

Полиморфизм интрона гена *rps16* у представителей рода *Malus* Mill. и родственных видов семейства Rosaceae Juss.

М.А. Филюшин¹✉, К.В. Борис^{1, 2}

¹ Институт биоинженерии, Федеральный исследовательский центр «Фундаментальные основы биотехнологии» Российской академии наук, Москва, Россия

² Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия

Интрон хлоропластного гена *rps16*, кодирующего рибосомный белок 16S, относится к интронам группы II, для которых характерна способность к автономному сплайсингу. Последовательность этого интрона может быть достаточно полиморфной и широко применяется для молекулярно-генетических исследований у различных семейств растений. Однако участки интрона, функционально значимые для сплайсинга и формирования вторичной структуры пре-мРНК у представителей семейства Rosaceae, в том числе рода *Malus*, на сегодняшний день не описаны. Целью работы была характеристика нуклеотидных последовательностей интрона *rps16*, оценка их полиморфизма и идентификация отдельных структурных элементов интрона, определяющих правильный сплайсинг гена *rps16* у представителей рода *Malus* и родственных видов Rosaceae. Проведены амплификация, секвенирование и анализ полиморфизма нуклеотидных последовательностей интрона группы II хлоропластного гена *rps16* у 32 видов рода *Malus* и 17 видов из других родов семейства Rosaceae. Показано, что у представителей рода *Malus* данный интрон имеет крайне низкий уровень вариативности (1.5 %), так же как и у представителей рода *Prunus*, у которых полиморфизм интрона гена *rps16* был несколько выше (2.25 %). В изученных последовательностях интрона гена *rps16* были идентифицированы границы шести функциональных доменов, характерные для интронов группы II, а также сайты связывания IBS/EBS. Определены уровни вариативности последовательностей всех шести доменов и междоменных участков. Наиболее консервативными оказались домены V и VI, что может быть связано с ролью этих доменов в позиционировании всех структур пре-мРНК и формировании каталитически активной пространственной структуры интрона. Таким образом, впервые изучен полиморфизм интрона хлоропластного гена *rps16* у видов *Malus* и родственных видов Rosaceae. Идентифицированы границы всех шести функциональных доменов, характерные для интронов группы II, и сайты связывания IBS/EBS. Показано, что наиболее консервативными в последовательности интрона являются домены V и VI.

Ключевые слова: интроны группы II; *Malus*; Rosaceae; вторичная структура пре-мРНК; генетический полиморфизм.

Polymorphism of the *rps16* gene intron in *Malus* Mill. and related Rosaceae Juss. species

M.A. Filyushin¹✉, K.V. Boris^{1, 2}

¹ Institute of Bioengineering, Research Center of Biotechnology RAS, Moscow, Russia

² Vavilov Institute of General Genetics RAS, Moscow, Russia

The intron of the chloroplast *rps16* gene encoding ribosomal protein 16S belongs to group II introns, which have the capacity to self-splice. The sequence of this intron may be quite polymorphic and is widely used for molecular-genetic studies of different plant families. But regions of the intron important for splicing and organization of the pre-mRNA secondary structure in Rosaceae species including the genus *Malus* have not yet been described. The aim of the work was to characterize the nucleotide sequences of the *rps16* gene intron, to study their polymorphism and to identify the individual structural elements of the intron determining the correct splicing of the *rps16* gene in *Malus* species and related Rosaceae species. Nucleotide sequences of the *rps16* chloroplast gene group II intron from 32 *Malus* species and 17 related Rosaceae species were amplified, sequenced and their polymorphism was analyzed. The *rps16* intron sequence in the genus *Malus* appeared to be very conservative (1.5 %), as well as in *Prunus* species, the level of the *rps16* intron polymorphism was slightly higher (2.25 %). In the studied sequences the boundaries of the six functional domains typical of group II introns and IBS/EBS binding sites were identified. Variability of all the domains and inter-domain regions was studied. Domains V and VI were the most conservative, which may be due to their role in the positioning of the pre-mRNA structures and formation of the catalytically active secondary structure. Thus, polymorphism of the *rps16* gene intron was for the first time studied in *Malus* species and related Rosaceae species. The boundaries of all six functional domains typical of group II introns and IBS/EBS binding sites were identified. The most conservative were domains V and VI.

