12, 57–65.

Mikami T., Sugiura M., Kinoshita T. 1984. Molecular heterogeneity in mitochondrial and chloroplast DNAs from normal and male sterile cytoplasms in sugar beets. Curr. Genet., 8, 319–322.

Mohammadi S.A., Prasanna B.M. 2003. Analysis of genetic diversity in crop plants – salient statistical tools and considerations. Crop Sci., 43, 1235–1248.

Nakajima Y., Yamamoto T., Oeda K. 1997. Genetic variation of mitochondrial and nuclear genomes in carrots revealed by random amplified polymorphic DNA (RAPD). Euphytica, 95, 259–267.

Owen F.V. 1942. Male sterility in sugar beets produced by complementary effects of cytoplasmic and Mendelian inheritance. Am. J. Bot., 29, 692.

Owen F.V. 1945. Cytoplasmically inherited male sterility in sugar beets. J. Agr. Res., 71, 423–440.

Riley J., Butler R., Ogilvie D., Finniear R., Jenner D., Powell S., Anand R., Smith J.C., Markham A.F. 1990. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res., 18, 2887–2890.

Santoni S., Berville A. 1992. Two different satellite DNAs in *Beta vulgaris* L.: evolution, quantification and distribution in the genus. Theor. Appl. Genet., 84, 1009–1016.

Satoh M., Kubo T., Nishizawa S., Estiati A., Itchoda N., Mikami T. 2004. The cytoplasmic male-sterile type and normal type mitochondrial genomes of sugar beet share the same complement of genes of known function but differ in the content of expressed ORFs. Mol. Genet. Genomics, 272, 247–256.

Schägger H., von Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem., 166, 368–379.

Schmidt T., Jung C., Metzlaff M. 1991. Distribution and evolution of two satellite DNAs in the genus *Beta*. Theor. Appl. Genet., 82, 793–799.

Schondelmaier J., Jung C. 1997. Chromosomal assignment of the nine linkage groups of sugar beet (*Beta vulgaris* L.) using primary trisomics. Theor. Appl. Genet., 95, 590–596.

Skaracis G.N., De Biaggi M. 2005. Conventional breeding methods. [In:] Genetics and breeding of sugar beet. Eds. E. Biancardi, L.G. Campbell, G.N. Skaracis, M. De Biaggi. Science Publishers, Enfield, 169–220.

Steinborn R., Weihe A., Börner T. 1992. Mitochondrial genome diversity within a cultivar of *Daucus carota* (ssp. *sativus*) revealed by restriction fragment analysis of single plants. Plant Breed., 109, 75–77.

Stojałowski S., Łapiński M., Szklarczyk M. 2006. Identification of sterility-inducing cytoplasms in rye using the plasmotype-genotype interaction test and newly developed SCAR markers. Theor. Appl. Genet., 112, 627–633.

Szklarczyk M., Oczkowski M., Augustyniak H., Börner T., Linke B., Michalik B. 2000. Organisation and expression of mitochondrial *atp9* genes from CMS and fertile carrots. Theor. Appl. Genet., 100, 263–270.

Wesołowski W., Szklarczyk M., Szalonek M., Słowińska J. 2015. Analysis of the mitochondrial proteome in cytoplasmic male-sterile and male-fertile beets. J. Prot., 119, 61–74.

Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A., Tingey S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., 18, 6531–6535.

Yamamoto M.P., Kubo T., Mikami T. 2005. The 5'-leader sequence of sugar beet mitochondrial *atp6* encodes a novel polypeptide that is characteristic of Owen cytoplasmic male sterility. Mol. Gen. Genomics, 273, 342–349.

Yoo K.S., Ok S.H., Jeong B.-C., Jung K.W., Cui M.H., Hyoung S., Lee M.-R., Song H.K., Shin J.S. 2011. Single cystathionine β-synthase domain–containing proteins modulate development by regulating the thioredoxin system in *Arabidopsis*. Plant Cell., 23, 3577–3594.

