

UDC 633.174:631.527.56:577.21

doi: 10.15389/agrobiology.2017.5.952rus

doi: 10.15389/agrobiology.2017.5.952eng

## POLYMORPHISM OF GRAIN SORGHUM FROM VIR WORLD COLLECTION FOR THE CHARACTERS ASSOCIATED WITH THE CMS-*Rf* GENETIC SYSTEM

I.N. ANISIMOVA<sup>1</sup>, D.N. RYABOVA<sup>1</sup>, E.V. MALINOVSKAYA<sup>2</sup>, N.V. ALPATIEVA<sup>1</sup>,  
Yu.I. KARABITSINA<sup>1</sup>, E.E. RADCHENKO<sup>1</sup>

<sup>1</sup>Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources, Federal Agency of Scientific Organizations, 42-44, ul. Bol'shaya Morskaya, St. Petersburg, 190000 Russia, e-mail irina\_anisimova@inbox.ru (corresponding author);

<sup>2</sup>Kuban Experiment Breeding Station, Branch of Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources, Federal Agency of Scientific Organizations, 2, ul. Tsentral'naya, pos. Botanika, Gul'kevichskii Region, Krasnodarskii krai, 352183 Russia

ORCID:

Anisimova I.N. [orcid.org/0000-0003-0474-8860](http://orcid.org/0000-0003-0474-8860)

Alpatieva N.V. [orcid.org/0000-0002-5531-2728](http://orcid.org/0000-0002-5531-2728)

Ryabova D.N. [orcid.org/0000-0002-1729-0900](http://orcid.org/0000-0002-1729-0900)

Karabitsina Yu.I. [orcid.org/0000-0002-8384-5134](http://orcid.org/0000-0002-8384-5134)

Malinovskaya E.V. [orcid.org/0000-0002-0547-0760](http://orcid.org/0000-0002-0547-0760)

Radchenko E.E. [orcid.org/0000-0002-3019-0306](http://orcid.org/0000-0002-3019-0306)

The authors declare no conflict of interests

Acknowledgements:

The sequencing was carried out using equipment of the ARRIAM Center of Genomic Technologies, Proteomics and Cell Biology (St. Petersburg).

The authors thank A.G. Pinaev (ARRIAM, St. Petersburg) for assistance in DNA sequencing

Received July 3, 2017

### Abstract

Seven different types of cytoplasmic male sterility (CMS) are known for the grain sorghum (*Sorghum bicolor* L. Moench), however, only A1 (milo) is used in heterotic hybrid breeding. The genetic control of pollen fertility restoration of CMS A1 is complex and determined by two or three *Rf* (*Restoration of Fertility*) genes, and also by a number of modifiers. It is very little known about molecular mechanisms of CMS A1 and fertility restoration. Only one gene, *Rf1*, is identified at the molecular level (R.R. Klein et al., 2005). In the present paper we have demonstrated for the first time the nucleotide polymorphism in the coding regions of the recessive and dominant alleles of *Rf2* gene and also of the candidate *RFL-PPR* gene homologous to the rice *Rf1* gene. Here, we studied polymorphism of the CMS-*Rf* genetic system related traits in sorghum accessions from the VIR collection, including the fertility restorers k-928 and k-929; a half-restorer k-1362, the sterile lines A-10598 and A-83 (CMS A1) and their fertile analogs, the F<sub>8</sub>-F<sub>12</sub> BC<sub>1</sub>-BC<sub>2</sub> sister lines resistant to *Schizaphis graminum* Rond. which have been isolated among the hybrids derived from crosses between the sterile (CMS A1) line N-81 and lines k-929 and k-928, and also hybrids between the sister lines. For investigating variability of candidate genes associated with the CMS-*Rf* genetic system the reference sequences were selected from the bioinformatic database (<http://www.ncbi.nlm.nih.gov>), the eighth specific primers were designed, and the fragments amplified on the DNA of genotypes differing by the ability to suppression of the CMS phenotype were sequenced. In the CMS lines and fertility restorers a significant polymorphism (18 polymorphic sites) was revealed in the 825 bp fragment of the *Rf2* coding region (reference fragment XM\_002459403.1, chromosome SDI02) and also in *RFL-PPR* candidate gene located in the chromosome 3 (reference fragment XM\_002458104.1). The sequenced regions of the structural nuclear gene *ALDH2b* encoding mitochondrial aldehyde dehydrogenase, the maize *Rf2* gene homolog, and also of the mitochondrial F0F1 ATPase alpha subunit were identical in the CMS and fertility restorer lines. Variability of pollen fertility indices was studied using acetocarmine stained cytological preparations. The lines resistant to *S. graminum*, and their hybrids differed in the percentage of stained (fertile) pollen grains, the presence of anomalous large pollen grains (54-70 μm in diameter), giant pollen grains (up to 84 μm in diameter) and deformed pollen grains. In the fertile F<sub>8</sub>-F<sub>12</sub> BC<sub>1</sub>-BC<sub>2</sub> lines which derived from the hybrids produced in crossings with fertility restorers, the frequency of stained pollen grains was relatively high and reached 72.2-83.8 % for k-929, and 57.4 and 63.4 % in two lines, respectively, for k-928; large pollen grains occurred at different frequency in five lines, and the giant ones were observed in two lines. The variability in pollen fertility could be due to the differences in the alleles derived from the recurrent parent.

