AGRICULTURAL BIOLOGY,
2021, V. 56, Iss. 5, pp. 885-898
[SEL'SKOKHOZYAISTVENNAYA BIOLOGIYA]

ISSN 2412-0324 (English ed. Online)
ISSN 0131-6397 (Russian ed. Print)
ISSN 2313-4836 (Russian ed. Online)

Genetics and genomics

UDC 635.64:581.192:575:577.2

doi: 10.15389/agrobiology.2021.5.885eng
doi: 10.15389/agrobiology.2021.5.885rus

STRUCTURAL AND FUNCTIONAL ANALYSIS OF *GME1* HOMOLOGOUS GENES AND ASCORBATE ACCUMULATION IN CULTIVATED AND WILD TOMATO SPECIES

D.Yu. TYAPKINA[1, 2] ✉, M.A. SLUGINA[2], E.Z. KOCHIEVA[2]

[1]*LLC Research Institute of Vegetable Breeding,* of. 1, 5, korp. 50, ul. 2-ya Entuziastov, Moscow, 111024 Russia, e-mail daria t@list.ru ✉;
[2]*Institute of Bioengineering, Federal Research Center Fundamentals of Biotechnology RAS,* 33/2, Leninskii prospect, Moscow, 119071 Russia, e-mail mashinmail@mail.ru, ekochieva@yandex.ru
ORCID:
Tyapkina D.Yu. orcid.org/0000-0001-8353-1170
Slugina M.A. orcid.org/0000-0003-1281-3837
Kochieva E.Z. orcid.org/0000-0002-6091-0765
The authors declare no conflict of interests
Received July 9, 2021

Abstract

Ascorbic acid (ascorbate, vitamin C) plays an important role in various metabolic processes both in plants and humans. Increasing the ascorbate content in plants using breeding approaches is important, both from the point of view of increasing the nutritional value of fruits, and from the point of view of plant resistance to stress. It is known that tomato has high potential as an ascorbate source in the human diet. Unfortunately, the ascorbate levels in ripe fruits of modern tomato (*Solanum lycopersicum*) cultivars and hybrids are low in comparison with relative wild tomato species. Use wild tomato accessions in breeding programs can significantly increase the ascorbate content of ripe fruits. However, for effective breeding for this trait, a more detailed study of the genetic determinants responsible for the ascorbate levels increase in in ripe fruits is necessary. In this study, we cloned and sequenced novel *GME1*gene homologous, which plays a key role in ascorbate biosynthesis in cultivated tomato and 11 wild tomato species. Structural analysis showed a low *GME1* variability level of in tomato species. In the *GME1* coding sequences, 28 SNPs were identified, of which only two led to nonsynonymous aminoacid substitutions (G2E and E281D) in *S. neorickii* and *S. peruvianum* var. *dentatum*. Analysis of GME1 motifs and domains did not reveal any specific motifs either at the interspecific level or at more distant taxonomic levels. The high *GME1* conservatism observed in quite evolutionarily distant tomato species is most likely due to the functional significance of this enzyme for the ascorbate synthesis and, indirectly, for protection from stress factors, primarily photostress. No correlation was found between amino acid or nucleotide substitutions and ascorbate levels in fruits. Expression analysis, including comparative interspecies organ-specific analysis and analysis of the dependence of the ascorbate content in mature fruits in tomato cultivars and wild species accessions and the *GME1* expression level, also did not reveal a relationship between transcriptional levels and ascorbate concentration. It can also be assumed that the final ascorbate content in the ripe tomato fruit may be influenced not by the intensity of GME1 expression at the last stage of fruit ripening, but by how this gene was active at earlier stages of ripening.

Keywords: *GME1,* gene expression, SNP, *Solanum lycopersicum*, wild tomato species, tomato cultivars, ripe fruits, ascorbate content

Ascorbic acid (ascorbate, vitamin C) plays an important role in various metabolic processes in plants in photosynthesis, photoprotection, resistance to stress, control of cell growth, biosynthesis of hormones and cell wall components [1-5]. In a plant cell, ascorbic acid acts as the main antioxidant due to its ability to reduce the content of reactive oxygen species, which are formed during photosynthesis and abiotic stresses (for example, at high light intensity, high temperatures, and strong ultraviolet radiation) [6, 7]. Humans and some primates are unable to synthesize ascorbate; therefore, plant foods rich in vitamin C are essential to maintain normal vital functions [8] and, thus, the search for new sources

of ascorbate is essential [9].

Tomato (*Solanum lycopersicum* L.) is one of the most commonly eaten vegetable crops with a wide cultivation area [10, 11]. Tomato varieties can serve as an additional source of vitamin C when consumed year-round [12]. However, in modern varieties and hybrids of cultivated tomato, the content of ascorbate in ripe fruits is low in comparison with that in related wild species [13]. Studies have shown that breeding programs involving wild tomato species, in particular *S. pennellii*, can significantly increase the ascorbate content in ripe fruits [14-18]. Nevertheless, effective breeding for this trait requires a more detailed study of the genetic determinants responsible for the ascorbate levels increase in ripe fruits.

One of the most important reactions of ascorbate biosynthesis in higher plant cells is considered to be the conversion of GDP-D-mannose to GDP-1-galactose, which is catalyzed by the enzyme GDP-mannose-3′, 5′-epimerase (GME; EC 5.1.3.18). Unlike most plants, which have only one copy of the GME gene, two copies have been identified in the tomato genome - GME1 and GME2 on chromosomes 1 and 9 [17, 19]. Analysis of QTL (quantitative trait loci) colocalization indicates that GME may serve as a genetic determinant of increased ascorbate levels in tomato fruits [17, 19]. This assumption is confirmed by work of L. Gilbert et al. [20], where the RNA silencing method was used to suppress the expression of both genes, *SlGME1* and *SlGME2*. Later C. Zhang et al. [21] demonstrated that overexpression of both genes significantly increased the ascorbate content in leaves and ripe tomato fruits. This directly affected the stress resistance of modified plants. A positive correlation between GME expression and ascorbate content was also shown for apple and blueberry [22, 23]. Analysis of the evolutionary patterns of the GME gene based on a comparison of 59 genomes of different species of higher plants and green algae revealed a high degree of conservatism of the sequence and structural organization of GME in higher plants [24]. However, the structure of GME gene homologous in wild tomato species has not yet been studied.