Abstract

The research was aimed at the identification of mitochondrial DNA (mtDNA) and protein features which differentiated cytoplasmic male-sterile (CMS) and male-fertile beets. The analysis encompassed CMS and maintainer lines of sugar beet, table (red) beet and fodder beet. The polymorphic mtDNA sequences were searched with the use of both RAPD- and vectorette PCR. The latter method was adopted for the analysis of plant mtDNAs by designing a series of universal mitochondrial primers (UMPs). The search resulted in the identification of 26 cytoplasm-specific markers – 5 RAPDs and 21 vectorette markers. A selection of these markers was subjected to cloning and sequencing. This revealed a few mtDNA sequence arrangements which were missing in the whole genome sequence records available in the databases. Specificity of the identified RAPD and vectorette markers was also verified with the use of next generation sequencing (NGS) – by mapping mitochondrial NGS reads to the marker sequences. Moreover, the RAPD and vectorette marker sequences were used to design plasmotype-specific SCAR markers. Five of these markers proved to be useful for the differentiation of S- and N-cytoplasmic beets. Finally, a proteomic study was performed in order to search for proteins associated with CMS in beets. The comparison of 2-DE protein maps revealed 13 proteins, which in certain genetic background(s) showed plasmotype association. One of these proteins – CBSX3 – was earlier linked with male sterility in *Arabidopsis thaliana*.

Streszczenie

Celem badań była identyfikacja cech mitochondrialnego DNA (mtDNA) i proteomu, które różnicowały cytoplazmatycznie męskosterylne (CMS) i męskopłodne formy buraka. Analizami objęto linie CMS i dopełniające buraka cukrowego, ćwikłowego oraz pastewnego. Do poszukiwania polimorficznych sekwencji mtDNA wykorzystano techniki RAPD- i vectorette PCR. Drugą z tych metod zaadaptowano do analizy roślinnych mtDNA poprzez zaprojektowanie zestawu uniwersalnych starterów mitochondrialnych (UMP). W wyniku przeprowadzonych badań zidentyfikowano 26 markerów plazmotypowo-specyficznych – 5 typu RAPD i 21 typu vectorette. Część tych markerów poddano klonowaniu i sekwencjonowaniu. W ten sposób wykryto kilka układów sekwencji mtDNA, których nie zawierały dostępne w bazach danych rekordy sekwencyjne całych genomów. Specyficzność zidentyfikowanych markerów RAPD i vectorette zweryfikowano także przy użyciu sekwencjonowania następnej generacji (NGS) – poprzez mapowanie mitochondrialnych odczytów NGS do sekwencji markerowych. Ponadto sekwencje markerów RAPD i vectorette zostały wykorzystane do zaprojektowania plazmotypowo-specyficznych markerów SCAR. Wykazano, iż pięć spośród tych markerów nadaje się do różnicowania roślin S- i N-plazmatycznych. Na koniec przeprowadzono analizy proteomiczne, w których poszukiwano białek związanych z cechą CMS u buraka. Porównując białkowe mapy 2-DE, znaleziono 13 białek, które w pewnym kontekście genetycznym wykazywały związek z plazmotypem. Jedno z tych białek – CBSX3 – było wcześniej wiązane z męską sterylnością u *Arabidopsis thalian*a.

Appendix I – Primer list

Primer sequences are given in the 5' to 3' direction.

RAPD primers

A-01	caggcccttc
A-02	tgccgagctg
A-03	agtcagccac
A-05	aatcgggctg
A-06	ggtccctgac
A-08	gtgacgtagg
A-09	gggtaacgcc
A-10	gtgatcgcag
A-11	caatcgccgt
A-12	tcggcgatag
B-01	gtttcgctcc
B-02	tgatccctgg
B-03	catccccctg
B-04	ggactggagt
B-05	tgcgcccttc

Universal mitochondrial primers

nad3-f	atttctatctatttagtgat
nad3-r	ttactcccgatccgaagcac
nad4L-f	gatcttatcaaatatttcac
nad4L-r	ttaaccttgaatgctattaa
nad6-f1	aaatccggtacattccgttt
nad6-f2	ttcaaatagcggagattcac
nad6-r1	tagtagatcgtgagtgggtc