Keywords: *Sorghum bicolor* L. Moench, grain sorghum, CMS, fertility restoration, *Rf*, pollen fertility, candidate genes, nucleotide polymorphism

The phenomenon of cytoplasmic male sterility (CMS) was discovered in the United States in 1931 by M. Rhoades and in the USSR by M.I. Khadzhinov (All-Union Institute of Plant Industry — VIR). To date, it has been described for more than 150 species of plants [1, 2]. CMS is widely used in the production of hybrid seeds of many crops (maize, rice, rape, cotton, sunflower, cabbage, etc.). The creation of heterotic hybrids based on CMS is considered as a priority of modern breeding programs for sorghum (*Sorghum bicolor* L. Moench), an important cereal crop widely spread in arid and semi-arid regions of the planet. The first mention of heterotic hybrids of sorghum dates back to 1927 [3, 4]; however, commercial hybrids were made only after the discovery of stably inherited cytoplasmic male sterility A1 (milo) in Kaffir sorghum and the sources of pollen fertility restoration [5]. Later, alternative types of CMS — A2-A6, 9E [6] were identified, but because of the difficulties in obtaining reliable sources of fertility restoration genes, as well as the epigenetic variability of the trait depending on environmental conditions [7-9], CMS A1 only is used for breeding programs [10]. At least two main genes, *Rf1* [11] and *Rf2* [10], whose dominant alleles are responsible for the restoration of the A1-milo CMS fertility, have been identified by hybridological analysis. Their manifestation depends on the environmental conditions and the action of the modifier genes. It is also suggested that in the case of CMS A1, pollen fertility restoration is controlled by the dominant alleles of the main and two duplicated complementary genes, whereas it is determined by the dominant alleles of the three genes interacting complementarily in CMS A2 and A3 [12]. The *Rf5* gene and a number of modifiers that restore the pollen fertility of CMS A1 and A2 have been found [13]. Fertility restoration with CMS A3 is controlled by dominant alleles of *Rf3* and *Rf4* genes at gametophyte level [14], and is explained by paramutations of *Rf* genes at sporophyte level [15]. The molecular mechanisms of CMS manifestation and fertility restoration in sorghum have been little studied. It has been shown that, like in other plants, sorghum CMS is caused by aberrant mitochondrial genes [16].

Most of the *Rf* genes characterized to date (in petunia, maize, rice, radish) encode proteins that contain repeating motifs of 35 amino acid residues (PPR, pentatricopeptide repeats) and regulate the coordinated work of the nucleus and mitochondria. *PPR* genes with the fertility restoration function are allocated to a separate subfamily *RFL-PPR* (*Restoration of Fertility Like-PPR*). The structural and functional diversity of *PPR-RFL* genes is supported by the variability of PPR motifs, as well as by the complex cluster organization of *Rf* locus in the genome [17-19].

Only one gene for pollen fertility restoration, *Rf1*, has been characterized for sorghum at the molecular level [11]. It was found that the locus *Rf1* is in the linkage group 08 and includes four open reading frames (ORFs) that code for Ca<sup>2+</sup>-ATPase of the plasma membrane, cyclin D-1, and also an unknown mitochondrial protein containing 13 PPR motifs and belonging to the E-type of *PPR* genes subfamily. There are 19 polymorphic sites in the coding sequence and in the 5'- or 3'-end flanking regions of the dominant and recessive alleles of the candidate gene *PPR13*. Sorghum *PPR13* gene differs significantly from other representatives of the *PPR-RFL* genes subfamily in the structure of PPR motifs [19-21]. Another candidate gene (*Rf2*) is located on 236,219 bp region of SD102 chromosome. The region includes 31 ORFs, including one *PPR* gene characterized by high similarity to the rice gene *Rf1* [10]. The polymorphism of the nucleotide sequences of the dominant and recessive alleles of the *Rf2* locus has not yet been studied thus limiting the development of specific molecular markers for their identification. Other sequences of sorghum genome potentially associated with pollen fertility restoration were not identified, too.

The VIR collection of sorghum totals about 9 thousand accessions. In the collection, there are CMS A1 (milo) sterile lines, fertility restorers and sterility maintainers. In sorghum, like many other plants, cytological analysis of pollen in F<sub>1</sub> hybrids is used, in addition to seed setting rate under self-pollination, to assess the male fertility restoration [8, 13, 22]. However, the variability of this trait and the polymorphism of genomic sequences potentially associated with the CMS-*Rf* genetic system have not yet been studied yet in the samples of VIR collection.

In this study, we first reported on significant nucleotide polymorphism in coding regions of *Rf2* gene recessive and dominant alleles (the reference fragment XM\_002459403.1, chromosome SDI02), and also of the candidate gene *RFL-PPR* (the reference fragment XM\_002458104.1) homologous to the rice gene *Rf1*.

In order to elucidate the character of variability associated with the CMS-*Rf* genetic system in sorghum lines, we studied nucleotide polymorphism of the candidate genes associated with the genetic system CMS-*Rf*, and compared the formation of pollen in fertile forms and plants with a sterile cytoplasm of type A1 (milo).

*Techniques.* Studied sorghum samples were the VIR collection accessions which differ in their ability to restore pollen fertility, i.e. fertility half-restorer k-1362; restorer lines k-928 and k-929; sterile lines A-10598 and A-83 based on CMS A1 (milo) and their respective fertile counterparts B-10598 and B-83; sister lines F<sub>8</sub>-F<sub>12</sub> BC<sub>1</sub>-BC<sub>2</sub> resistant to *Schizaphis graminum* Rond. which were derived from the hybrids produced by crossing sterile line (CMS A1) Nizkorosloe 81 (N-81) with the restorers k-928 and k-929; F<sub>1</sub> hybrids from crosses of the sterile and the fertile lines [23]. The lines k-928 and k-929 selected from the samples of grain sorghum Dzhugara Belaya from Western China possess different alleles of for resistance to *S. graminum* [24]. The sample k-1362 (Dzhugara Belaya, Syria) is a donor of resistance to *S. graminum* and the fertility half-restorer. The sterile lines of grain Kaffir sorghum A-10598 and A-83 and their fertile analogues B-10598 and B-83 entered the VIR collection from India in the 1980s. The lines and hybrids were grown in 2014-2016 (field tests, VIR Kuban Experimental Station, Krasnodar Krai).