In this study, we were the first to identify and describe in detail *GME1* gene homologous in cultivated tomato cultivars *S. lycopersicum* and samples of wild species characterized by a high content of ascorbate in fruits. The variability of nucleotide and amino acid sequences, the composition of functional motifs in genes-homologs of *GME1* were determined. An analysis of the expression of such homologous genes in tomato samples with different contents of ascorbate in commercial products had shown that there is no relationship between this trait and the transcriptional activity of *GME1* in mature fruits.

Purpose of this study was to identify novel homologous GME1 gene and to conduct comparative analysis of their structure and transcriptional activity in different tomato species to identify possible correlations between the differences in these genes and the accumulation of ascorbate in fruits.

Materials and methods. Plant material included 15 tomato samples (section *Lycopersicon*, genus *Solanum*) — 11 wild species *S. chmielewskii* (VIR13725), *S. neorickii* (VIR5033), *S. chilense* (VIR4300), *S. corneliomulleri* (VIR4367), *S. peruvianum* (VIR4361), *S. peruvianum* var. *dentatum* (VIR3966), *S. arcanum* (VIR13958), *S. habrochaites* (VIR13964), *S. cheesmaniae* (VIR3969), *S. galapagense* (VIR3970), *S. pimpinellifolium* var. *racemigerum* (VIR1018), wild tomato specie *S. lycopersicum* var. *humboldtii* (VIR2912), tomato of *S. lycopersicum* Silvestre recordo (VIR1580), Bychye serdtse and Yellow belorus varieties (provided by the Vavilov All-Russian Institute of Plant Genetic Resources, Saint Petersburg and Federal Scientific Vegetable Center Gavrish LLC, Moscow Region). Plants were grown in a greenhouse (23 °C, day length 16 h, illumination 5 thousand lux/m^2).

When determining the content of ascorbate in fruits, the method of G. Giovanelli et al. [25] was used for its extraction with modifications. An aqueous solution of metaphosphoric acid (1 g/100 ml) was added to fresh tomatoes in a ratio of 1:10 (w/v) and homogenized in a Waring 8011 blender (Waring, USA). The resulting homogenate was quickly transferred into polypropylene tubes and centrifuged at 11000 g for 20 min at 4 °C. The analysis was performed by high performance liquid chromatography (HPLC, liquid chromatography with mass detector LC-QQQ Agilent 6460, Agilent Technologies, USA; ACE® 5 C18 column, Advanced Chromatography Technologies, Ltd., UK) (Shared-Access Equipment Centre Industrial Biotechnology of Federal Research Center Fundamentals of Biotechnology RAS, Moscow). The mobile phase was ultrapure water (pH 3), isocratic elution (flow rate 0.5 ml/min, temperature 35 °C). The supernatants were preliminarily filtered through membranes with a pore size of 0.45 rm (Amersham, USA). The volume of the analyzed samples was 20 ml. The concentration of ascorbic acid was expressed per 100 g of homogenized fruit pulp. The ascorbate content was determined in two biological replicates.

Genomic DNA was isolated from freshly harvested tomato leaves by method proposed by K. Edwards et al. [26]. The purity was evaluated and the concentration of the obtained DNA preparations was determined (spectrophotometer DU 530, Beckman, USA).

For the isolation of RNA with subsequent synthesis of cDNA, leaves, roots, flowers, and fruits (at the stage of full ripeness) of wild tomato species *S. lycopersicum* var. *humboldtii*, *S. peruvianum*, *S. habrochaites* and three tomato species var. *S. lycopersicum* Silvestre record, Bychye serdtse, and Yellow belorus were used. RNA was isolated using the RNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's recommendations. The resulting RNA preparations were treated with DNase I (10 units) (Invitrogen, USA) in accordance with the attached description. The cDNA preparations were obtained using the GoScript™ Reverse Transcription System kit (Promega, USA) according to the manufacturer's method. The concentration of RNA and cDNA was determined on a Qubit 4 fluorometer (Thermo Fisher Scientific, USA) using the appropriate reagents (Invitrogen, USA). Additionally, the quality of RNA was checked by electrophoresis in 1.5% agarose gel.

PCR amplification was performed on a Veriti™ 96-well Thermal Cycler (Applied Biosystems, USA) in a 15 µl reaction mixture containing 1.5 µl Dream™ Taq buffer 10×, 20 mM dNTP, 10 µM of specially developed GME primers F (5′-CACTGTATTAGTGCC-TCATC-3′) and GME R (5′-CAATTACCAGAATCT-AACACATC-3′), 0.25 µl Dream™ Taq DNA Polymerase (5 UI/µl) (Fermentas, Lithuania) and ~ 100 ng of genomic DNA. Amplification conditions: 5 min at 95 °C; 15 s at 95 °C, 10 s at 58 °C, 1 min at 72 °C (40 cycles); final elongation for 4 min at 72 °C.

The resulting PCR products were analyzed by electrophoresis in 1.5% agarose gel in 1× TBE buffer stained with ethidium bromide. The results were documented using the Gel Doc™ XR + Imaging System (Bio-Rad, USA). Commercial standards were used to determine the size of amplified DNA fragments 1 Kb DNA Ladder and 100 bp DNA Ladder (Fermentas, Lithuania)

PCR products were cloned using the pGEM®-T Easy Vector plasmid vector system (Promega, USA) according to the manufacturer's procedure, with additional deproteinization with a phenol: chloroform mixture (1:1).

The analysis of the expression of homologous *GME1* gene was carried out in four organs of tomato - root, leaf, flower, and fruit in biological ripeness by quantitative real-time PCR (RT-qPCR) using the kit Reaction mixture for carrying out RT-PCR in the presence of SYBR GreenI and ROX (Syntol LLC, Russia)

on a CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories, USA). The reactions were carried out in two biological and three analytical repeats. The following protocol was used to analyze the expression: 5 min at 95 °C; 15 s at 95 °C, 30 s at 60 °C (40 cycles), followed by reading the results. To obtain the melting curves of the amplification products, the temperature was increased from 55 °C to 95 °C (with a step of 0.5 °C every 5 s), followed by reading the results. To analyze the expression of homologous *GME1* gene, a pair of primers GMErnaF (5′-AGAATGGGAAGCTCTGGTGG-3′) and GMErnaR (5′-GGCTTCCAA-TTGAAATGATGACAG-3′) was developed allowing the amplification of a 185 bp fragment. During normalization, the expression value of the reference genes *Expressed* and *Actin 2/7* was used according to the description [27, 28].

