

DOI:10.30901/2227-8834-2018-4-91-103

ОРИГИНАЛЬНАЯ СТАТЬЯ

УДК 575.11:577.21:633.854

**И. Н. Анисимова<sup>1</sup>,  
Н. В. Алпатьева<sup>1</sup>,  
С. В. Горюнова<sup>2,4</sup>,  
Д. В. Горюнов<sup>3,4</sup>,  
Ал. В. Конарев<sup>5</sup>,  
В. А. Гаврилова<sup>1</sup>,  
Е. Е. Радченко<sup>1</sup>**

<sup>1</sup>Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений имени Н. И. Вавилова, 190000, Россия, Санкт-Петербург, ул. Б. Морская, д. 42, 44, e-mail: irina\_anisimova@inbox.ru

<sup>2</sup>Федеральное государственное бюджетное учреждение науки Институт общей генетики им. Н. И. Вавилова Российской академии наук, 119991, Россия, Москва, ул. Губкина, д. 3, e-mail: orang2@yandex.ru

<sup>3</sup>Научно-исследовательский институт физико-химической биологии имени А. Н. Белозерского Московского государственного университета имени М. В. Ломоносова, 119992, Россия, Москва, Ленинские горы, д. 1, e-mail: s.arosarpum@gmail.com

<sup>4</sup>Автономная некоммерческая образовательная организация высшего профессионального образования «Сколковский институт науки и технологий», 121205, Россия, Москва, ул. Нобеля, д. 3

<sup>5</sup>Федеральное государственное бюджетное научное учреждение Всероссийский научно-исследовательский институт защиты растений, 196608, Россия, Санкт-Петербург, Пушкин, шоссе Подбельского, д. 3, e-mail: al\_konarev@hotmail.com

**Ключевые слова:**

подсолнечник, 2S альбумины, богатый метионином белок SFA8, ген, полиморфизм, нуклеотидная и аминокислотная последовательности, интрон, структура

**Поступление:**

28.11.2018

**Принято:**

10.12.2018

## СТРУКТУРНАЯ ИЗМЕНЧИВОСТЬ ГЕНА БОГАТОГО МЕТИОНИНОМ АЛЬБУМИНА SFA8 ПОДСОЛНЕЧНИКА

**Актуальность.** 2S альбумины семян подсолнечника и других масличных растений обладают высокой питательной ценностью, защитной активностью против возбудителей грибных болезней, а также ценными функциональными свойствами. Главный компонент альбуминовой фракции – белок SFA8 содержит 103 аминокислотных остатка, из которых 15 приходится на долю метионина – незаменимой аминокислоты. Структурный ген SFA8 представлен в генофонде однолетнего культурного подсолнечника двумя аллелями, продукты которых имеют различные изоэлектрические точки и отличаются по электрофоретической подвижности, однако молекулярные механизмы полиморфизма неизвестны. **Результаты.** Секвенированы последовательности амплифицированных фрагментов гена SFA8 семи образцов *Helianthus annuus* L. и трех образцов диких видов рода *Helianthus* L. из коллекции ВИР. Впервые установлено, что в центральной части гена SFA8 находится интрон, длина которого у разных генотипов варьирует от 258 до 303 пн. Длина первого экзона составляет 99 пн., длина второго – 210 пн. Нуклеотидные и транслированные аминокислотные последовательности полиморфны у различных генотипов. Линия ВИР130, у которой ранее обнаружены два экспрессирующихся варианта белка – нормальный с изоэлектрической точкой (pI) 6,0 («дикий тип») и его аллельный вариант с pI 6,5 («вариантный белок») – имеет два варианта кодирующей последовательности. В одной из них присутствует замена 108С→G, приводящая к замене полярной незаряженной аминокислоты серина на положительно заряженный аргинин и, соответственно, изменению заряда белка и его изоэлектрической точки. Последовательность интрона также полиморфна и характеризуется наличием инделей длиной около 45 пн. У всех образцов последовательности интрона содержат динуклеотиды GT на 5' и AG на 3' конце, характерные для консенсусных последовательностей сайтов сплайсинга интронов U2-типа. Варианты вторичной структуры последовательностей интрона SFA8 *H. argophyllus* Torr. & A.Gray и всех проанализированных генотипов *H. annuus* сходны между собой, но отличаются от вариантов *H. petiolaris* Nutt. и *H. giganteus* L. **Выводы.** Полученные данные о полиморфизме последовательностей гена SFA8 важны для объяснения молекулярных механизмов генотипических различий биохимических и функциональных свойств кодируемого белка, а выявленные различия вторичной структуры интрона – для понимания особенностей экспрессии этого белка.

DOI:10.30901/2227-8834-2018-4-91-103

ORIGINAL ARTICLE

I. N. Anisimova<sup>1</sup>,  
 N. V. Alpatieva<sup>1</sup>,  
 S. V. Goryunova<sup>2,4</sup>,  
 D. V. Goryunov<sup>3,4</sup>,  
 Al. V. Konarev<sup>5</sup>,  
 V. A. Gavrilova<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, 42, 44 Bolshaya Morskaya St., St. Petersburg, 190000, Russia, e-mail: irina\_anisimova@inbox.ru

<sup>2</sup>Vavilov Institute of General Genetics Russian Academy of Sciences, 3 Gubkina str., Moscow, 119333, Russia e-mail: orang2@yandex.ru

<sup>3</sup>A. N. Belozersky Institute of Physico-Chemical Biology MSU, 1 Leninskye gory, building 40, Moscow, 119992, Russia e-mail: s.apocarpum@gmail.com

<sup>4</sup>Skolkovo Institute of Science and Technology, 3 Nobelya str., Moscow, 121205, Russia

<sup>5</sup>All-Russian Institute of Plant Protection, 3 Podbelskogo schosse, St.Petersburg – Pushkin, 196608, Russia e-mail: al\_konarev@hotmail.com

**Key words:**

sunflower, 2S albumins, methionine-rich protein SFA8, gene, polymorphism, nucleotide and amino acid sequences, intron, structure