For estimation of pollen fertility, mature anthers were collected early in the morning in the period of mass flowering and fixed in 70 % ethanol. The proportion of pollen fertile grains was counted according to Navashin method [cited from 25] in preparations stained with aceto-carmin glycerol (Zeiss Axioplan 2 imaging, Carl Zeiss, Germany). The percentage of fully stained (fertile), weakly stained and unstained pollen grains (PGs) was calculated based on at least 30 fields of view at a magnification of ×20 and considering pollen diameter, uniformity in diameter and the presence of abnormal PGs.

DNA was extracted according to a CTAB protocol [26]. Eleven sequences were identified as a result of bioinformatic search (GenBank database, NCBI, USA) (<http://www.ncbi.nlm.nih.gov>). These sequences have a homology to the nuclear *Rf* genes and mitochondrial genes associated with CMS in sorghum [27] and other plant species. Eight pairs of specific primers flanking the full-sized and internal fragments were designed based on four selected reference sequences (Table 1). Amplicons produced with the developed primers using template DNA of genotypes, contrasting in pollen fertility, were purified in 1 % agarose gel and sequenced on a genetic analyzer ABI 3500xl (Applied Biosystems, USA). The alignment and analysis of the sequences were implemented using Mega 5.1 software (<http://www.megasoftware.net/>) [28].

The means (*M*) and standard error of means ( $\pm m$ ) were calculated for

the fertile pollen grains number.

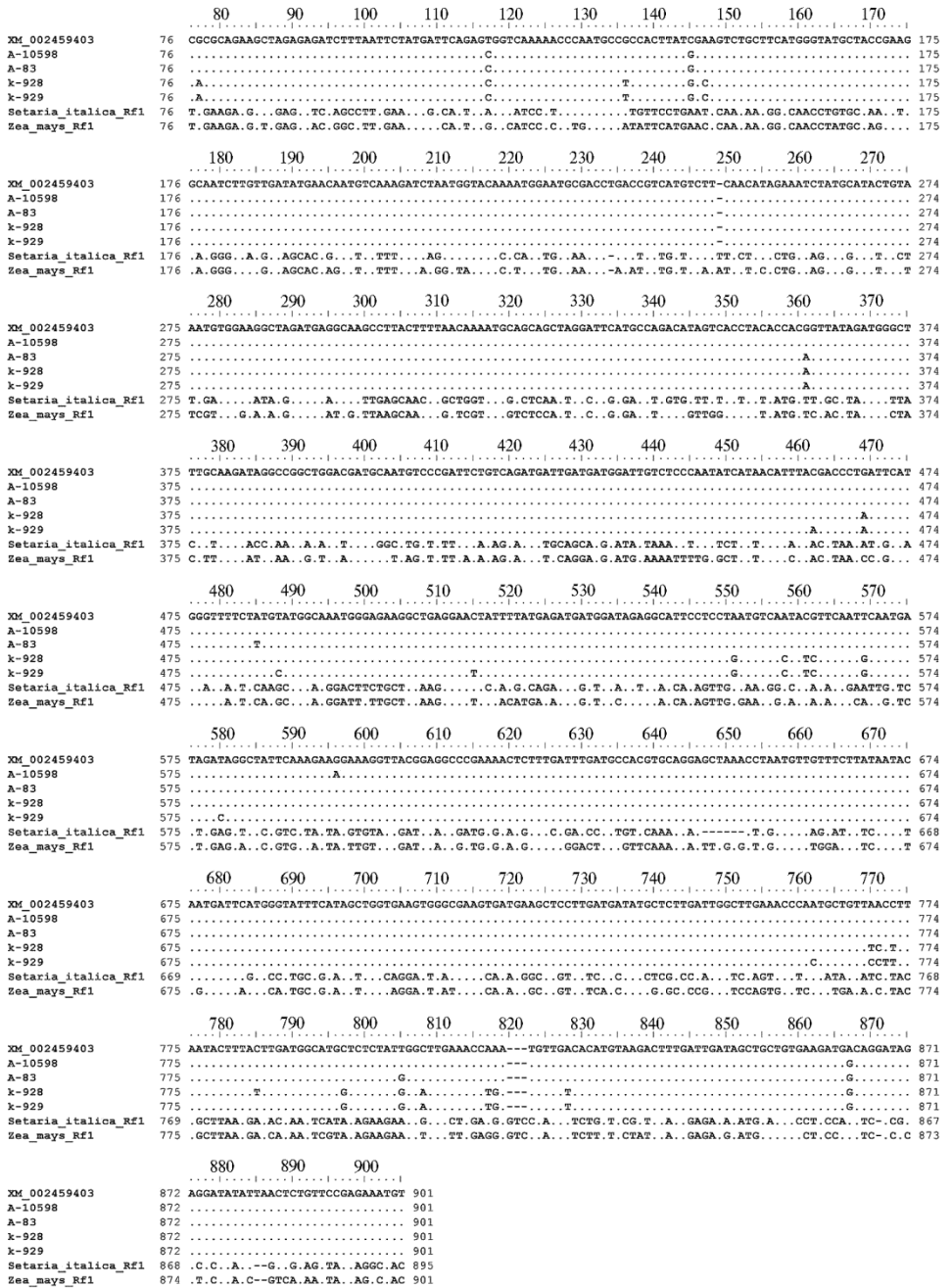
**Results.** To determine the variability of *Rf2*, one of the main genes for CMS A1 fertility restoration located on chromosome 8, we studied the polymorphism of fragments amplified using designed primers specific for the *S. bicolor* genomic sequence (XM\_002459403.1) which contains PPR motifs and is the closest [10] to the sequence of rice *Oryza sativa* L. *Rf1* gene restoring the fertility of CMS type BTII (Boro II) (Table 1). In this, we revealed a similarity of the reference fragment with several fragments of sorghum genome (presumably, the gene *Rf1* sequences) and also with the predicted fertility restorer genes of *Setaria italica* (L.) P. Beauv. and *Zea mays* L. The reference fragment length was 951 bp and the sequenced fragment size was 825 (positions 76-901). The length of in silico translated sequence was 275 bp.