Bioinformatic analysis of nucleotide sequences was performed using the MEGA 7.0 software [29] (https://www.kent.ac.uk/soft-ware/mega-7). Analysis of amino acid sequences and prediction of the presence of substitution sites that significantly affect the functionality of the enzyme were performed using the PROVEAN program [30] (http://provean.jcvi.org/index.php). The search for common hidden motives was carried out using the MEME 5.3.2 program [31] (https://meme-suite.org/meme/tools/meme). For additional analysis of amino acid sequences, the NCBI-CDD resource (https://www.ncbi.nlm.nih.gov/cdd/) and the UniProt database (https://www.uniprot.org) were used. For comparative analysis of nucleotide and amino acid sequences, sequences of *GME* genes-homologues from the NCBI database were also used, namely *S. lycopersicum* cv. Heinz 1706; *S. pimpinellifolium* (LA0480), *S. arcanum* (LA2157); *Oryza sativa* and *Arabidjpsis thaliana*.

The arithmetic mean values were calculated at determination of the content of ascorbate.

Results. Accumulation of ascorbate in ripe fruits. Table 1 shows the distribution of the studied samples by the content of vitamin C in fruits. Biochemical analysis showed that in the evolutionarily younger red-fruited tomato species *S. lycopersicum* and *S. pimpinellifolium* the accumulation of ascorbate occurs evenly as the fruit ripens, with a maximum in the fruit of biological ripeness. In the majority of green-fruited cross-pollinated tomatoes, the opposite dynamics is observed with the maximum content of ascorbate in the early stages of fruit development. Samples with the lowest and highest content of ascorbate in the fruit were identified at the stage of biological ripeness. Among the representatives of S. *lycopersicum*, the smallest amount of ascorbate was detected in the Yellow belorus (28.1 mg/100 g of homogenized fruit pulp) and Bychye serdtse (42.8 mg/100 g) samples, and the highest in *S. lycopersicum* var. *humboldtii* (120.1 mg/100 g). The obtained values slightly exceeded the data obtained earlier for cultivars and samples of wild species *S. pennellii* and *S. pimpinellifolium* [32], which may be due to the specificity of the manifestation of this trait in the studied cultivars and accessions, as well as the conditions for growing plants.

1. Vitamin C accumulation (*M*, mg/100 g pulp homogenate) **in tomato (*Solanum* L.) fruits at full ripeness**

Genotype	IG	MG	Br	Red ripe
S. lycopersicum Silvestre recordo (VIR1580) R/C	0.2	11.9	23.2	68.5
S. pimpinellifolium var. *racemigerum* (VIR1018) R/W	36.5	77.4	99.3	150.3
S. lycopersicum var. *humboldtii* (VIR2912) R/W	19.3	14.7	17.1	120.1
S. lycopersicum сорт Бычье сердце R/C	2.5	6.8	12.3	42.8
S. lycopersicum сорт Желтый белорус R/C	0.4	13.1	16.8	28.1
S. cheesmaniae (VIR3969) R/W	9.5	32.4	23.3	130.1
S. galapagense (VIR3970) R/W	34.2	46.4	65.7	84.2
S. chmielewskii (VIR13725) G/W	63.3	74.8	94.1	132.7
S. chilense (VIR4300) G/W	21.8	43.7	63.2	93.3

Continued Table 1

S. corneliomulleri (VIR4367) G/W	14.9	18.7	27.2	104.8
S. peruvianum (VIR4361) G/W	18.8	17.6	36.4	85.0
S. peruvianum var. *dentatum* (VIR3966) G/W	50.0	84.7	87.3	119.7
S. arcanum (VIR13958) G/W	66.9	65.2	37.8	44.7
S. habrochaites (VIR13964) G/W	0.2	0.2	0.3	0.3
S. neorickii (VIR5033) G/W	58.0	78.4	98.1	127.3

Note. R — red-fruited, G — green-fruited, W — wild, C — cultivated, IG — immature green, MG — mature green, Br — breaker. The measurements were carried out in two biological replicates (in generalized samples).

Identification and characterization of variability of GME1 homologous genes. When designing primer for DNA amplification of GME genes, being an enzyme involved in ascorbate synthesis, we used the genomes of tomato *S. lycopersicum* variety Heinz 1706 (NM_001247914.2), wild tomato *S. pennellii* (XM_015214839.2), and potatoes presented in the NCBI database. *S. tuberosum* (GCA_000226075.1) and mRNA sequences. As a result, a pair of flanking primers GME F–GME R was developed to amplify the *GME1* gene (Table 2). In addition, internal primers with localization in exon sequences were developed for sequencing. As a result, seven internal primers were obtained for amplification and sequencing of the *GME1* genes (see Table 2, Fig. 1).

2. Primers designed for amplification and sequencing of *GME1* homologous genes in the studied varieties and wild-growing tomato (*Solanum* L.) species

Primers	Sequences (5′→ 3′)
GME F	CACTGTATTAGTGCCTCATC
GME R	CAATTACCAGAATCTAACACATC
GME1exF	CAGCTTCAATATGATGGAAGCT
GME2exF	CAGCCTGAAAGAAGCTGATG
GME3exR	ACCTTTCCAGGTTCCAAAAGG
GME4exF	GCACAAACTGCAACGGATAAG
GME5exR	ATTGTTGGAGCCCAACCAAG
GME5exF	GCCAGTAAACATTGGAAGTGA
GME4inR	CTCCAGTAATTCAGAATAAGTGG

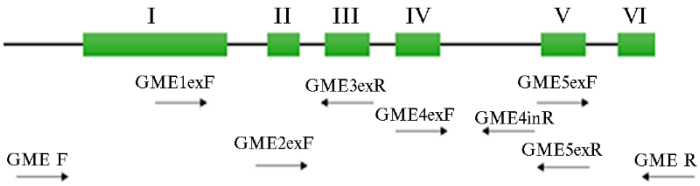

Fig. 1. Schematic representation of the structure of the *GME1* gene obtained during the analysis of *GME1* homologous genes in 11 wild tomato species and the tomato cultivar Silvestre recordo using the developed primers. The location of flanking and internal primers with localization in exons (green rectangles) and intron (lines) is marked. Exons are marked with Roman numerals. For a description of species and varieties, see the *Materials and methods* and Table 1.