Key words: group II introns; *Malus*; Rosaceae; pre-mRNA secondary structure; genetic polymorphism.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Филюшин М.А., Борис К.В. Полиморфизм интрона гена *rps16* у представителей рода *Malus* Mill. и родственных видов семейства Rosaceae Juss. Вавиловский журнал генетики и селекции. 2017;21(5):596-600. DOI 10.18699/VJ17.20-o

HOW TO CITE THIS ARTICLE:

Filyushin M.A., Boris K.V. Polymorphism of the *rps16* gene intron in *Malus* Mill. and related Rosaceae Juss. species. Vavilovskii Zhurnal Genetiki i Selektzii = Vavilov Journal of Genetics and Breeding. 2017;21(5):596-600. DOI 10.18699/VJ17.20-o (in Russian)

К интронам группы II относятся автономно сплайсируемые интроны, встречающиеся в бактериальных геномах, а также в цитоплазматических геномах растений (Carrillo et al., 2001; Brown et al., 2014). Точное вырезание интронов группы II в процессе сплайсинга достигается за счет формирования специфической консервативной вторичной структуры пре-мРНК, состоящей из шести доменов (Dai et al., 2008).

В митохондриальном и хлоропластном геномах высших растений насчитывается около 20 интронов группы II, но в процессе эволюции многие из них утратили способность к автономному сплайсингу, и их вырезание происходит аналогично ядерным интронам с помощью рибонуклеопротеидных комплексов (Brown et al., 2014; Zimmerly, Sempere, 2015). В пластидном геноме высших растений способность к автономному сплайсингу, по-видимому, сохранил только интрон гена *trnK* за счет наличия в его последовательности гена матуразы *matK* (Kelchner, 2002). Считается, что одного гена, кодирующего матуразу, достаточно для вырезания всех интронов группы II в хлоропластном геноме растений (Vogel et al., 1997), однако до сих пор не показано автономного вырезания ни одного интрона группы II у растений *in vitro* (Zimmerly, Sempere, 2015).

Интрон хлоропластного гена *rps16*, кодирующего рибосомный белок 16S, относится к интронам группы II. Интересно отметить, что у некоторых растений интрон в гене *rps16* может отсутствовать (Jansen et al., 2007), как и сам ген *rps16* (Lei et al., 2016).

В настоящее время последовательность и вторичная структура интрона гена *rps16* охарактеризованы у ряда однодольных и двудольных растений. Показано, что данный интрон может быть достаточно полиморфным у разных родов растений и содержать как единичные нуклеотидные замены, так и индели (Michel et al., 1989; Рыжова и др., 2009, 2013; Yuhua et al., 2011). Это обуславливает использование последовательностей интрона гена *rps16* для молекулярно-генетических исследований растений на различных таксономических уровнях (Smedmark et al., 2008; Рыжова и др., 2009, 2013).

У представителей Rosaceae интрон *rps16*, наряду с другими последовательностями хлоропластного и ядерного генома, ранее был использован для изучения филогении трибы Ругеае (Lo, Donoghue, 2012). Однако участки этого интрона, функционально значимые для сплайсинга и формирования вторичной структуры пре-мРНК у представителей семейства Rosaceae, в том числе рода *Malus*, на сегодняшний день не описаны. Целью данной работы была характеристика нуклеотидных последовательностей интрона *rps16*, оценка их полиморфизма и идентификация отдельных структурных элементов интрона, определяющих правильный сплайсинг гена *rps16* у представителей рода *Malus* и родственных видов Rosaceae.

Материалы и методы

Для анализа полиморфизма и вторичной структуры интрона гена *rps16* было отобрано 38 образцов 32 видов пяти основных секций рода *Malus*, а также 10 образцов видов других родов семейства Rosaceae (табл. 1). Для сравнительного анализа из базы данных NCBI были взяты

последовательности интрона гена *rps16* семи видов рода *Prunus* (см. табл. 1).

ДНК выделяли из свежих листьев и гербарного материала СТАВ методом с дополнительной очисткой хлороформом. Амплификацию последовательности интрона гена *rps16* проводили по стандартной методике с использованием реактивов производства «Диалат ЛТД» (Россия). Для амплификации и секвенирования полученных продуктов использовали праймеры *rps16F* (5'-AAA CGATGTGGTARAAAGCAAC-3') и *rps16R* (5'-AACATCW ATTGCAASGATTTCGATA-3'), соответствующие 3'-концу первого экзона и 5'-концу второго экзона гена *rps16* (Shaw et al., 2005).