### 1. Primers designed to amplify fertility restoration gene homologs in *Sorghum bicolor* L. Moench

Reference fragment (length)	Gene, protein	Primers, nucleotide sequence 5'→3'	Sequenced fragment, bp	Positions in the reference sequence
XM_002458104.1 (2801 bp)	Not identified, PPR protein	02458104fw1: CACCCAATTCTCCAGACCAT 02458104rev1: ACATCTGCCGTACATAGCC 02458104fw2: GGCTATGTACCGGCAGATGT 02458104rev2: GATGGGATCAAATGGAATGG 104_inner_fw: TTGCTTGCATGGAGAAATTG 104_inner_rev: CTGCGAGATCACAGCAGTTG	818	301-1119
XM_002459403.1 (951 bp)	<i>Rf2</i> , PPR protein	2459403fw: CAGGGGCCAAATGTTGTTC 2459403rev: CACAGTTTTATATTTCCGTGAT-AGTG	825	76-901
AJ278689.1 (1324 bp)	<i>atpA</i> , $\alpha$ subunit of ATPase synthase F0F1	AJ278689fw: AACTTTTACCGAATTTCAAGTGG AJ278689rev: TGACAGCAGCATAAATAACAACAA AJ_inner_fw: TCCTATAGGCCGTGGTCAAC AJ_inner_rev: CGTCTCCAGCTTGTGTTTCA	1183	60-1243
AB084898.1 (2159 bp)	<i>ALDH2b</i> , mitochondrial aldehyde dehydrogenase	AB084898fw: TTCTGGTTTTGGCCCTACTG AB084898rev: CTCTTCTAACAAATGTTTTTCAT-AAT AB_inner_fw: AACCATACGAATAAAGCCTTGC AB_inner_rev: CTCGCATTTGCCCTCTTAAT	738	858-1596

For the restorer lines k-928 and k-929, presumably carrying dominant alleles of *Rf2*, the studied sequences differed in six nucleotide substitutions, whereas sterile lines A-10598 and A-83, the carriers of recessive alleles, differed in three nucleotide substitutions. The reference sequence which originates from of BT  $\times$  623 line, a maintainer of sterility with a putative genotype *rf2rf2*, was more similar to the sequences of the sterile lines A-10598 and A-83 and differed from them in two and four polymorphic nucleotide positions, respectively. At the same time, the sequenced fragments of k-928 and k-929 genomes differed from the reference fragment and fragments found in sterile lines in 18 nucleotides substitutions and 7 amino acid substitutions (Fig. 1). It is known that the polymorphism of the coding sequences of *Rf* genes of plants, particularly *Rf1* gene in sorghum, is associated with their functional status [11, 18]. We can assume that the polymorphic sequences identified in CMS lines and fertility restorers, are different allelic variants of gene *Rf2* which products are involved in editing mitochondrial RNA. In the sequences, we found six and five PPR repeats for the carriers dominant and recessive *Rf2* alleles, respectively. In general, the *Rf2* locus polymorphism in the dominant and recessive allele carriers was relatively high (2.2 % of polymorphic sites).

We studied polymorphism of 2,801 bp XM\_002458104.1 fragment located on chromosome 3, using three pairs of primers (Table 1, Fig. 2). XM\_002458104.1 contains 17 PPR repeats and is the closest to nuclear gene *Rfo* of radish *Raphanus*



**Fig. 1. Alignment of the nucleotide sequences of the candidate gene *Rf2* of *Sorghum bicolor* L. Moench lines from the world VIR collection: k-928 and k-929 — carriers of dominant allele, A-10598 and A-83 — carriers of recessive allele; XM\_002459403 — reference fragment. For comparison, homologous fragments of the genes *Rf Setaria italica* and *Zea mays* are given.**

*sativus* (L.) Domin., restoring the fertility of CMS Ogura [29], and the gene *Rf1* of rice *O. sativa* Japonica Group. The reference fragment and 818 bp DNA fragment of line A-83 genome were identical, but differed from the DNA fragment of the fertility restorer k-929 (17 polymorphic sites and 8 amino acid substitutions revealed) (Fig. 2). The highest similarity (90 % identical nucleotide positions) was found between the identified fragment and XM\_008677263.2 sequence from Gen-