Using the developed primers GME F–GME R, we have amplified the sequences of homologous *GME1* gene in 11 wild tomato species and the vegetable tomato cultivar Silvestre recordo as a representative of the *lycopersicum* species. The resulting fragments were cloned and sequenced using the internal primers.

Comparison of the structure of *GME1* homologous genes in the studied tomato samples showed that in all analyzed species this gene consists of six exons (see Fig. 1), which coincided with the data on the analysis of the exon-intron structure of the *S. lycopersicum GME1* gene obtained earlier [21]. The gene length varied from 1962 bp (*S. galapagense*) to up to 1985 bp (*S. peruvianum*). The size of the gene did not depend on whether the species was evolutionarily older (green-fruited) or younger (red-fruited). Comparison with the previously obtained data on the *GME* gene in the cultivated tomato cultivar Ailsa Craig also did not reveal a difference in the nucleotide sequences [21].

In the whole genome sequence of *GME1,* we found 139, or 6.9% single nucleotide polymorphisms (SNPs). Of these, 134 SNPs were found in green-fruited tomato species and only 5 in red-fruited species. The vast majority (108 SNPs) were found in intron sequences. Earlier, it was shown that the frequency of occurrence of nucleotide substitutions in intron sequences is on average 8±2% [33]. The level of intron polymorphism in *GME1* homologous genes described by us in this work generally corresponds to the given value. Also, as a result of the project The 100 Tomato Genome Sequencing Consortium, it was shown that the variability of gene sequences in red-fruited tomatoes is 20 times lower than in green-fruited species [33]. In our study, this difference for *GME1* turned out to be 26.8-fold, which is also consistent with previously published data [33].

In the coding region, 28 single nucleotide substitutions were found, the two SNPs in red-fruited species and 26 SNPs in green-fruited species (Fig. 2, see http://www.agrobiology.ru). The number of synonymous substitutions in the exon region ranged from 7-8 SNPs in *S. peruvianum* accessions to 12 and 13, respectively, in S. *arcanum* and *S. neorickii*. Only two substitutions, G2E and E281D, were found in green-fruited species *S. neorickii* and *S. peruvianum* var. *dentatum* were found to be nonsynonymous.

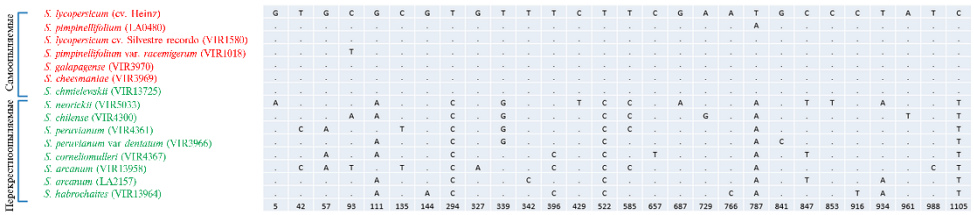

Fig. 2. Distribution of SNPs in the coding region of *GME1* homologues genes in the studied cross- and self-pollinated wild-growing species and varieties of tomato (Solanum L.). Numerical values indicate the location of single nucleotide substitutions. See the description of species and varieties in the *Materials and methods* and in Table 1.

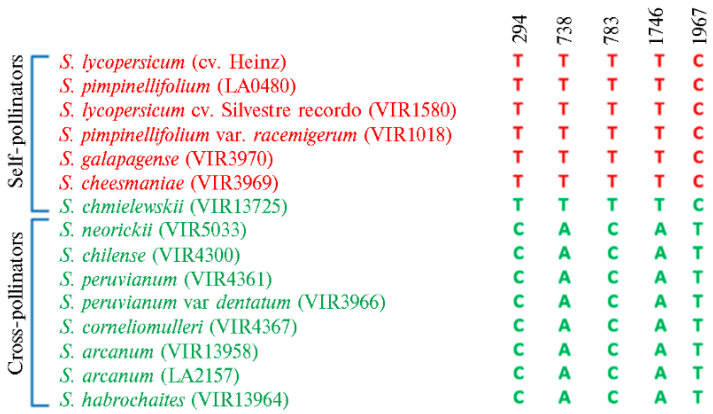

Fig. 3. Distribution of SNPs in *GME1* homologues genes in the studied cross- and self-pollinated wild-growing species and varieties of tomato (*Solanum* L.). The color indicates red (orange)- and green-fruited species. Numerical values indicate the location of single nucleotide substitutions. See the description of species and varieties in the *Materials and methods* and in Table 1.

In tomato, in addition to nucleotide substitutions in the analyzed sequences of *GME1* homologous gene, we found 22 indels ranging in size from 1 to 10 bp. All indels, as expected, were localized in introns, the maximum number of indels (seven) was in intron V. Interestingly, a 9-nucleotide indel (CCCTTGTA) was found in intron IV (at position 1237 from the start codon), which is present only in self-pollinated species, including the red-fruited *S. lycopersicum* (varieties

Heinz 1706 and Silvestre recordo), *S. pimpinellifolium* (LA0480), *S. pimpinellifolium* var. *racemigerum, S. cheesmaniae*, *S. galapagense*, and green-fruited *S. chmielewskii*. In addition to indels, in the analyzed *GME1* sequences, we found five SNPs, which are also characteristic of self-pollinated tomato species (Fig. 3).