Первичные нуклеотидные последовательности амплифицированных фрагментов определяли с помощью системы BigDye на секвенаторе Applied Biosystems 3730 DNA Analyzer (Applied Biosystems, США) с использованием соответствующих праймеров. Выравнивание и анализ полученных нуклеотидных последовательностей проводили в программе MEGA 7.0 (Kumar et al., 2016). В этой же программе определялось число вариабельных сайтов (нуклеотидные сайты выровненных последовательностей, в которых наблюдаются замены; синоним SNPs) и инсерций/делений (индели).

Результаты и обсуждение

Последовательность интрона гена *rps16* была амплифицирована и секвенирована у всех взятых в анализ образцов. У представителей рода *Malus* длина интрона варьировала от 859 п. н. у Антоновки обыкновенной и Антоновки Ольгинской до 887 п. н. у *M. soulandii*. У видов *Prunus* длина интрона менялась от 856 п. н. у *P. padus* до 871 п. н. у *P. persica*, а у видов из других родов Rosaceae этот показатель варьировал от 851 п. н. у *Chaenomeles japonica* до 889 п. н. у *Cydonia oblonga* (см. табл. 1). Выравненная длина интрона составила 932 п. н. Всего в последовательности интрона идентифицировано 98 вариабельных сайтов (10.5 % от выравненной длины интрона), из них для анализируемых видов *Malus* вариабельными были только 14 сайтов (1.5 %).

У всех анализируемых образцов на 5'-конце последовательности интрона выявлен мотив GTGCGA, соответствующий каноническому мотиву GUGYGA, участвующему в образовании междоменных связей при формировании вторичных структур, а на 3'-конце все последовательности имели мотив AAT, что соответствует каноническому мотиву RAY (Michel et al., 1989). Наличие этих мотивов позволяет отнести интрон гена *rps16* исследуемых образцов Rosaceae к типу IIB (Michel et al., 1989; Toor et al., 2001). Также у всех образцов в последовательности интрона были идентифицированы канонические полуконсервативные сайты связывания EBS1 (GTTGC) и EBS2 (TCTT), которые взаимодействуют при сплайсинге с сайтами связывания IBS1 (GCAAC) и IBS2 (AGAA), расположенными в первом экзоне гена *rps16* (Toor et al., 2001).

В последовательности интрона гена *rps16* у всех изучаемых образцов Rosaceae были определены границы шести доменов, характерных для интронов группы II (Michel et al., 1989) (табл. 2). Наиболее протяженным оказался домен I, длина которого у образцов *Malus* и

Table 1. Rosaceae accessions taken for chloroplast gene *rps16* intron polymorphism analysis