Bank database for *Zea mays* PPR protein.

```

..... 310      320      330      340      350      360      370
XM_002458104.1  301  GGGCTCGCGATGACCTCGCTTCCCTCCCTCCGCGCACTCCTCGCATCCTCGCCGACGGCCGCGACGCCT 370
A-83            301  ..... 370
k-929          301  .....ATCG..... 370
Zea mays XM_008677263.2 301  .....C.....AC..T..A..... 370

..... 380      390      400      410      420      430      440
XM_002458104.1  371  TCCACCTGCTACGTTCCGCTGCGCTGGATACGCGCTCCCGCGAAGAGCTCGTGCAGCGCTCCTCTC 440
A-83            371  ..... 440
k-929          371  ..... 440
Zea mays XM_008677263.2 371  .....T.....C..... 440

..... 450      460      470      480      490      500      510
XM_002458104.1  441  CGTCGCAGACGCTGGCTCGCCGGCCGCGCCACGCTCCTCAGTCAGTCCCTCACCCTGCCTCTCCAGCGCC 510
A-83            441  ..... 510
k-929          441  ..... 510
Zea mays XM_008677263.2 441  ..AC..TG.....TT..T.....C.....C.....T 510

..... 520      530      540      550      560      570      580
XM_002458104.1  511  GCCCGGACTCGCGGGCCGCGCTGGCCGATACTCTCGCATGGTCCAGGGGGGTTGTCGCCGACGCCA 580
A-83            511  ..... 580
k-929          511  ..... 580
Zea mays XM_008677263.2 511  .....A.....C..... 580

..... 590      600      610      620      630      640      650
XM_002458104.1  581  AATCTCGCACCGACTGCTCGTCGCCACGGCACGGGGCGCATCGGCTGCGGATCGCTCAGCTGTTCGA 650
A-83            581  ..... 650
k-929          581  ..... 650
Zea mays XM_008677263.2 581  ..G.....A..A.....G.....G.....A..... 650

..... 660      670      680      690      700      710      720
XM_002458104.1  651  CGAGATCCGGTGCAAGGGGTACTACGCGGATGCTAAGATGTACGACGTCGTGATCGGGGCTCGCTCGTG 720
A-83            651  ..... 720
k-929          651  ..... 720
Zea mays XM_008677263.2 651  .....G.....A.G.....G.....A.....AG. 720

..... 730      740      750      760      770      780      790
XM_002458104.1  721  GGACGGATGCACGGTACGCGCTCAGGCTGTTCGACGAAATGGCCGGTGCCGGAGTCAAGCCTGACGAGC 790
A-83            721  ..... 790
k-929          721  ..... 790
Zea mays XM_008677263.2 721  .....G.....T.....T.....T..T.....T.....G.....C..... 790

..... 800      810      820      830      840      850      860
XM_002458104.1  791  GCGTTTATGCCATCACAATCACAGGTTTGTGCAAGTACCGGATCAGACCGGGCAGTCCAGTCTGGG 860
A-83            791  ..... 860
k-929          791  ..... 860
Zea mays XM_008677263.2 791  .....T.....T.....C.....T..... 860

..... 870      880      890      900      910      920      930
XM_002458104.1  861  GAAGATGAGGGAGGCAGGGTTGAAGCCACGGGATTTACCTACAATTCTGTGGTGGATGTGCTGTGAAG 930
A-83            861  ..... 930
k-929          861  ..... 930
Zea mays XM_008677263.2 861  .....CG...GT...G..A..G...G..... 930

..... 940      950      960      970      980      990      1000
XM_002458104.1  931  GTGGGAGGATGGATGAGGCATTGCGGCTGAAGGATCAGATGCTCTGGCCACGGGGAAGAGATGGATG 1000
A-83            931  ..... 1000
k-929          931  ..... 1000
Zea mays XM_008677263.2 931  .....A.....T.....A..... 1000

..... 1010     1020     1030     1040     1050     1060     1070
XM_002458104.1  1001 TGTTTCTCGCGACGACGTTGATGACAGGATATTGCTTGCATGGAGAAATGGGAAAGCATTAGATTGTGTT 1070
A-83            1001  ..... 1070
k-929          1001  ..... 1070
Zea mays XM_008677263.2 1001 ..G.....T.....C..G.....G.....G..... 1070

..... 1080     1090     1100     1110
XM_002458104.1  1071 TGATGAGGCTGTCAGGATGGTGTGACACCGACCAATGTGACATAT-AC 1118
A-83            1071  ..... 1118
k-929          1071  ..A.A.G.....AA...T.....AC.TCT.....A..T 1118
Zea mays XM_008677263.2 1071 .....T.....T.....-GG 1118

```

Fig. 2. Alignment of the nucleotide sequences of the locus encoding mitochondrial PPR protein in *Sorghum bicolor* L. Moench lines from the world VIR collection: k-929 — fertility restorer, A-83 — sterile line; XM\_002458104.1 — reference fragment. For comparison, the homologous sequence of *Zea mays* is given.

*Rf2*, one of maize genes restoring fertility of T-type CMS line, encodes the aldehyde dehydrogenase (ALDH), the enzyme catalyzing oxidation of alde-

hydes. It is known that ALDH is involved in the detoxification of acetaldehyde, produced during the development of pollen, also ALDH may be involved in cell energy metabolism, especially during anther development, and possibly interacts with the mitochondrial protein URF13 related with the T-type CMS in maize [30, 31]. It is believed that the action of ALDH as a fertility restorer controlled by gene *Rf2* is due to the tunnel cavities on the surface of the enzyme molecule, through which the protein can bind long-chain ligands of different lengths and/or potentially toxic products of the mitochondrial CMS gene expression which are harmful for pollen development [32].

To test the hypothesis about the possible involvement of sorghum aldehyde dehydrogenase gene in control of pollen fertility restoration, we designed primers specific for coding sequence of the enzyme based on *ALDH2b* mRNA of *S. bicolor* (AB084898.1) from GenBank database. In sorghum *ALDH2b*, a gene of subfamily 2 of the extensive ALDH superfamily of higher plants, nucleotides located at the positions 858-1596 was homologous to the ALDH sequences of other cereal species including maize (94 % identical amino acid residues). At the same time, we found no differences between the fertility restorers and CMS lines in nucleotide sequences of this fragment. It should be noted that we succeeded to determine the sequence of an extended 2055 bp region of *ALDH2b* gene of line A-83, but sequenced and reference fragments differed considerably in the N-terminal regions. It may be due to a considerable variability of residues in the specified part of the molecule [32] or the presence in the amplified product of an intron or introns information of which is still lacking for sorghum and maize ALDH genes.

It is known that promoter regions of ATP-synthase genes or parts of these genes are found in many mitochondrial loci associated with CMS. For example, gene *urf13-T* associated with T-type CMS in maize contains 59 nucleotides from the regulatory region of the *atp6* gene. Gene *orf221* linked to and co-transcribed with *urf13-T* was identified as a fragment of gene encoding a subunit of F0 component of ATP synthase F0F1 [16]. We used AJ278689.1 sequence (GenBank database) of gene *atpA* encoding the  $\alpha$  subunit of ATP-synthase F0F1 of line CS3541, the fertility restorer for CMS A1, as a reference sequence to analyze fragments of *S. bicolor* mitochondrial genome potentially associated with CMS A1. Aligned fragments had a length of 1183 bp and were identical in epy forms with sterile (A-83) and fertile (k-929) cytoplasm types. The translation products of these sequences in the CMS line and the restorer line, as well as the reference sequence AJ278689.1 and sequences coding  $\alpha$  subunit of ATP-synthase F0F1 in maize, rice, *Triticum aestivum* L., *T. durum* L., and *Secale cereale* L. were highly homologous.