3. Characterization of identified *GME1* homologous genes in the studied varieties and wild tomato (*Solanum* L.) species

Species	Gen-Bank NCBI record	Gene length, bp	Length			1	2	Isoelectric point, pI	Molecular weight, kDa
			introns, bp	cDNA, bp	protein, aar				
Red-fruited *Self-pollinated*									
S. lycopersicum cv. Silvestre recordo (VIR1580)	MK895092	1963	832	1131	376	0	0	5.65	42.456
S. pimpinellifolium var. *racemigerum* (VIR1018)	MK895094	1963	832	1131	376	1	0	5.65	42.456
S. cheesmaniae (VIR3969)	MK895095	1963	832	1131	376	0	0	5.65	42.456
S. galapagense (VIR3970)	MK895096	1962	831	1131	376	0	0	5.65	42.456
Green-fruited *Self-pollinated*									
S. chmielewskii (VIR13725)	MK895097	1966	836	1131	376	0	0	5.65	42.456
Cross-pollinated									
S. chilense (VIR4300)	MK895098	1973	842	1131	376	10	0	5.65	42.456
S. corneliomulleri (VIR4367)	MK895099	1981	850	1131	376	10	0	5.65	42.456
S. peruvianum (VIR4361)	MK895100	1985	854	1131	376	7	0	5.65	42.456
S. peruvianum var. *dentatum* (VIR3966)	MK895101	1982	851	1131	376	8	1	5.64	42.515
S. arcanum (VIR13958)	MK895102	1982	851	1131	376	12	0	5.65	42.456
S. habrochaites (VIR13964)	MK895103	1975	844	1131	376	10	0	5.65	42.456
S. neorickii (VIR5033)	MK895093	1981	850	1131	376	13	1	5.56	42.601

Note. 1 and 2 stand for the number of SNPs and amino acid substitutions compared to the reference sequence (*S. lycopersicum*, cv. Heinz 1706; GeneBank NCBI). The data for the Silvestre recordo cultivar are given as typical for the studied cultivars. See the description of species and varieties in the *Materials and methods* and in Table 1.

The coding sequences of *GME1* homologous gene were translated. The obtained protein sequences were of the same length, 376 amino acid residues, which corresponded to a molecular weight of 42 kDa (Table 3). In general, the amino acid sequences of the proteins were found to be similar. In GME1 proteins in green-fruited species *S. neorickii* and *S. peruvianum* var. *dentatum*, two amino acid substitutions were present, G2E and E281D. The analysis of these substitutions using the PROVEAN program showed that they are not radical and, presumably, cannot affect the functionality of GME1. The results obtained confirm the data that GME is one of the most conserved proteins involved in the biosynthesis of ascorbate [24, 34].

It should be noted that both samples in which nonsynonymous substitutions were identified, *S. neorickii* and *S. peruvianum* var. *dentatum* did not differ significantly from other samples in the content of ascorbic acid (see Table 1) and the dynamics of its accumulation.

In all studied tomato species, analysis of the amino acid sequences of GME1 using the UniProt database [35] revealed an epimerase domain (positions 30-269 aar). Within the domain, for all analyzed proteins, five binding sites for the GME substrate, GDP-mannose (G105, N204, K226, R307, and S357), as well as four binding sites for NAD+, the GME cofactor (D60, D80, Y175, and

K179).

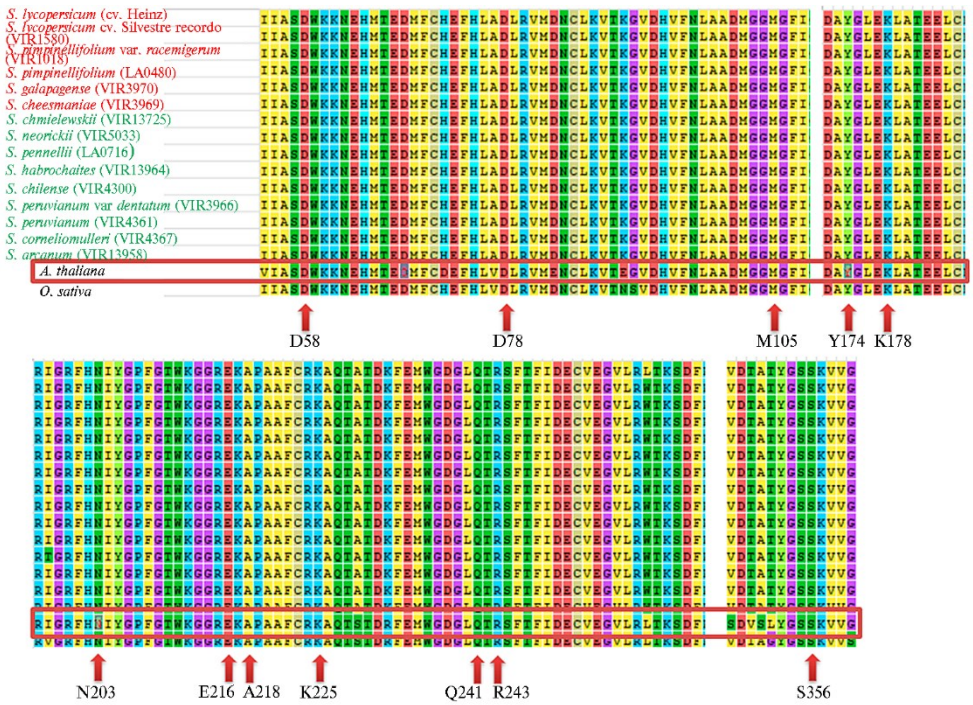

Fig. 4. Conservative amino acid residues in GME1 proteins in the studied cultivars and wild-growing tomato (*Solanum* L.) species, which are critical for the functionality of the GME enzyme (GDP-mannose-3′,5′-epimerase) (shown by arrows). For comparison, the amino acid sequences of the GME homologues of rice and *Arabidopsis* were taken from the NCBI database. *Materials and methods* and in Table 1.

The structure and function of GME in *Arabidopsis thaliana* was previously determined [36]. Our comparative analysis using the MEGA 7.0 software package showed a high degree of homology of this protein in *Arabidopsis* and tomato, which suggests a functional similarity of the amino acid sequence of GME1 in tomato and *Arabidopsis*. Thus, based on the data of site-directed mutagenesis of the GME gene in *Arabidopsis*, it was shown that amino acids C145 and K217 are responsible for the activity of oxidative and reductive epimerization [36]. These amino acid residues coincide with those in the tomato GME1 protein and, therefore, can also be decisive for its functional activity. Amino acids C145 and K217 were invariant in all GME1 sequences analyzed by us in the studied tomato samples (Fig. 4).

The analysis also showed that GME1 is a member of the extended short chain dehydratase/reductase (SDR) family, albeit with a modified glycine-rich nucleotide binding motif [37] (GAGGFIA instead of GXXGXXG) [36]. In the tomato species we studied, GME1 also revealed a conservative motif GAGGFIA at positions 34-40 aar. Such a high conservatism of this motif in GME proteins is most likely determined by its critical function, which is due to its participation in the formation of hydrogen bonds with NAD+.