nad6-r2	agtactatagccccaatcat
nad9-f	actttacccaagaaatgggt
nad9-r	attcttgggtcatctcaatg
cob-f1	cgagcaatcttagttattgg
cob-f2	gcattgcatcaatatggatc
cob-r1	atgaagaagactggcgccta
cob-r2	cacaggttgacatccgatcc
cox1-f1	tcggtgccattgctggagtg
cox1-f2	tacttcaccgcagctaccat
cox1-r1	acagtaaacatatgatgagc
cox1-r2	gaggactttgtaccatccat
cox3-f	tgaatctcagaggcattctt
cox3-r	ttcaaagccaacgtgatgct
atp1-f1	aagtggatgagatcggtcga
atp1-f2	gtggcatatcgacaaatgtc
atp1-r1	gaagcggtggctgctacaag
atp1-r2	agccttgcacctctattgag
atp6-f	gtaaaggaacaaataggtgg
atp6-r	agcatcattcaagtaaatac
atp8-f	acttgggatcagcagaattc
atp8-r	ttccattcctcgtgagccac
atp9-f1	atgttagaaggtgcaaaatc
atp9-f2	cggagctgctacaattgctt
atp9-r1	tcagaatacgaataagatca
atp9-r2	atcattggggcaaacaatgc
ccb203-f	ttggaccgccggcgcgaaca
ccb203-r	agatgcacaagagtacttcg
ccb206-f	ctccacaccaatcacgagtt
ccb206-r	actaatcgagaccgaaattg
ccb256-f	gttgttcttaacagcgatgg
ccb256-r	gaagttagcaaagttagaca
ccb382-f	tgcattcacttacaacaaga
ccb382-r	gatccgacgcatccagcaga
rpl2-f	caagtgaaatccaatccacg
rpl2-r	tcgacagggttcataactac
rpl5-f	tcaattttcattacgaagat
rpl5-r	tcgaagatctcgaaatgatc
rps4-f	cattaagatttcaaacttgt
rps4-r	gaataatgagtaggtagttc
rps7-f	atgtcacgccgaggtactgc
rps7-r1	gcccctaataaccaacgaat

rps7-r2	ttgcctctgccattctatga
rps13-f	atatatttcaggagctagat
rps13-r	gcattagtatgagttcgttg
rrn5-f	gcactacggtgagacgtgaa
rrn5-r	tcagtacatatggcgcaaga
rrn18-f1	gctcagaaggaacgctagct
rrn18-f2	tgagtgttcgcccttggtct
rrn18-r1	ttacggcatggactaccagg
rrn18-r2	tgctgatccgcgattactag
rrn26-f1	gcgcactcactctaacggcg
rrn26-f2	cgacacaggtgaacaagtag
rrn26-r1	tgtacccatcggacggcagc
rrn26-r2	gatgtgatgagtcgacatcg
orf25-f	cgaagaagatctcaatctat
orf25-r	taagtaatcgttgttgttca
orfX-f	cgttcggatattgatcggtc
orfX-r	ttggcaccagatatccggag
matR-f1	agaggcgatcagaatggtac
matR-f2	caacaatagctgcacggagt
matR-r1	tctctgacctgataactagt
matR-r2	agcttcccaagctctatgct
nad1ex1-f	gtacatagctgttccagctg
nad1ex1-r	acctagcgaagatatggcaa
nad1ex2-f	tctaggagcattacgatctg
nad1ex2-r	ataagaataagaccaataga
nad1ex3-f	ctgtactaatatgtgtaggt
nad1ex3-r	tctacattatagcctgcaac
nad2ex1-f	agcggtttccccagagatct
nad2ex1-r	acactaagtaatccaagcca
nad2ex2-f	gcgcacctctcctaactatt
nad2ex2-r	aatgcacctaagatcaaata
nad2ex3-f	ttccatgatctatgggtcta
nad2ex3-r	ataggaatcctacagcgata
nad4ex1-f	gaacatttctgtgaatgcta
nad4ex1-r	aacatagggattggcacgct
nad4ex2-f	agatcaaggcagcatatcag
nad4ex2-r	ctaaacataccaatagtcac
nad4ex3-f	atgttaagtcatggactggt
nad4ex3-r	ccagaaaccacacgattata
nad4ex4-f	gcatgcatacatccgtaagt
nad4ex4-r	ccatgttgcactaagttact

nad7ex1-f	ggacctcaacatcctgctgc
nad7ex1-r	caatatgtggttccgcacgt
nad7ex3-f	gaacacgctcattcttcagc
nad7ex3-r	tcttaacattacaccactga
nad7ex4-f	gagcagcaccttacgatgtt
nad7ex4-r	gattccatggatagtttcat
nad7ex5-f	tctacctataccgcagttga
nad7ex5-r	acacaatatcttgagtacct
cox2ex1-f	gttcatggaactactatcga
cox2ex1-r	tcactgcactgaccatagta
cox2ex2-f	ggtatccaatcaattaatcc
cox2ex2-r	aacagcttctacgacgatag
ccb452ex1-f	aatggttcgtcagtagagat
ccb452ex1-r	gacagaagaacacccaacat
ccb452ex2-f	tcactcaacgattgcctttg
ccb452ex2-r	tcgaagcatatagatccgta