Thus, in this work, the polymorphism of nucleotide sequence fragments of a candidate gene *Rf2* which controls fertility restoration for CMS A1 in sorghum was revealed for the first time. The greatest number of polymorphic sites was found when comparing these sequences in sterile lines and restorers. In particular, the nucleotide polymorphism was found in the fertility restorers k-928 and k-929 which served as parental forms for *S. graminum* resistant lines F<sub>8</sub>-F<sub>12</sub> BC<sub>1</sub>-BC<sub>2</sub>. These lines have sterile cytoplasm and, apparently, differ in allelic composition of *Rf* genes derived from the male parent. Despite the high homozygosity of the lines F<sub>8</sub>-F<sub>12</sub> BC<sub>1</sub>-BC<sub>2</sub>, their pollen fertility varied significantly.

The studied genotypes differed significantly both in the amount of fertile (stained) pollen grains and in their diameter (Table 2). It has been reported [33] that normally the diameter of fertile PGs in diploid sorghum does not depend on the year condition and varies within 37.5-54.2  $\mu\text{m}$ . PGs of more than 54.2  $\mu\text{m}$  in diameter (large pollen grains) occur in diploid sorghum quite rarely and, appar-

ently, contain an unreduced number of chromosomes. B-10598 line, the fertile analog of CMS A-10598 line, as well as k-1362 and k-928 (75.5 %, 100.0 % and 83.3 % of fertile PGs, respectively), was characterized by high pollen quality. At the same time, k-1362 and k-928 had a large number of large PGs with a diameter of 55 to 70  $\mu\text{m}$ . Seven fertile lines F<sub>8</sub>-F<sub>12</sub> BC<sub>1</sub>-BC<sub>2</sub> resistant to *S. graminum* which were selected from the hybrids N-81  $\times$  k-928 and N-81  $\times$  k-929, showed a relatively high percentage of fertile PGs. This index was slightly higher for five lines in which k-929 served as the donor of the *Rf* genes (72.2-83.8 %, respectively) than that for two lines from crosses with k-928 (57.4 % and 63.4 %, respectively). Different frequency of large PGs of 55-70  $\mu\text{m}$  in diameter showed five lines, and two lines had giant PGs of up to 84  $\mu\text{m}$  in diameter.

## 2. Pollen grains of sorghum (*Sorghum bicolor* L. Moench) lines with different types of cytoplasm from the world VIR collection

Sample	Type of cytoplasm	Genotype	Characterization of pollen grains					
			F PGs, % ( $M \pm m$ )	D PG, $\mu\text{m}$		G PGs, %	L PGs, %	D uniformity/abnormal PGs
				min	max			
k-1362 Dzhugara Belaya	F	Fertility half-restorer	100	51.8	68.9	0	70.0	+/-
k-928 Dzhugara Belaya	F	Fertility restorer	83.3 $\pm$ 3.73	40.1	60.5	0	60.0	-/-
B-10598	F	Sterility maintainer	75.5 $\pm$ 9.71	42.2	48.3	0	0	+/-
2146/15	S	Fertile line F <sub>9</sub> BC <sub>2</sub> (N-81 $\times$ k-929)	75.3 $\pm$ 4.65	19.9	50.3	0	0	-/-
2148/15	S	Fertile line F <sub>10</sub> BC <sub>2</sub> (N-81 $\times$ k-929)	76.6 $\pm$ 5.56	26.5	58.5	1.8	0	+/+
2149/15	S	Fertile line F <sub>10</sub> BC <sub>2</sub> (N-81 $\times$ k-929)	75.6 $\pm$ 3.30	27.6	52.2	0	0	+/-
2150/15	S	Fertile line F <sub>10</sub> BC <sub>2</sub> (N-81 $\times$ k-929)	83.8 $\pm$ 3.66	28.7	84.3	2.7	8.2	-/-
2151/15	S	Fertile line, F <sub>12</sub> BC <sub>2</sub> (N-81 $\times$ k-929)	72.2 $\pm$ 6.34	16.4	57.2	0	2.1	-/-
2152/15	S	Fertile line F <sub>12</sub> BC <sub>1</sub> (N-81 $\times$ k-928)	63.4 $\pm$ 3.90	20.8	56.1	0	2.8	-/+
2153/15	S	Fertile line F <sub>12</sub> BC <sub>1</sub> (N-81 $\times$ k-928)	57.4 $\pm$ 3.20	28.5	56.1	0	3.1	-/-
73/16	S	Hybrid F <sub>1</sub> between sterile line and fertility restorer	27.0 $\pm$ 5.88	13.4	52.5	0	0	-/-
74/16	S	Hybrid F <sub>1</sub> between sterile line and fertility restorer	58.3 $\pm$ 8.96	28.5	60.2	0	3.0	-/-

Note. F PGs — fertile (stained) pollen grains, D PG — diameter of pollen grains, G PGs — giant pollen grains (diameter > 70-80 microns), L PGs — large pollen grains (diameter > 55 microns); “+” or “-” mean presence and absence of traits.

Only the line 2149/15 of seven analyzed sister lines had PGs uniform in the diameter. Since fertile lines have sterile cytoplasm inherited from the maternal form N-81, it can be assumed that their genotypes contain alleles of nuclear *Rf* genes derived from the male parents which to varying degrees influence pollen fertility restoration under sterile cytoplasm. Deformed PGs were encountered in two lines. The hybrids from the crosses of sterile lines resistant to *S. graminum* with presumably fertility-restoring sister lines also significantly differed in the proportion of fertile pollen, from 9.0 % to 58.3 % for different F<sub>1</sub> plants. One of the hybrids from crossing the sterile line with the sister maintainer line was completely sterile and did not form pollen while the other had pollen in the anther, which, however, was not stained, that is, was sterile. Thus, the variability in pollen fertility rates observed in the *S. graminum* resistant sister lines and in the interlinear hybrids derived from crossing with these lines indicates differences in the alleles of genes participating in fertility restoration (possibly minor genes or modifier genes), which is consistent with the hypothesis of complex genetic con-



trol of the trait [13].