Besides, we analyzed the amino acid sequences that we obtained for GME1 in the studied tomato species and in homologues of other species (*O. sativa, A. thaliana*) using the MEME 5.3.2 program, which made it possible to identify 10 conservative motifs. All detected motifs (see additional materials on the website http://www.agrobiology.ru) showed high conservatism in comparison with homologues in other representatives of the genus *Solanum* and with the GME homologue of *Arabidopsis*. Our results confirmed the data on the similarity of both

GME1 and *GME2* homologues in plants [24]. In the second motif (12-61 aar), the sequence GAGGFIA was identified, the nucleotide-binding functions of which were previously determined for *A. thaliana* [36].

We also compared the identified motifs using the NCBI-CDD resource, which made it possible to classify them as NAD-binding domains — NAD(P)(+)-binding Rossmann-fold (NADB_Rossmann). The NADB domain is characteristic of many metabolic pathway dehydrogenases and various redox enzymes. It is believed that the presence of this domain is due to the epimerase activity of GME1 [24].

Thus, the analysis of the motifs and domains of the GME1 protein in the studied tomato samples and the comparison with the homologous proteins of rice and Arabidopsis did not reveal any specific motifs either at the interspecific level or at more distant taxonomic levels. The observed low level of variability of GME1 proteins in evolutionarily quite distant tomato species, for example, *S. neorickii* and S. *lycopersicum* (cv. Silvestre recordo), and the presence of only two radical substitutions confirms its conservatism, shown earlier [34], and, most likely, indicates on the functional significance of this enzyme (as well as the biochemical reaction catalyzed by it) for the synthesis of ascorbate and, indirectly, for protection from stress factors, primarily from photostress.

Analysis of the expression of *GME1* homologous genes. Since the structural analysis of the sequences of *GME1* homologues genes did not reveal polymorphisms correlated with the accumulation of ascorbate, we assumed that the effect of the GME1 enzyme on the biosynthesis of ascorbate could occur at the transcriptional level.

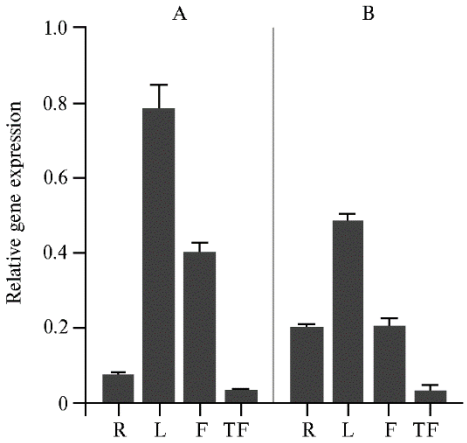

Fig. 5. Expression of homologous *GME1* genes in organs of two tomato species with a contrasting content of ascorbate in fruits : A — *Solanum lycopersicum* (var. Silvestre recordo), B — *S. peruvianum*; R — root, L — leaf, Fl — flower, Fr — fruit in biological ripening phase. Expression was normalized using two standard reference genes, the *Expressed* and *Actin 2/7*.

When studying the organ-specificity of the transcriptional activity of *GME1* homologous genes, we compared the transcriptional activity of this gene in four organs (root, leaf, flower, and fruit at the stage of biological ripeness) in *S. lycopersicum* plants of the Silvestre recordo cultivar and a sample of the wild species *S. peruvianum* with a contrasting content of ascorbate in fruits (see Table. 1, Fig. 5).

The analysis revealed the highest transcriptional activity of *GME1* in leaves regardless of the tomato species and the accumulation of ascorbate in fruits. The result is anticipated, since the leaves are most susceptible to photooxidative stress. The lowest level of *GME1* gene expression was observed in fruits (see Fig. 5).

However, it is the content of ascorbate in fruits (and in particular, the elucidation of the mechanisms that ensure the formation of this trait) that is of the greatest practical interest. To study in more detail the possible relationship between the accumulation of ascorbate in commercial products and the number of *GME1* transcripts, we determined the level of *GME1* expression in fruits of biological ripeness in five samples, the vegetable tomato *S. lycopersicum* var. Silvestre recordo, Bychye serdtse, and Yellow belorus, and samples of wild species *S. lycopersicum* var. *humboldtii* and *S. peruvianum* (Fig. 6, A). These samples are contrasting in the ascorbate content in fruits

and represented both varieties and a wild-growing sample of a cultivated red-fruited species and a wild-growing green-fruited species. In the same five samples, we annotated the expression of *GME1* in mature fruits (see Fig. 6, B).

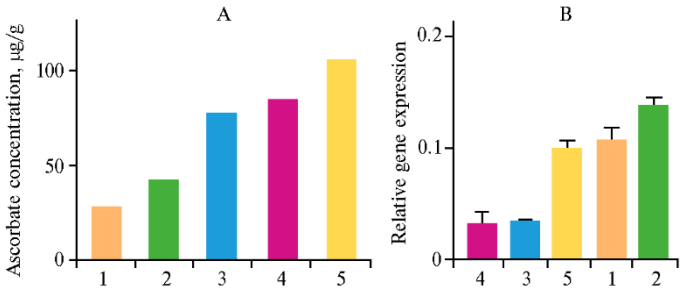

Fig. 6. Comparison of the ascorbate content (A) **and *GME1* expression** (B) **in fruits at the stage of biological maturity of tomato lines and cultivars *Solanum lycopersicum* Yellow belorus** (1)**, Bychye serdtse** (2)**, Silvestre recordo** (3) **and wild species *S. peruvianum*** (4)**, and *S. lycopersicum* var. *humboldtii*** (5). Expression was normalized using two standard reference genes, the *Expressed* and *Actin 2/7*.

Comparison of the results of *GME1* gene expression and biochemical analysis (see Fig. 6) does not reveal a clear correlation between the obtained indicators; we can only speak of a weak inverse relationship between gene expression in mature fruits and the content of ascorbate in them. It can also be assumed that the final ascorbate content in the ripe fruit in the analyzed tomato species may be influenced not by the intensity of *GME1* expression at the last stage of fruit ripening, but by the extent to which this gene was active at earlier stages of ripening, which was shown earlier for *S. lycopersicum* cv. West Virginia 106 [38]. To confirm or to refute this assumption, it is necessary to study the activity of *GME1* in the process of fruit ripening. In addition, possible alternative metabolic pathways for the accumulation of ascorbate in ripe tomato fruits should be considered, for example, the L-galactose pathway, which uses polysaccharides from the fetal cell wall as a substrate.