SCAR primers

O_A08_scar-f	tagacacaagccgtgacgta
O_A08_scar-r	atgagaaggtgggttgcgta
S_B01_scar-f	gataatcaggctcgaactga
S_B01_scar-r	tgactcagacctgttctaag
O_B05_scar-f	tgcaggatgaatatcgtgct
O_B05_scar-r	taggtgcaggtacatgtaca
O_atp8-f_scar	ttgactcctcgaatacaatg
O_cob-r1_scar	ctcttcttctagttctagca
O_rps7-r1_scar	actgagaacacctgatcagt
O_atp6-r_scar	ctctcgacagttagctagag
O_atp9-f2–38_scar	cgtctgtttcgtgctcatct
O_atp9-f2–39_scar	cgttataacgcaatatagag
S_atp6-r_scar	atggacgattacggtattcg
S_cob-r1_scar	tctctccgtgatcgtagttc
S_nad3-r_scar	taaggatagccatttagttg
S_rrn26-f1_scar	gtaattcaaagtacaggaac

Appendix II

Table A1. Database sequence records producing the most significant alignments with the sequences of the identified RAPD and vectorette markers. Homology searches were performed using the Blastn algorithm which returned the presented similarity parameters

Marker	Record title	Accession	Max score	Total score	Query cover [%]	E value	Identity [%]
N_A-08	*Beta vulgaris* subsp. *maritima* genotype male-sterile E mitochondrion, partial genome	FQ014226	435	435	97	1e-118	99
N_B-05	*Beta vulgaris* subsp. *vulgaris* uncharacterized LOC104898410 (LOC104898410), mRNA	XM_010685495	518	518	60	2e-142	68
S_B-01_c29	*Beta vulgaris* subsp. *vulgaris* mitochondrial DNA, complete genome	BA000024	3427	3528	99	0.0	99
S_B-01_c3	*Beta vulgaris* subsp. *vulgaris* genes for 16S rRNA, tRNA-Val, ribosomal protein S12, ribosomal protein S7, NADH dehydrogenase ND2 subunit, partial and complete cds	AB032426	4158	4158	99	0.0	99
N_atp6-r	*Beta macrocarpa* mitochondrion, complete genome	FQ378026	3094	5194	100	0.0	100

Table A1. cont.

Marker	Record title	Accession	Max score	Total score	Query cover [%]	E value	Identity [%]
N_atp8-f	*Beta vulgaris* subsp. *maritima* genotype male-sterile E mitochondrion, partial genome	FQ014226	3297	3733	100	0.0	99
N_atp9-f2_c38	*Beta vulgaris* subsp. *maritima* genotype male-sterile E mitochondrion, partial genome	FQ014226	2693	3151	90	0.0	99
N_atp9-f2_c39	*Beta vulgaris* subsp. *maritima* genotype male-sterile E mitochondrion, partial genome	FQ014226	2722	3174	100	0.0	99
N_cob-r1	*Beta vulgaris* subsp. *maritima* genotype male-sterile E mitochondrion, partial genome	FQ014226	1689	1741	98	0.0	99
N_rps7-r1	*Beta vulgaris* subsp. *maritima* genotype male-sterile E mitochondrion, partial genome	FQ014226	1896	2657	100	0.0	99
S_atp6-r	*Beta vulgaris* subsp. *vulgaris* mitochondrial DNA, complete genome	BA000024	2863	2863	100	0.0	99
S_cob-r1	*Beta vulgaris* subsp. *vulgaris* mitochondrial DNA, complete genome	BA000024	663	1326	96	0.0	100
S_nad3-r	*Beta vulgaris* subsp. *vulgaris* mitochondrial DNA, complete genome	BA000024	1462	2128	100	0.0	99
S_rrn26-f1	*Beta vulgaris* subsp. *vulgaris* mitochondrial DNA, complete genome	BA000024	1997	3124	99	0.0	100