**Received:**

28.11.2018

**Accepted:**

10.12.2018

## STRUCTURAL VARIABILITY OF SUNFLOWER GENE FOR METHIONINE-RICH ALBUMIN SFA8

**Background.** The 2S albumins of sunflower and other oilseed plants possess a high nutritional quality, the defense activity against fungi diseases casual agents and also valuable functional properties. The major component of albumin fraction, the SFA8 protein consists of 103 amino acid residues among which methionine constitutes 15 Mole %. In the cultivated sunflower gene pool the *SFA8* structural gene is represented by the two alleles the products of which have different isoelectric points and differ by the electrophoretic mobility, however molecular mechanisms of the polymorphism are still unknown. **Results.** The amplified sequences of the *SFA8* gene from seven *Helianthus annuus* L. accessions and three accessions of wild *Helianthus* L. species from VIR collection were sequenced. The intron of 258-303 bp length depending on the genotype was firstly found in the central part of the gene. The length of the first exon constitutes 99 bp, the second exon is of 210 bp length. The nucleotide and translated amino acid sequences are polymorphic among different genotypes. The line VIR 130 in which the two expressing SFA8 proteins, the normal polypeptide with isoelectric point (pI) approximately 6.0 (normal SFA8) and its allelic variant with pI 6.5 (variant SFA8) have been earlier revealed possesses two types of the *SFA8* encoding sequence. In one sequence the substitution 108C→G is present that results in the substitution of the polar uncharged amino acid serine for the positively charged arginine and respectively in alteration of the protein charge and isoelectric point. The intron sequence is also polymorphic and characterized by the presence of indels of approximately 45 bp. The intron sequences of all accessions contain dinucleotides GT at the 5' end and AG at the 3' end which are characteristic for consensus sequences of splicing sites in the U2-type introns. The variants of the secondary structure of the *SFA8* intron sequences of *H. argophyllus* Torr. & A. Gray and all the analyzed *H. annuus* genotypes are similar and differ from those of *H. petiolaris* Nutt. and *H. giganteus* L. **Conclusions.** The data on the *SFA8* gene sequence polymorphism are important understanding the molecular mechanisms of genotypic differences in biochemical and functional properties of the protein, and he revealed differences in the intron secondary structure can be important for understanding expression patterns of the protein.

## Введение

Семена растений накапливают значительные количества запасных белков, необходимых для питания прорастающего зародыша. Запасные белки представляют значительную ценность как источники незаменимых аминокислот в питании человека и животных. Белковая фракция семян масличных растений, наряду с компонентами глобулиновой природы, включает и низкомолекулярные водорастворимые белки, относящиеся к классу 2S альбуминов. 2S альбумины семян масличных растений обладают рядом уникальных свойств – высокой питательной ценностью за счет исключительно высокого содержания незаменимой аминокислоты метионина (Youle, Huang, 1981), а также защитной активностью против возбудителей грибных болезней, которая определяется наличием доменов, характерных для семейства ингибиторов трипсина/альфа-амилазы (Agizzio et al., 2006; Odintsova et al., 2010). Следует отметить также высокую аллергенность 2S альбуминов, обусловленную присутствием в молекуле консервативных IgE-связывающих эпитопов (Radauer, Breiteneder, 2007). Поэтому в последние годы резко возрос интерес к этой группе белков как объектам для изучения питательных и функциональных свойств, а также клинических исследований. К настоящему времени охарактеризованы главные альбумины клещевины, рапса, желтой горчицы, бразильского ореха и других растений; установлено структурное сходство этих белков (Moreno, Clemente, 2008). У большинства растений 2S альбумины состоят из двух полипептидных цепей – малой ( $M_r \sim 3000-5000$ ) и большой ( $M_r \sim 7000-9000$ ). Субъединицы синтезируются в виде одного предшественника (Shewry et al. 1995), который затем расщепляется, образуя две субъединицы.

Запасные 2S альбумины семян относят к семейству ингибиторов трипсина/альфа-амилазы (Shewry, Tatham, 1999), насчитывающему 33% пищевых аллергенов. Эти белки эволюционно родственны, содержат консервативный «скелет» из восьми остатков цистеина. Считается, что гены 2S альбуминов растений, за редкими исключениями (Gander et al., 1999) не содержат интронов.

Kreis et al. (1985) предположили эволюционное родство 2S альбуминов и «двухголового» ингибитора трипсина Bowman-Birk (BBI) из сои. Не всегда это находило подтверждение, но в пользу данного предположения говорит частичная гомология BBI и ингибитора трипсина из семян *Zinnia elegans* Jacq. (ZTI, 11350), представителя близкого подсолнечнику рода *Zinnia* L. (Konarev et al., 2002). ZTI также гомологичен 2S альбуминам подсолнечника. Представители семейства BBI чрезвычайно изменчивы у растений и, возможно, родственные им структуры путем дупликации и дивергенции участвовали в становлении разнообразных белков, в частности, у подсолнечника. Mylne et al. (2011) установили, что полипептиды-предшественники 2S альбуминов разных видов подсолнечника содержат с N-конца последовательности, которые в результате процессинга дают, наряду с альбуминами, пептиды, обладающие определенной биологической активностью. Одним из таких пептидов является найденный ранее уникальный низкомолекулярный (1514 Да) циклический ингибитор трипсина SFTI-1 (Luckett et al., 1999; Konarev et al., 2000, 2002), представляющий большой интерес для медицинской и аграрной биоинженерии. Важно отметить, что SFTI-1 практически полностью соответствует ингибиторной петле BBI.

В отличие от 2S альбуминов других растений, 2S альбумины подсолнечника имеют уникальную структуру: они существуют в виде единичных полипептидов, а их вторичная структура поддерживается за счет внутримолекулярных дисульфидных связей (Kortt, Caldwell, 1990, Anisimova et al., 1995; Pandya et al.,

2000). Фракция 2S альбуминов подсолнечника содержит два уникальных богатых серосодержащими аминокислотами белка – SFA7 и SFA8 ( $M_r \sim 12000$ ), характеризующиеся почти идентичным аминокислотным составом. Содержание остатков метионина в этих полипептидах составляет 15 моль% (Kortt, Caldwell, 1990). Белок SFA8 является главным компонентом альбуминовой фракции. Молекулярная масса белка составляет 12,133 кД, а длина полипептидной цепи – 103 аминокислотных остатка, из которых 16 составляет метионин и 8 – цистеин (Kortt et al., 1991). В генофонде подсолнечника обнаружены два аллельных варианта белка SFA8 – SFA8<sub>n</sub> нормальный («дикого типа») и «вариантный» (SFA8<sub>v</sub>), которые незначительно различались по подвижности в полиакриламидном геле (SDS-трис-трициновая система, pH 8,8), а при изоэлектрофокусировании имели разные значения изоэлектрических точек – 6,0 у белка «дикого типа» и 6,5 – у вариантного (Anisimova et al., 2003). При этом у всех генотипов гомозиготных по аллелю SFA8<sub>v</sub> в спектрах присутствовала слабая полоса, соответствующая белку SFA8<sub>n</sub>. Причины выявленного полиморфизма пока неизвестны.