The biochemical analysis of the ascorbate content in mature fruits in the studied species of tomato did not show a correlation with the levels of *GME1* transcription (see Fig. 6) or identified SNPs. A possible explanation, in addition to the hypothesis formulated above that the transcriptional activity of the *GME1* gene can affect the ascorbate content in fetuses at stages preceding full maturity, is that *GME1*, in addition to biosynthesis of ascorbate, is actively involved in cell wall synthesis during fruit growth [38]. It should also be taken into account that, in addition to the L-galactose pathway, other alternative pathways of biosynthesis, for example, myo-inositol pathway can be used for the formation of ascorbate in tomato fruits [39].

Thus, in this work, we were the first to obtain and analyze the sequences of *GME1* homologous genes in 11 species of cultivated and wild tomato. The performed structural and functional analysis showed an extremely high conservatism of *GME1* homologous genes and proteins encoded by them both in the evolutionarily more ancient green-fruited species and evolutionarily younger red-fruited tomatoes. In the *GME1* homologous genes, 28 SNPs were found that are characteristic of self-pollinated tomato species, which can be used in breeding work. No correlation was found between the content of ascorbate in mature fruits, on the one hand, and the detected SNPs or levels of the *GME1* transcription, on the other hand. The results of this study show that the GME1 protein in the studied tomato species does not have a significant effect on the accumulation of

ascorbate in ripe fruits and indirectly confirm that the biosynthesis of ascorbate in tomato fruits can be provided by other alternative pathways, such as myo-inositol or L-galactose.

REFERENCES

1. Barth C., De Tullio M., Conklin P. The role of ascorbic acid in the control of flowering time and the onset of senescence. *Journal of Experimental Botany*, 2006, 57(8): 1657-1665 (doi: 10.1093/jxb/erj198).
2. Fotopoulos V., De Tullio M.C., Barnes J., Kanellis K. Altered stomatal dynamics in ascorbate oxidase over-expressing tobacco plants suggest a role for dehydroascorbate signalling. *Journal of Experimental Botany*, 2008, 59(4): 729-737 (doi: 10.1093/jxb/erm359).
3. Foyer C.H., Noctor G. Ascorbate and glutathione: the heart of the redox hub. *Plant Physiology*, 2011, 155(1): 2-18 (doi: 10.1104/pp.110.167569).
4. Fotopoulos V., Kanellis A.K. Altered apoplastic ascorbate redox state in tobacco plants via ascorbate oxidase overexpression results in delayed dark-induced senescence in detached leaves. *Plant Physiology and Biochemistry*, 2013, 73: 154-160 (doi: 10.1016/j.plaphy.2013.09.002).
5. Smirnoff N. Ascorbic acid metabolism and functions: a comparison of plants and mammals. *Free Radical Biology and Medicine*, 2018, 122: 116-129 (doi: 10.1016/j.freeradbiomed.2018.03.033).
6. Conklin P.L., Barth C. Ascorbic acid, a familiar small molecule intertwined in the response of plants to ozone, pathogens, and the onset of senescence. *Plant Cell and Environment*, 2004, 27(8): 959-970 (doi: 10.1111/j.1365-3040.2004.01203.x).
7. Müller-Moulé P., Golan T., Niyogi K.K. Ascorbate-deficient mutants of *Arabidopsis* grow in high light despite chronic photooxidative stress. *Plant Physiology*, 2004, 134(3): 1163-1172 (doi: 10.1104/pp.103.032375).
8. Linster C.L., Van Schaftingen E. Vitamin C. Biosynthesis, recycling and degradation in mammals. *FEBS Journal*, 2007, 274(1): 1-22 (doi: 10.1111/j.1742-4658.2006.05607.x).
9. Granger M., Eck P. Dietary vitamin C in human health. *Advances in Food and Nutrition Research*, 2018, 83: 281-310 (doi: 10.1016/bs.afnr.2017.11.006).
10. Celma A.R., Cuadros F., López-Rodríguez F. Characterisation of industrial tomato by-products from infrared drying process. *Food and Bioproducts Processing*, 2009, 87(4): 282-291 (doi: 10.1016/J.FBP.2008.12.003).
11. FAOSTAT. Food and Agriculture Organization of the United Nations Statistical Databases, 2019. Available: http://faostat.fao.org/. No date.
12. Wang F., Kang S.Z., Du T.S., Li F.S., Qiu R.J. Determination of comprehensive quality index for tomato and its response to different irrigation treatments. *Agricultural Water Management*, 2011, 98(8): 1228-1238 (doi: 10.1016/j.agwat.2011.03.004).
13. Di Matteo A., Sacco A., Anacleria M., Pezzotti M., Delledonne M., Ferrarini A., Frusciante L., Barone A. The ascorbic acid content of tomato fruits is associated with the expression of genes involved in pectin degradation. *BMC Plant Biology*, 2010, 10(1): 1-11 (doi: 10.1186/1471-2229-10-163).
14. Rousseaux M.C., Jones C.M., Adams D., Chetelat R., Bennett A., Powell, A. QTL analysis of fruit antioxidants in tomato using *Lycopersicon pennellii* introgression lines. *Theoretical and Applied Genetics*, 2005, 111(7): 1396-1408 (doi: 10.1007/s00122-005-0071-7).
15. Schauer N., Semel Y., Roessner U., Gur A., Balbo I., Carrari F., Pleban T., Perez-Melis A., Bruedigam C., Kopka J., Willmitzer L., Zamir D., Fernie A.R. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology*, 2006, 24(4): 447-454 (doi: 10.1038/nbt1192).
16. Schauer N., Semel Y., Balbo I., Steinfath M., Repsilber D., Selbig J., Pleban T., Zamir D., Fernie A.R. Mode of inheritance of primary metabolic traits in tomato. *The Plant Cell*, 2008, 20(3): 509-523 (doi: 10.1105/tpc.107.056523).
17. Stevens R., Buret M., Duffe P., Garchery C., Baldet P., Rothan C., Causse M. Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three tomato populations. *Plant Physiology*, 2007, 143(4): 1943-1953 (doi: 10.1104/pp.106.091413).
18. Stevens R., Page D., Gouble B., Garchery C., Zamir D., Causse M. Tomato fruit ascorbic acid content is linked with monodehydroascorbate reductase activity and tolerance to chilling stress. *Plant, Cell and Environment*, 2008, 31(8): 1086-1096 (doi: 10.1111/j.1365-3040.2008.01824.x).
19. Zou L.P., Li H.X., Ouyang B., Zhang J.H., Ye Z.B. Cloning, expression, and mapping of GDP-D-mannose pyrophosphorylase cDNA from tomato (*Lycopersicon esculentum*). *Yi Chuan Xue Bao*, 2006, 33(8): 757-764 (doi: 10.1016/S0379-4172(06)60108-X).
20. Gilbert L., Alhagdow M., Nunes-Nesi A., Quemener B., Guillon F., Bouchet B., Faurobert M., Gouble B., Page D., Garcia V., Petit J. GDP-D-mannose 3,5-epimerase (GME) plays a key role at the intersection of ascorbate and non-cellulosic cell-wall biosynthesis in tomato. *The Plant Journal*, 2009, 60(3): 499-508 (doi: 10.1111/j.1365-313X.2009.03972.x).