Thus, the nucleotide sequence of *Rf2*, which is a candidate gene for fertility restoration of CMS A1 in sorghum, is polymorphic in the sterile lines A-10598 and A-83 and in the pollen fertility restorers k-928 and k-929 (Dzhugara Belaya). A 825 bp fragment of coding sequence of *Rf2* gene in the carriers of recessive and dominant alleles differs in 18 polymorphic sites, and the translated sequence differed in seven amino acid substitutions. The genomic fragment of one of the *PPR* genes homologous to the rice gene *Rf1* (XM\_002458104.1) is highly polymorphic in the lines differing in ability to suppress CMS phenotype, and contains 17 polymorphic sites. Identified polymorphism can be used in the development of allele-specific molecular markers for *Rf2* locus. The sequenced 1183 bp fragment of mitochondrial gene *atpA*, which encodes ATP synthase F0F1 subunit, and the 738 bp fragment of *ALDH2b* aldehyde dehydrogenase nuclear gene are identical in the studied lines. Lines F<sub>8</sub>-F<sub>12</sub> BC<sub>1</sub>-BC<sub>2</sub> resistant to *Schizaphis graminum*, which were selected among hybrids derived from crosses of CMS line Nizkorosloe 81 with fertility restorers k-928 and k-929, as well as hybrids between these lines differ in pollen fertility and the presence of large, giant and abnormal pollen grains. Since these lines have sterile cytoplasm, it can be assumed that they carry different alleles of genes derived from the recurrent parent that affect the manifestation of pollen fertility restoration. Grain sorghum lines resistant to *S. graminum* can serve as a model for studying mechanisms of pollen fertility restoration.

## REFERENCES

1. Ivanov M.K., Dymshits G.M. Cytoplasmic male sterility and restoration of pollen fertility in higher plants. *Russian Journal of Genetics*, 2007, 43(4): 354-368 (doi: 10.1134/S1022795407040023).
2. Chen L., Liu Y.G. Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.*, 2014, 65: 579-606 (doi: 10.1146/annurev-arplant-050213-040119).
3. Conner A.B., Karper R.E. Hybrid vigour in sorghum. *Texas Agricultural Experiment Station Bulletin*, 1927, 359: 21-26 (tsit. po B.V.S. Reddy et al., 2006).
4. Reddy B.V.S., Sharma H.C., Thakur R.P., Ramesh S., Rattunde F., Mgonja M. *Sorghum hybrid parents research at ICRISAT — strategies, status, and impacts. Journal of SAT Agricultural Research*, 2006, 2: 1-24.
5. Stephens J.C., Holland R.F. Cytoplasmic male sterility for hybrid sorghum seed production. *Agron. J.*, 1954, 46: 20-23 (doi: 10.2134/agronj1954.00021962004600010006x).
6. Pring D.R., Tang H.V., Schertz K.F. Cytoplasmic male sterility and organelle DNAs of sorghum. In: *Molecular biology of plant mitochondria*. C.S. Leavings III, I.K. Vasil (eds.). Kluwer, Dordrecht, The Netherlands 1995.
7. Elkonin L.A., Tsvetova M.I. Heritable effect of plant water availability conditions on restoration of male fertility in the 9E CMS-inducing cytoplasm of sorghum. *Front. Plant Sci.*, 2012, 3: 91 (doi: 10.3389/fpls.2012.00091).
8. Elkonin L.A., Domanina I.V., Gerashchenkov G.A., Rozhnova N.A. Inheritance of reversions to male fertility in male-sterile sorghum hybrids with 9E male-sterile cytoplasm induced by environmental conditions. *Russian Journal of Genetics*, 2015, 51(3): 251-261 (doi: 10.7868/S0016675815030030).
9. Kozhemyakin V.V., Elkonin L.A., Dahlberg J.A. Effect of drought stress on male fertility restoration in A3 CMS-inducing cytoplasm of sorghum. *The Crop Journal*, 2017, 5(4): 282-289 (doi: 10.1016/j.cj.2017.02.003).
10. Jordan D.R., Mace E.S., Henzell R.G., Klein P.E., Klein R.R. Molecular mapping and candidate gene identification of the *Rf2* gene for pollen fertility restoration in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor. Appl. Genet.*, 2010, 120(7): 1279-1287 (doi: 10.1007/s00122-009-1255-3).
11. Klein R.R., Klein P.E., Mullet J.E., Minx P., Rooney W.L., Schertz K.F. Fertility restorer locus *Rf1* of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the collinear region of rice chromosome 12. *Theor. Appl. Genet.*, 2005, 111(6): 994-1012 (doi: 10.1007/s00122-005-2011-y).
12. Reddy P.S., Rao D.M., Reddy V.S.B., Kumar A.A. Inheritance of male-fertility restoration in A1, A2, A3 and A4(M) cytoplasmic male-sterility systems of sorghum [*Sorghum bicolor* (L.) Moench]. *Indian Journal of Genetics*, 2010, 70(3): 240-246.
13. Jordan D.R., Klein R.R., Sakrewski K.G., Henzell R.G., Klein P.E., Mace E.S. Mapping and characterization of *Rf5* a new gene conditioning pollen fertility resto-