Последовательность первичного продукта трансляции SFA8 (600 пн) впервые была определена в работе (Kortt et al., 1991) и помещена в базу данных NCBI (<http://www.ncbi.nlm.nih.gov/nucore>) под номером X56686. В референсной последовательности генома *H. annuus* HanXRQr1.0 ген SFA8 локализован в группе сцепления 11 в промежутке 89,644,488–89,645,412 (локус LOC110892047), однако аннотированная последовательность XM\_022139425.1 также представлена предсказанной последовательностью мРНК длиной 623 пн.

В результате комплексного анализа генома, транскриптома и протеома установлено, что из 26 генов 2S альбуминов, обнаруженных в геноме подсолнечника, экспрессировались лишь 15. В белковых профилях зрелых семян были представлены продукты 11 генов, причем в протеоме семян доминировали всего три белка – SESA2, SESA20-2, SESA3 (SFA8), кодируемые тремя генами (Franke et al., 2016; Jayasena et al., 2016). Гены этих доминирующих белков отличались от генов с низким уровнем экспрессии наличием интронов, что свидетельствовало о возможном влиянии интронов на экспрессию (Jayasena et al., 2016). Полиморфизм 2S альбуминов подсолнечника авторы связывают преимущественно с посттрансляционными модификациями. Работы по биохимической характеристике и анализу экспрессии этих белков были выполнены на единичных генотипах. Данные, иллюстрирующие полиморфизм альбуминовой фракции семян подсолнечника, весьма ограничены и основаны исключительно на результатах одномерного и двумерного электрофореза в полиакриламидном геле, изоэлектрофокусирования, а также RP-HPLC (Anisimova et al., 1995, 2003). Однако, учитывая уникальность альбумина SFA8, его биологическую ценность как источника незаменимой аминокислоты метионина, а также ценные функциональные свойства и потенциальную аллергенность, необходимо выяснение молекулярных основ изменчивости этого белка. В задачи настоящего исследования входило: определение экзон-интронной структуры гена SFA8, а также анализ изменчивости геномных последовательностей у различных генотипов культурного подсолнечника и образцов диких видов рода *Helianthus* L.

## Материалы и методы

Материал исследования включал 6 линий *H. annuus* L. из генетической коллекции ВИР, несущих, по данным гибридологического анализа, различные аллели в локусе SFA8 (Anisimova et al., 2003): ВИР130 и ВИР365 с аллелем SFA8<sub>v</sub>,

ВИР262, ВИР114А (ЦМС РЕТ1), ВИР471Ар (ЦМС РЕТ1), ВИР471Аrig (ЦМС RIG0) с аллелем дикого типа *SFA8n*, а также образцы сорта 'Передовик' и трех диких видов – однолетних *H. argophyllus* Torr. & A.Gray и *H. petiolaris* Nutt. и многолетнего *H. giganteus* L. Суммарную ДНК выделяли из этиолированных проростков с использованием модифицированного СТАВ-метода (Анисимова и др., 2010).

Праймеры для амплификации геномной последовательности SFA8 (прямой 5'-ТСТАСААТGGCAAGGTTTTC, обратный 5'-ТААСGAGACACACGATGGAC-3'), конструировали на основе информации о последовательности мРНК SFA8 (образец X56686, депонированной в международном генном банке GenBank NCBI (<http://www.ncbi.nlm.nih.gov/nucscore>) с использованием программы Primer3Plus (<http://www.bioinformatics.nl/>). Продукт амплификации включал кодирующую последовательность, 6 нуклеотидов 5' лидерной последовательности и 120 нуклеотидов 3' трейлерной последовательности. Его предсказанная длина составляла около 550 пн. Полимеразную цепную реакцию проводили в объеме 25 мкл по протоколам, подобранным нами. Реакционная смесь (25 мкл) содержала 25 нг геномной ДНК, 1,5 мМ MgCl<sub>2</sub>, по 0,1 мМ каждого из dNTP, по 0,5 мМ праймера и 1 е.а. Taq-полимеразы. Реакцию проводили при следующих условиях: 36 циклов, денатурация при 94°C (30 с), отжиг при 55°C (45 с), элонгация при 72°C (1 мин), финальная элонгация – 10 мин.

Очистку ПЦР-продуктов из амплификационной смеси проводили в 1% агарозном геле. Амплифицированные фрагменты были клонированы в векторе pAL-TA Vector по протоколу производителя (компания Евrogen, <https://www.evrogen.ru>). Для трансформации использовали клетки *E. coli* штамма DH5L. Отбор рекомбинантных клонов проводили при помощи ПЦР с универсальными праймерами M13dir и M13rev. ПЦР-продукты разделяли электрофорезом в 1% агарозном геле и отбирали клоны со вставкой нужного размера. Фрагменты были секвенированы с использованием генетического анализатора ABI PRISM 3500xl в ЦКП «Геномные технологии и клеточная биология» (ФГБНУ Всероссийский научно-исследовательский институт сельскохозяйственной микробиологии). Во избежание ошибок при амплификации и чтении последовательностей секвенирование проводили в двух повторностях и каждый фрагмент секвенировали с обоих концов. Число отсеквенированных клонов было различным для изученных генотипов. Поскольку наибольший интерес представляло изучение связи между генетическим полиморфизмом белка SFA8 и полиморфизмом его кодирующих последовательностей, максимальное число клонов (10) было получено для линии ВИР130 с вариантным белком SFAv. Для остальных генотипов анализировали по 3–4 клон.

Выравнивание последовательностей и их анализ проводили с помощью программы MEGA version 5.1 (Kumar et al., 2016). Для расчета биохимических и физико-химических показателей предсказанных белков использовалась программа <http://cn.exPASy.org/tools/#proteome>. Вторичная структура интрона предсказывалась с использованием программы MFOLD (Zuker, 2003).

## Результаты и обсуждение

Продукты ПЦР представителей однолетнего культурного вида *H. annuus* и наиболее близкого к нему дикорастущего однолетнего *H. argophyllus*, полученные с помощью праймеров, фланкирующих кодирующую последовательность гена *SFA8*, имели одинаковую длину размера – примерно 850 пн (рис. 1). В то же время длина продуктов ПЦР образцов многолетнего *H. giganteus* и однолетнего

*H. petiolaris* оказалась примерно на 30–40 пн меньше и составила около 790 пн. Размер полученных ампликонов значительно отличался от ожидаемого (550 пн). Отсюда мы предположили, что геномная последовательность *SFA8* содержала вставку размером около 300 пн.