21. Zhang C., Liu J., Zhang Y., Cai X., Gong P., Zhang J., Wang T., Li H., Ye Z. Overexpression of SlGMEs leads to ascorbate accumulation with enhanced oxidative stress, cold, and salt tolerance in tomato. *Plant Cell Rep.*, 2011, 30(3): 389-398 (doi: 10.1007/s00299-010-0939-0).
22. Li M., Ma F., Liang D., Li J., Wang Y. Ascorbate biosynthesis during early fruit development is the main reason for its accumulation in kiwi. *PLoS ONE*, 2010, 5(12): 14281 (doi: 10.1371/journal.pone.0014281).
23. Liu F., Wang L., Gu L., Zhao W., Su H., Cheng X. Higher transcription levels in ascorbic acid biosynthetic and recycling genes were associated with higher ascorbic acid accumulation in blueberry. *Food Chemistry*, 2015, 188: 399-405 (doi: 10.1016/j.foodchem.2015.05.036).
24. Tao J., Wu H., Li Z., Huang C., Xu X. Molecular evolution of GDP-D-mannose epimerase (GME), a key gene in plant ascorbic acid biosynthesis. *Frontiers in Plant Science*, 2018, 9: 1293 (doi: 10.3389/fpls.2018.01293).
25. Giovanelli G., Zanoni B., Lavelli V., Nani R. Water sorption, drying and antioxidant properties of dried tomato products. *Journal of Food Engineering*, 2002, 52(2): 135-141 (doi: 10.1016/S0260-8774(01)00095-4).
26. Edwards K., Johnstone C., Thompson C. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, 1991, 19(6): 1349 (doi: 10.1093/nar/19.6.1349).
27. González-Aguilera K.L., Saad C.F., Chávez Montes R.A., Alves-Ferreira M., de Folter S. Selection of reference genes for quantitative real-time RT-PCR studies in tomato fruit of the genotype MT-Rg1. *Frontiers in Plant Science*, 2016, *7*: 1386 (doi.org/10.3389/fpls.2016.01386).
28. Slugina M.A., Filyushin M.A., Shchennikova A.V., Kochieva E.Z., Skryabin K.G. FAS, YABBY2, and YABBY5 gene expression profile correlates with different fruit locule number in tomato. *Russian Journal of Genetics*, 2020, 56(4): 401-407 (doi: 10.31857/S0016675820030157).
29. Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 2016, 33(7): 1870-1874 (doi: 10.1093/molbev/msw054).
30. Choi Y., Sims G.E., Murphy S., Miller J.R., Chan A.P. Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE,* 2012, 7(10): e46688 (doi: 10.1371/journal.pone.0046688).
31. Bailey T.L., Boden M., Buske F.A. "MEME SUITE: tools for motif discovery and searching". *Nucleic Acids Research*, 2009, 37(2): 202-208 (doi: 10.1093/nar/gkp335).
32. Mellidou I., Koukounaras A., Kostas S., Patelou E., Kanellis A.K. Regulation of vitamin C accumulation for improved tomato fruit quality and alleviation of abiotic stress. *Genes*, 2021, 2(5): 694 (doi: 10.3390/genes12050694).
33. The 100 Tomato Genome Sequencing Consortium, 2014. Available: https://solgenomics.net/projects/tomato100/. No date.
34. Wolucka B.A., Van Montagu M. The VTC2 cycle and the de novo biosynthesis pathways for vitamin C in plants: an opinion. *Phytochemistry*, 2007, 68(21): 2602-2613 (doi: 10.1016/j.phytochem.2007.08.034).
35. Bairoch A., Apweiler R., Wu C.H., Barker W.C., Boeckmann B., Ferro S., Gasteiger E., Huang H., Lopez R., Magrane M., Martin M.J., Natale D.A., O'Donovan C., Redaschi N., Yeh L.S. The Universal Protein Resource (UniProt). *Nucleic Acids Research*, 2004, 33(1): 154-159 (doi: 10.1093/nar/gki070).
36. Major L.L., Wolucka B.A., Naismith J.H. Structure and function of GDP-mannose-3',5'-epimerase: an enzyme which performs three chemical reactions at the same active site. *Journal of the American Chemical Society*, 2005, 127(51): 18309-18320 (doi: 10.1021/ja056490i).
37. Wierenga R.K., De Maeyer M.C.H., Hol W.G.J. Interaction of pyrophosphate moieties with α-helixes in dinucleotide binding proteins. *Biochemistry*, 1985, 24(6): 1346-1357 (doi: 10.1021/bi00327a012).
38. Mounet-Gilbert L., Dumont M., Ferrand C., Bournonville C., Monier A., Jorly J., Lemaire-Chamley M., Mori K., Atienza I., Hernould M., Stevens R. Two tomato GDP-D-mannose epimerase isoforms involved in ascorbate biosynthesis play specific roles in cell wall biosynthesis and development. *Journal of Experimental Botany*, 2016, 67(15): 4767-4777 (doi: 10.1093/jxb/erw260).
39. Munir S., Mumtaz M.A., Ahiakpa J.K., Liu G., Chen W., Zhou G., Zheng W., Ye Z., Zhang Y. Genome-wide analysis of Myo-inositol oxygenase gene family in tomato reveals their involvement in ascorbic acid accumulation. *BMC Genomics*, 2020, 21(1): 1-15 (doi: 10.1186/s12864-020-6708-8).