- ration in A<sub>1</sub> and A<sub>2</sub> cytoplasm in sorghum (*Sorghum bicolor* (L.) Moench). *Theor. Appl. Genet.*, 2011, 123(3): 383-396 (doi: 10.1007/s00122-011-1591-y).
14. Pring D.R., Tang H.V., Howad W., Kempken F. A unique two-gene gametophytic male sterility system in sorghum involving a possible role of RNA editing in fertility restoration. *J. Hered.*, 1999, 90: 386-393 (doi: 10.1093/jhered/90.3.386).
  15. Tang H.K., Pederson J.F., Chase C.D., Pring D.R. Fertility restoration of the sorghum A3 male-sterile cytoplasm through a sporophytic mechanism derived from sudangrass. *Crop Science*, 2007, 47: 943-950 (doi: 10.2133/cropsci2006.08.0542).
  16. Hanson M.R., Bentolila S. Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell*, 2004, 16(Suppl. 1): 154-169 (doi: 10.1105/tpc.015966).
  17. O'Toole N., Hattori M., Andres C., Iida K., Lurin C., Schmitz-Linneweber C, Sugita M, Small I. On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.*, 2008, 25(6): 1120-1128 (doi: 10.1093/molbev/msn057).
  18. Fujii S., Bond Ch.S., Small I.D. Selection patterns on restorer-like genes reveals a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. *PNAS*, 2011, 108(4): 1723-1728 (doi: 10.1073/pnas.1007667108).
  19. Dahan J., Mireau H. The *Rf* and *Rf*-like PPR in higher plants, a fast-evolving subclass of PPR genes. *RNA Biol.*, 2013, 10(9): 1469-1476 (doi: 10.4161/rna.25568).
  20. Gaborieau L., Brown G.G., Mireau H. The propensity of pentatricopeptide repeat genes to evolve into restorers of cytoplasmic male sterility. *Front. Plant Sci.*, 2016, 7: 1816 (doi: 10.3389/fpls.2016.01816).
  21. Kaur P., Verma M., Chaduvula P.K., Saxena S., Baliyan N., Junaid A., Mahato A.K., Sing N.K., Gaikwad K. Insights into PPR gene family in *Cajanus cajan* and other legume species. *Journal of Data Mining in Genomics and Proteomics*, 2016, 7: 203 (doi: 10.4172/2153-0602.1000203).
  22. Kibal'nik O.P. *Materialy Mezhdunarodnoi nauchno-prakticheskoi konferentsii, posvyashchennoi 128-i godovshchine so dnya rozhdeniya akademika N.I. Vavilova* [Proc. Int. Conf., dedicated to 128th Anniversary of N.I. Vavilov]. Saratov, 2015: 124-125 (in Russ.).
  23. Radchenko E.E., Malinovskaya E.V. *Zashchita i karantin rastenii*, 2012, 10: 24-25 (in Russ.).
  24. Radchenko E.E. Inheritance of greenbug resistance in several forms of grain sorghum and sudangrass. *Russian Journal of Genetics*, 2006, 42(1): 55-59 (doi: 10.1134/S1022795406010078).
  25. Roskin G.I. *Mikroskopicheskaya tekhnika* [Microscopy techniques]. Moscow, 1951 (in Russ.).
  26. Anisimova I.N., Alpat'eva N.V., Timofeeva G.I. *Skrining genicheskikh resursov rastenii s ispol'zovaniem DNK-markerov: osnovnye printsipy, vydelenie DNK, postanovka PTSR, elektroforez v agaroznom gele. Metodicheskie ukazaniya VIR* /Pod redaktsiei E.E. Radchenko [Screening plant genetic resources using DNA markers: basic principles, DNA isolation, PCR, agaros gel electrophoresis. E.E. Radchenko (ed.)]. St. Petersburg, 2010 (in Russ.).
  27. Paterson A.H., Bowers J.E., Bruggmann R., Dubchak I., Grimwood J., Gundlach H., Haberer G., Hellsten U., Mitros T., Poliakov A., Schmutz J., Spannagl M., Tang H., Wang X., Wicker T., Bharti A.K., Chapman J., Feltus F.A., Gowik U., Grigoriev I.V., Lyons E., Maher C.A., Martis M., Narechania A., Ollillar R.P., Penning B.W., Salamov A.A., Wang Y., Zhang L., Carpita N.C., Freeling M., Gingle A.R., Hash C.T., Keller B., Klein P., Kresovich S., McCann M.C., Ming R., Peterson D.G., Mehboob-ur-Rahman, Ware D., Westhoff P., Mayer K.F., Messing J., Rokhsar D.S. The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, 2009, 457(7229): 551-556 (doi: 10.1038/nature07723).
  28. Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, 2011, 28: 2731-2739 (doi: 10.1093/molbev/msr121).
  29. Desloire S., Gherbi H., Laloui W., Marhadour S., Clouet V., Cattolico L., Falentin L., Giancola S., Renard M., Budar F., Small I., Caboche M., Delourme R.M., Bendahmane A. Identification of the fertility restoration locus, *Rfo*, in radish, as a member of the pentatricopeptide-repeat protein family. *EMBO Rep.*, 2003; 4(6): 588-594 (doi: 10.1038/sj.embor.embor848).
  30. Cui X., Wise R.P., Schnable P.S. The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. *Science*, 1996, 272(5266): 1334-1336 (doi: 10.1126/science.272.5266.1334).
  31. Liu F., Cui X., Horner H.T., Weiner H., Schnable P.S. Mitochondrial aldehyde dehydrogenase activity is required for male sterility in maize. *The Plant Cell*, 2001, 13: 1063-1078 (doi: 10.1105/tpc.13.5.1063).
  32. Jimenez-Lopez J.C., Gachomo E.W., Seufferheld M.J., Kotchoni S.O. The maize ALDH protein superfamily: linking structural features to functional specificities. *BMC Struct. Biol.*, 2010, 10: 43 (doi: 10.1186/1472-6807-10-43).
  33. Tsvetova M.I., Ishin A.G. Large pollen grains as indicators of increased ploidy level due to colchicines treatment. *International Sorghum and Millets Newsletter*, 1995, 36: 77.