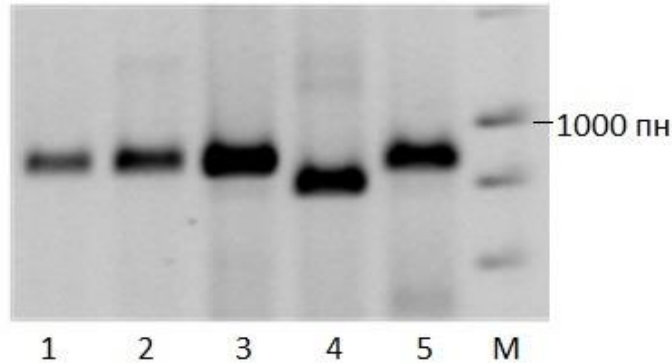


Рис. 1. Продукты ПЦР, полученные с использованием пары праймеров *SFA8* у линий *H. annuus*: ВИР130 (1), ВИР365 (2), ВИР262 (3), ВИР114А (ЦМС PET1) (5) и многолетнего вида *H. giganteus* (4). М – маркер молекулярного веса 100bp DNA Ladder (Fermentas)

Fig. 1. PCR products obtained using primer pair *SFA8* in *H. annuus* lines VIR130 (1), VIR365 (2), VIR262 (3), VIR114A (CMS PET1) (5) and perennial species *H. giganteus* (4). M – molecular weight marker 100bp DNA Ladder (Fermentas)

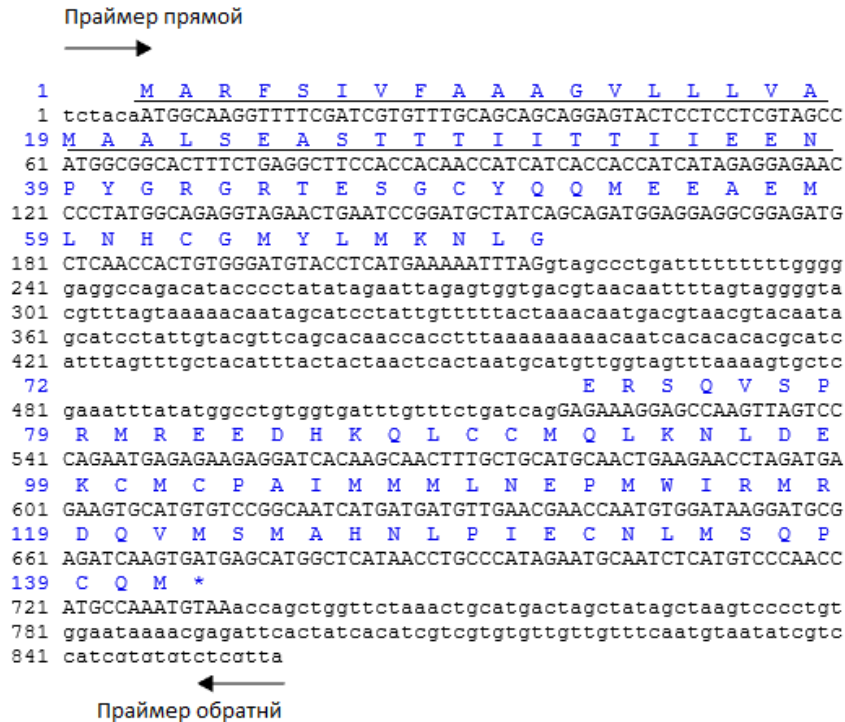
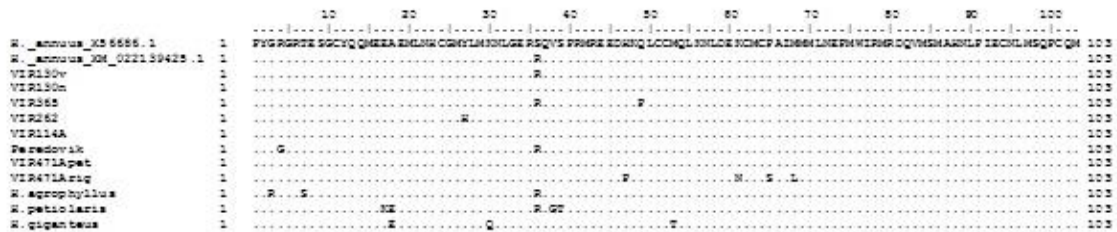


Рис. 2. Нуклеотидная и транслированная *in silico* аминокислотная последовательности гена *SFA8* (линия ВИР114А). Последовательности экзонов указаны прописными буквами, последовательность интрона и 5' и 3' нетранслируемых районов – строчными. Предполагаемые аминокислотные остатки приведены над нуклеотидной последовательностью

Fig. 2. Nucleotide and *in silico* translated amino acid sequences of the *SFA8* gene (line VIR 114A). Exon sequences are indicated by uppercase letters, sequences of the intron and untranslated regions are shown by lowercase letters. Predicted amino acids are written above the nucleotide sequence

Полученные после амплификации фрагменты гена *SFA8* были клонированы и секвенированы. Сравнительный анализ нуклеотидных последовательностей изученных генотипов и двух депонированных в базе NCBI последовательностей мРНК (X56686 и XM\_022139425.1) показал, что ген *SFA8* имеет интрон протяженностью 250–300 нуклеотидов (рис. 2). Длина первого экзона составляет 99 пн, длина второго – 210 пн. Интрон находится в центральной части гена, вблизи одного из двух эволюционно консервативных сайтов посттрансляционного расщепления на предполагаемые «малую» и «большую» субъединицы. В последовательности *SFA8* обнаружены два таких сайта – SQ (позиции 36, 37) и SP (позиции 39, 40) аминокислотной цепи зрелого белка (Kortt et al., 1991). Эти сайты идентичны у предшественников *SFA8* и 2S альбумина семян бразильского ореха, однако, в отличие от гомологичных белков других двудольных растений, у подсолнечника данная область процессингу не подвергается. Отсутствие процессинга можно объяснить отсутствием соответствующей протеазы в семенах либо структурными отличиями данного района, последовательность которого у подсолнечника, как и у схожего по структуре одноцепочечного ингибитора трипсина ячменя, содержит меньше аминокислотных остатков (Kortt et al., 1991).

Последовательность интрона в референсных последовательностях X56686 и XM\_022139425.1 из базы данных NCBI отсутствовала и идентифицирована нами впервые. Она оказалась полиморфной и характеризовалась наличием инделей длиной около 45 пн. при сравнении диких видов *H. argophyllus*, *H. petiolaris* и *H. giganteus* с однолетним культурным подсолнечником *H. annuus*.



**Рис. 3. Выравнивание предсказанных аминокислотных последовательностей белка SFA8 из различных генотипов (позиции аминокислот 1–103)**

**Fig. 3. Alignment of predicted amino acid sequences of the SFA protein from various genotypes (amino acid (positions 1–103))**

В результате сравнительного анализа нуклеотидных последовательностей и последовательности мРНК *SFA8* (сорт ‘Hysun’) (Kortt et al., 1991) у изученных генотипов выявлены замены отдельных нуклеотидов в районах как в 5', так и в 3' нетранслируемых последовательностях. Транслированные *in silico* аминокислотные последовательности характеризовались полиморфизмом у различных генотипов (рис. 3). Отдельные аминокислотные замены обнаружены у сорта ‘Передовик’ и линии VIP262, дикие виды отличались от культурного подсолнечника по большему числу позиций. Секвенированные последовательности одного и того же генотипа – линии VIP130 (характеризуется двумя электрофоретическими вариантами белка *SFA8*, различающимися по изоэлектрической точке) – оказались полиморфными. Большая часть клонов (восемь из десяти проанализированных) содержала замену 108С—G, приводящую

к замене полярной незаряженной аминокислоты серина на положительно заряженный аргинин (позиция 36 полипептидной цепи), но в одной из клонированных последовательностей подобная замена не выявлена. Полученные данные согласуются с полученными нами ранее результатами биохимического анализа расщепляющейся гибридной популяции  $F_2$  (ВИР130 × ВИР104) (Anisimova et al., 2003). В спектрах изоэлектрофокусирования 2S альбуминов, выделенных из семян генотипов гомозиготных по вариантному аллелю *SFA8v*, присутствовали оба компонента (*SFA8v* и *SFA8n*), причем интенсивность компонента дикого типа была заметно ниже интенсивности вариантного белка. Этот факт можно объяснить наличием в генотипе нескольких отличающихся по последовательности копий гена *SFA8* либо наличием всего двух копий – дикого типа и вариантной с мутацией 108С—G, но отличающихся по уровню экспрессии. Такая же замена обнаружена в проанализированных последовательностях из других генотипов, у которых ранее (Anisimova et al., 2003) мы обнаружили вариантный белок *SFA8v* – у линий ВИР365 и сорта ‘Передовик’. Замена 108С—G также обнаружена в изученных последовательностях *SFA8* однолетних диких видов *H. argophyllus* и *H. petiolaris*. Следует отметить, что та же замена присутствует и в референсной последовательности XM\_022139425.1 из аннотированного генома подсолнечника, источником которого служила инбредная линия XRQ, созданная в Национальном институте сельскохозяйственных исследований Франции (INRA) на основе российского сорта ‘Прогресс’ (Badouin et al., 2017). Интересно, что и линия ВИР365, у которой обнаружена мутация 108С—G, также создана при участии сорта ‘Прогресс’.

**Таблица. Физико-химические и биохимические характеристики вариантов белка *SFA8*, рассчитанные на основании их кодирующих последовательностей с помощью программы EXPASY (<http://cn.expasy.org/tools/#proteome>)**  
**Table. Physicochemical and biochemical characteristics of *SFA8* protein variants calculated from their coding sequences using EXPASY software (<http://cn.expasy.org/tools/#proteome>)**

Образец Accession	pI	Заряд молекулы при pH 7,80 Molecule charge at pH 7.80	Показатель гидрофобнос- ти (GRAVY) Hydrophobicity index (GRAVY)	Позиции сайтов гидролиза трипсином (общее число) Positions and number of trypsin cleavage sites
SFA8 X56686.1	5,91	-1,715	-0,586	4 6 30 35 41 43 48 56 61 78 80 (11)
SFA8 XM022139425.1	6,35	-0,715	-0,622	4 6 30 35 36 41 43 48 56 61 78 80 (12)
ВИР130 ( <i>SFA8v</i> )*	6,35	-0,715	-0,622	4 6 30 35 36 41 43 48 56 61 78 80 (12)
ВИР130 ( <i>SFA8n</i> )	5,91	-1,715	-0,586	4 6 30 35 41 43 48 56 61 78 80 (11)
ВИР365 ( <i>SFA8v</i> )	6,35	-0,715	-0,604	4 6 30 35 36 41 43 56 61 78 80 (11)
ВИР262 ( <i>SFA8n</i> )	6,09	-1,623	-0,605	4 6 30 35 41 43 48 56 61 78 80 (11)
ВИР114А ( <i>SFA8n</i> )	5,91	-1,715	-0,586	4 6 30 35 41 43 48 56 61 78 80 (11)
Передовик ( <i>SFA8n</i> )	5,91	-1,715	-0,583	6 30 35 36 41 43 48 56 61 78 80 (11)
ВИР471Ар ( <i>SFA8n</i> )	5,91	-1,715	-0,586	4 6 30 35 41 43 48 56 61 78 80 (11)
VIR471Arig ( <i>SFA8n</i> )	5,32	-2,805	-0,541	6 30 35 41 43 48 56 78 80 (10)
<i>H. argophyllus</i>	7,19	+0,285	-0,663	3 4 6 30 35 36 41 43 48 56 61 78 80 (13)
<i>H. petiolaris</i>	7,19	+0,285	-0,721	4 6 17 30 35 36 41 43 48 56 61 78 80 (13)
<i>H. giganteus</i>	5,30	-3,713	-0,659	4 6 35 41 43 48 56 61 78 80 (10)

\*в скобках отмечен электрофоретический вариант полипептида (*SFA8n* – нормальный, *SFA8v* – мутантный).



Аминокислотный состав белковых молекул определяет значение изоэлектрической точки. Расчетное значение изоэлектрической точки полипептидов SFA8, у которых выявлена замена S36R составило 6,35, тогда как значение изоэлектрической точки вариантов SFA8 «дикого типа» – 5,91 (таблица). Эти значения соответствуют оценкам 6,5 и 6,0, полученным ранее при изофокусировании белков (Anisimova et al., 2003). Можно предположить, что изменение изоэлектрической точки сопровождается изменением заряда в щелочной среде, что также отражается на его подвижности в полиакриламидном геле. Рассчитанные изоэлектрические точки белка SFA8 у различных линий *H. annuus* колебались от 5,32 до 6,35, заряд – от –2,805 до –0,715, а показатель гидрофильности/гидрофобности (GRAVY) – от –0,541 до –0,622 (таблица). Число и позиции сайтов расщепления пищеварительным ферментом трипсином было одинаковым для белков из разных генотипов *H. annuus*, за исключением вариантного SFA8 полипептида, но отличалось у диких видов. Состав пептидов, образующихся в результате расщепления пептидных связей в белке, может определять его аллергенные свойства. Пептиды, образующиеся при гидролизе полиморфных белков, могут отличаться по составу, что может приводить к изменению потенциальных аллергенных свойств белка.

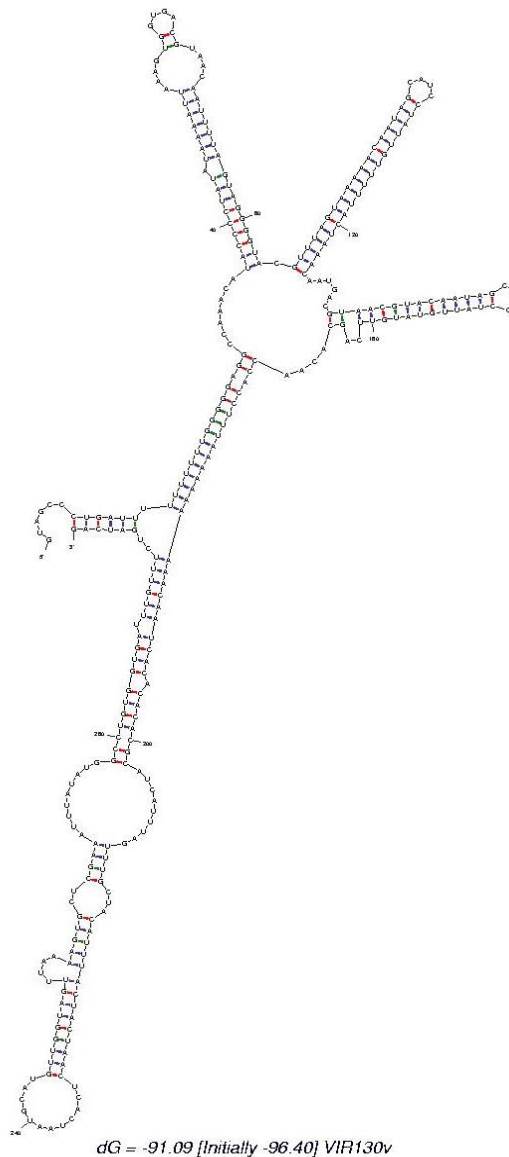
Нуклеотидные последовательности интрона гена SFA8 различались у исследованных образцов. Длина интрона составляла от 258 пн у *H. petiolaris* и *H. giganteus* до 303 у последовательности VIR130v линии ВИР130 *H. annuus*. У всех образцов последовательности интрона содержали динуклеотиды GT на 5', а также AG на 3' конце интрона, характерные для консенсусных последовательностей сайтов сплайсинга интронов U2-типа, которые являются основным типом ядерных интронов эукариот (рис. 4, 5).

Вторичная структура интронов пре-мРНК гена SFA8 у различных образцов подсолнечника существенно отличались. Для образца *H. argophyllus* и всех проанализированных образцов *H. annuus* (линии ВИР130, ВИР365, ВИР262, ВИР114А, ВИР471Apet, ВИР471Arig, сорт 'Передовик'), последовательности интронов которых имеют сходную длину и отличаются короткими инделями и единичными нуклеотидными заменами, было выявлено от одного (для последовательности VIR130v линии ВИР130) до трех (последовательности сорта 'Передовик' и линии ВИР262) вариантов вторичной структуры. При этом обладающие минимальной энергией ( $\Delta G$  от –92 до –87) оптимальные варианты вторичной структуры последовательностей интрона образца вида *H. argophyllus* и всех проанализированных образцов *H. annuus* были сходны между собой. 5' и 3' концы интрона были сближены за счет формирования небольшой шпильки I и двух более протяженных соседних шпилек II и III (рис. 4).

Последовательности интрона гена SFA8 у видов *H. petiolaris* и *H. giganteus* отличались от последовательностей *H. argophyllus* и *H. annuus* наличием делеции, длина которой составляла 45 пн у *H. petiolaris* и 46 пн у *H. giganteus*. Для последовательности *H. petiolaris* было выявлено пять возможных вариантов вторичной структуры с  $\Delta G$  от –59 до –61. Для последовательности *H. giganteus* было выявлено восемь возможных вариантов вторичной структуры с  $\Delta G$  от –47 до –55. При этом 5' и 3' концы интрона последовательности *H. petiolaris* не формировали шпильку ни в одном из вариантов вторичной структуры.

У *H. giganteus*, несмотря на наличие делеции в последовательности интрона, семь из восьми вариантов вторичной структуры были сходны со вторичной структурой последовательностей *H. argophyllus* и *H. annuus*. Как и у *H. argophyllus* и *H. annuus*, последовательность интрона *H. giganteus* формировала три шпильки, в результате чего 5' и 3' концы интрона также были сближены (рис. 5). Однако,

структура шпилек интрона *H. giganteus* отличалась от таковой *H. argophyllus* и *H. annuus*. Таким образом, были выявлены отличия во вторичной структуре интрона гена *SFA8* у образцов видов *Helianthus*.

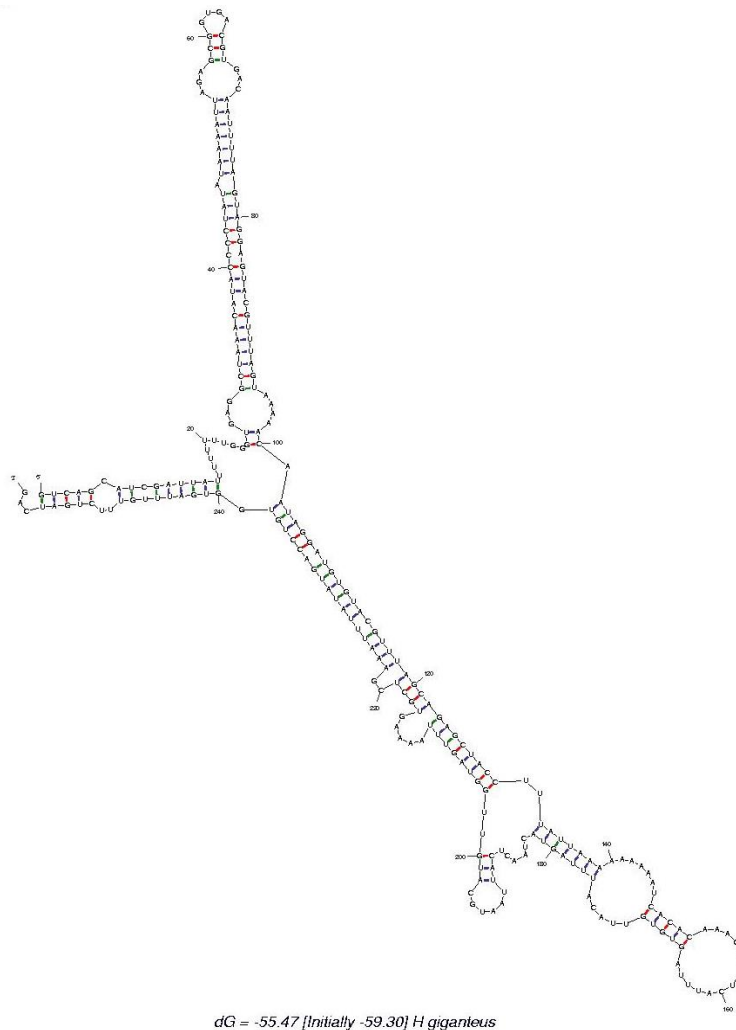


**Рис. 4. Вторичная структура интрона гена *SFA8* последовательности VIR130v линии ВИР130 *Helianthus annuus* L.**

**Fig. 4. Secondary structure of intron in the VIR130v sequence from *Helianthus annuus* L. line VIR130**

Ранее с помощью двумерного электрофореза альбуминовой фракции 2S альбуминов подсолнечника, продемонстрированы значительные межвидовые отличия по экспрессии *SFA8*. Компонент, соответствующий белку *SFA8*, был значительно ослаблен у образцов однолетних диких видов и практически не виден – у образцов многолетних (Gavrilova, Anisimova, 2003). В последнее время появились данные о влиянии вторичной структуры интрона на сплайсинг у некоторых видов животных, дрожжей и растений. На настоящий момент выявлено несколько механизмов влияния вторичной структуры интрона на сплайсинг и показано, что формирование вторичной структуры может как повышать, так и понижать эффективность сплайсинга (Lin et al., 2016). Кроме того,

у *Saccharomyces cerevisiae* формирование шпилек между сайтом ветвления и 3'-сайтом сплайсинга может играть роль термосенсора, влияя на альтернативный сплайсинг (Meyer et al., 2011). Эффективность экспрессии SFA8 у трансгенных растений люцерны, в генотип которых был включен кодирующий район гена, оказалась низкой (Tabe et al., 1993). Впоследствии было показано, что наличие интрона усиливает экспрессию SFA8 в трансгенных растениях *Arabidopsis* (Sun et al., 2015), что свидетельствовало о возможном влиянии интрона на экспрессию.



**Рис. 5. Вторичная структура вариантной последовательности интрона гена SFA8 *Helianthus giganteus* L.**  
**Figure 5. Secondary structure of the SFA8 gene intron in the VIR130v sequence from *Helianthus giganteus* L.**

### Заключение

В настоящей работе впервые изучен структурный полиморфизм последовательностей гена богатого метионином альбумина SFA8. Установлено, что ген SFA8 имеет два экзона длиной 99 пн и 210 пн и содержит в центральной части интрон, длина которого у изученных образцов представителей рода *Helianthus* варьирует в пределах 285–303 пн, а у образцов культурного вида

*H. annuus* составляет 300–303 пн. Варианты интрона полиморфны у разных генотипов и отличаются по вторичной структуре. У линий подсолнечника, характеризующихся различными электрофоретическими вариантами SFA8, выявлены различия кодирующих последовательностей и обнаружена мутация, приводящая к скачкообразному изменению заряда. Выявленные нами различия вторичной структуры интрона могут оказаться важными для понимания особенностей экспрессии гена SFA8, а обнаруженный полиморфизм кодирующих последовательностей – для выяснения молекулярных механизмов различий биохимических и функциональных свойств этого белка у генотипов подсолнечника.

*Благодарности:* Работа выполнена в рамках государственного задания согласно тематическому плану ВИР по теме № 0662-2018-0009 «Развитие современных технологий паспортизации сортов и гибридов, генотипирования мировых растительных ресурсов и их фитосанитарного мониторинга с использованием молекулярных маркеров», номер государственной регистрации ЕГИСУ НИОКР АААА-А16-116040710356-4 и согласно тематическому плану Института общей генетики им. Н.И. Вавилова по теме 0112-2018-0004 «Исследование генома культурных растений и их сородичей, применительно к генетической теории селекции», номер государственной регистрации ЕГИСУ НИОКР АААА-А16-116111610169-8.

## References/Литература

- Anisimova I. N., Alpat`eva N. V., Timofeeva G. I. Skrining geniticheskikh resursov rastenii s ispolzovaniem DNK markerov: osnovnye principy, vydelenie DNK, PCR, elektroforez v agaroznom gele. Metodicheskie ukazaniya VIR (Ed. by E.E. Radchenko). St. Petersburg : VIR, 2010, 30 p. [in Russian] (Анисимова И. Н., Алпатьева Н. В., Тимофеева Г. И. Скрининг генетических ресурсов растений с использованием ДНК-маркеров: основные принципы, выделение ДНК, постановка ПЦР, электрофорез в агарозном геле: Методические указания ВИР / под ред. Е. Е. Радченко. СПб. : ВИР, 2010. 30 с.).
- Gavrilova V. A., Anisimova I. N. Genetics of cultivated plants. Sunflower. St. Petersburg : VIR, 2003, 197 p. [in Russian] (Гаврилова В. А., Анисимова И. Н. Генетика культурных растений. Подсолнечник Санкт-Петербург : ВИР, 2003. 197 с.).
- Agizzio A. P., Da Cunha M., Carvalho A. O., Oliveira M. A., Ribeiro S. F., Gomes V. M. The antifungal properties of a 2S albumin-homologous protein from passion fruit seeds involve plasma membrane permeabilization and ultrastructural alterations in yeast cells // Plant Sci., 2006, vol. 171, no. 4, pp. 515–522. DOI: 10.1016/j.plantsci.2006.06.001.
- Anisimova I. N., Fido R. J., Tatham A. S., Shewry P. R. Genotypic variation and polymorphism of 2S albumins of sunflower // Euphytica, 1995, vol. 83, pp. 15–23.
- Anisimova I. N., Konarev A. V., Gavrilova V. A., Rozhkova V. T., Fido R. J., Tatham A. S., Shewry P. R. Polymorphism and inheritance of methionine-rich 2S albumins in sunflower // Euphytica, 2003, vol. 129, no. 1, pp. 99–107. DOI: 10.1023/A:1021562712945.
- Badouin H., Gouzy G., Grassa C. J. et al. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution // Nature, 2017, vol. 546, pp. 148–152. DOI: 10.1038/nature22380.
- Elliott A. G., Delay C., Liu H., Phua Z., Rosengren K. J., Benfield A. H., Panero J. L., Colgrave M. L., Jayasena A. S., Dunse K. M., Anderson M. A. Evolutionary origins of a bioactive peptide buried within preproalbumin // The Plant Cell, 2014, vol. 26, pp. 981–995. DOI 10.1105/tpc.114.123620.
- Franke B., Colgrave M. L., Mylne J. S., Rosengren K. J. Mature forms of the major seed storage albumins in sunflower: A mass spectrometric approach // 2016, vol. 147, no. 1, pp.177–186. DOI: 10.1016/j.jprot.2016.05.004.
- Jayasena A. S., Franke B., Rosengren J., Mylne J. S. A tripartite approach identifies the major sunflower seed albumins // Theor. Appl. Genet., 2016, vol. 129, no. 3, pp. 613–629. DOI: 10.1007/s00122-015-2653-3.
- Kortt A. A., Caldwell J. B. Low molecular weight albumins from sunflower seed: identification of a methionine-rich albumin // Phytochemistry, 1990, vol. 29, no. 9, pp. 2805–2810.

- Kortt A. A., Caldwell J. B., Lilley G. G., Higgins T. J. V. Amino acid and cDNA sequences of a methionine-rich 2S protein from sunflower seed (*Helianthus annuus* L.) // Eur. J. Biochem., 1991, vol. 195, pp. 329–334. DOI: 10.1111/j.1432-1033.1991.tb15710.x.
- Konarev A. V., Gavrilova V. A., Rozhkova V. T., Fido R. J., Tatham A. S., Shewry P. R. Novel proteinase inhibitors in seeds of sunflower (*Helianthus annuus* L.): polymorphism, inheritance and properties. Theor. Appl. Genet., 2000, vol. 100, no. 1, pp. 82–88. DOI: 10.1007/s001220050012.
- Konarev A. V., Anisimova I. N., Gavrilova V. A., Vachrusheva T. E., Konechnaya G. Y., Lewis M., Shewry P. R. Serine proteinase inhibitors in the Compositae: Distribution, polymorphism and properties // Phytochemistry, 2002, vol. 59, pp. 279–291. DOI: 10.1016/S0031-9422(01)00463-0.
- Kreis M., Forde B. G., Rahman S., Mifflin B. J., Shewry P. R. Molecular evolution of the seed storage proteins of barley, rye and wheat // J. Mol. Biol., 1985, vol. 183, no. 3, pp. 499–502.
- Lin C. L., Taggart A. J., Fairbrother W. G. RNA structure in splicing: An evolutionary perspective // RNA Biol., 2016, vol. 13, no. 9, pp. 766–771. DOI: 10.1080/15476286.2016.1208893.
- Luckett S., Garcia R. S., Barker J. J., Konarev A. V., Shewry P. R., Clarke A. R., Brady R. L. High resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds // J. Mol. Biol., 1999, vol. 290, pp. 525–533. DOI: 10.1006/jmbi.1999.2891.
- Meyer M., Plass M., Pérez-Valle J., Eyraes E., Vilardell J. Deciphering 3' splice site selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing // Mol. Cell., 2011, vol. 43, no. 6, pp. 1033–1039. DOI: 10.1016/j.molcel.2011.07.030.
- Moreno F. J., Clemente A. 2S Albumin Storage Proteins: What makes them food allergens? // Open Biochem. J., 2008, vol. 2, pp. 16–28. DOI: 10.2174/1874091X00802010016.
- Mylne J. S., Colgrave M. L., Daly N. L., Chanson A. H., Elliott A. G., Mc Callum E. J., Jones A., Craik D. J. Albumins and their processing machinery are hijacked for cyclic peptides in sunflower // Nat. Chem. Biol., 2011, vol. 7, pp. 257–259. DOI: 10.1038/nCHEMBIO.542.
- Odintsova T. I., Rogozhin E. A., Sklyar I. V., Musolyamov A. K., Kudryavtsev A. M., Pukhalsky V. A., Smirnov A. N., Grishin E. V., Egorov T. A. Antifungal activity of storage 2S albumins from seeds of the invasive weed dandelion *Taraxacum officinale* Wigg. // Protein and Peptide Letters, 2010, vol. 17, no. 4, pp. 522–529. DOI: 10.2174/092986610790963591.
- Pandya M. J., Sessions R. B., Williams P. B., Dempsey C. E., Tatham A., Shewry P. R., Clarke A. R. Structural characterization of a methionine-rich, emulsifying protein from sunflower seed // Proteins: Structure Function, and Bioinformatics, 2000, vol. 38, no. 3, pp. 341–349. DOI: 10.1002/(SICI)1097-0134(20000215)38:33.0.CO;2-D.
- Radauer C., Breiteneder H. Evolutionary biology of plant food allergens // J. Allergy Clin. Immunol. 2007, vol. 120, no. 3, pp. 518–525. DOI: 10.1016/j.jaci.2007.07.024.
- Shewry P. R., Napier J. A., Tatham A. S. Seed Storage Proteins: Structures and Biosynthesis // The Plant Cell, 1995, vol. 7, no. 7, pp. 945–956. DOI: 10.1105/tpc.7.7.945.
- Shewry P. R., Tatham A. S. The characteristics, structures and evolutionary relationships of prolamins. In : Seed proteins (eds. P. R. Shewry, R. Casey). Kluwer Academic Publishers, the Netherlands, pp. 11–33.
- Sun J., Zhao H., Nie L., Yi J., Zhang Q.-L. The intron in an albumin gene from sunflower increases expression of SFA8. In: Molecular breeding of forage and turf (eds. H. Budak, G. Spangenberg) 2015, pp. 183–191. DOI: 10.1007/978-3-319-08714-6\_16.
- Tabb L. M., Higgins C. M., McNabb W. C., Higgins T. J. Genetic engineering of grain and pasture legumes for improved nutritive value // Genetica, 1993, vol. 90, no. 2–3, pp. 181–200.
- Zuker M. Mfold web server for nucleic acid folding and hybridization prediction // Nucleic Acids Res., 2003, vol. 31, no. 13, pp. 3406–3415. DOI: 10.1093/nar/gkg595.
- Youle R. J., Huang A. H. C. Occurrence of low molecular weight and high cysteine containing albumin storage proteins in oilseeds of diverse species // Am. J. Bot., 1981, vol. 68, no. 1, pp. 44–48. DOI: 10.2307/2442990.