No	Species / variety	VIR accession number (if provided)	<i>rps16</i> intron length	No	Species / variety	VIR accession number (if provided)	<i>rps16</i> intron length
Section <i>Malus</i>				28	<i>M. sargentii</i>	2428	862
1	<i>M. caspiensis</i>	14943	860	29	<i>M. scheideckeri</i>	2407	862
2	<i>M. orientalis</i>	29484	864	30	<i>M. sieboldii</i>	2322	866
3	<i>M. prunifolia</i>	2430	865	31	<i>M. toringoides</i>	3109	866
4	<i>M. pumila</i>	2383	863	32	<i>M. transitoria</i>	2424	862
5	<i>M. purpurea</i>	2392	863	33	<i>M. zumi</i>	2427	866
6	<i>M. purpurea</i> var. <i>aldenhamensis</i>	2393	863	Section <i>Chloromeles</i>			
7	<i>M. ringo</i>	41280	863	34	<i>M. coronaria</i>	2336	885
8	<i>M. sieversii</i>	13280	863	35	<i>M. ioensis</i>	2352	886
9	<i>M. spectabilis</i>	2415	863	36	<i>M. platycarpa</i>	362179	884
10	<i>M. sylvestris</i>	14981A	864	37	<i>M. soulardii</i>	2414	887
11	<i>M. sylvestris</i>	41639	863	Section <i>Docyniopsis</i>			
12	<i>M. turkmenorum</i>	13283	862	38	<i>M. sikkimensis</i>	2412	863
<i>M. domestica</i> varieties				Other <i>Rosaceae</i>			
13	Антоновка обыкновенная	21190	859	39	<i>Aronia melanocarpa</i>		863
14	Антоновка Ольгинская	56	859	40	<i>Chaenomeles japonica</i>		851
15	Антоновка Полуторафунтовая	76	860	41	<i>Cotoneaster lucidus</i>		853
16	Антоновка из Севастопольской	8920	861	42	<i>Crataegus molis</i>		864
17	Golden Delicious	17477	862	43	<i>C. monogyna</i>		864
Section <i>Gymnomeles</i>				44	<i>C. submolis</i>		864
18	<i>M. baccata</i> var. <i>coerulescens</i>	2333	865	45	<i>Cydonia oblonga</i>		889
19	<i>M. cerasifera</i>	29494	864	46	<i>Mespilus germanica</i>		859
20	<i>M. denticulata</i>	29416	862	47	<i>Prunus kansuensis</i> (NC_023956)		869
21	<i>M. hupehensis</i>	14945A	863	48	<i>P. maximowiczii</i> (NC_026981)		863
22	<i>M. mandshurica</i>	14947	860	49	<i>P. mume</i> (NC_023798)		868
23	<i>M. robusta</i>	43199	865	50	<i>P. padus</i> (NC_026982)		856
Section <i>Sorbomalus</i>				51	<i>P. persica</i> (NC_014697)		871
24	<i>M. arnoldiana</i>	2312	863	52	<i>P. pseudocerasus</i> (NC_030599)		863
25	<i>M. florentina</i>	2345	885	53	<i>P. yedoensis</i> (NC_026980)		862
26	<i>M. floribunda</i>	2346	863	54	<i>Pyrus elata</i>		864
27	<i>M. fusca</i>	–	863	55	<i>Sorbus aucuparia</i>		861

Plant material was provided by: (1–26, 28–38) the Vavilov Research Institute of Plant Industry (VIR); (27, 42–44, 55) the Tsitsin State Botanical Garden of the Russian Academy of Sciences; (39–41, 45, 46, 54) the Michurin All-Russian Research Institute of Genetics and Breeding of Fruit Crops; (47–53) NCBI sequences.

большинства исследуемых представителей других родов *Rosaceae* составила 417 п. н. В этом домене у видов *Malus* было идентифицировано четыре SNPs: из них две замены (C242A, A287T) характерны для *M. ioensis*, *M. soulardii*, *M. coronaria*, *M. platycarpa* и *M. florentina*, а одна (T286G) – для *M. denticulata*, *M. orientalis*, *M. purpurea* var. *aldenhamensis*, *M. sieversii* и *M. turkmenorum*. Еще одна замена (C461A) была специфична для *M. florentina* (здесь и

далее положение замен дается относительно выравненной последовательности интрона).
Длина домена I у образцов *Prunus* варьировала за счет дубликации (*P. kansuensis* и *P. persica*) и трех копий (*P. mume*) фрагмента TCAAA. У образцов *Prunus* в последовательности домена идентифицированы две однонуклеотидные и трехнуклеотидная делеции, а также 14 SNPs, характерные только для образцов данного рода.

Table 2. Characterization of the chloroplast *rps16* intron in the Rosaceae species studied

Domain	Domain boundaries	<i>Malus</i> species		Other Rosaceae species (outgroup)	
		Length, bp	SNPs/indels	Length, bp	SNPs/indels
I	5'-CTTGGC 3'-CTGGAACAAG	417	4/0	406; 412; 417; 422	26/6
II	5'-TACCR TT 3'-AAAAGGGTT	79; 103	3/1	79; 80; 81; 103	10/4
III	5'-ACTCAATAWA 3'-AACTTGAGT	63	1/0	63; 68	9/1
IV	5'-GAGGRYCAAA 3'-TTTTTC	141–148	2/0	141; 145; 146; 147; 149; 153	36/5
V	5'-GAGCCGTACG 3'-ATACGTTTC	34	0/0	34	0/0
VI	5'-GGGGGGGT 3'-TCTATCCC	24	0/0	24	0/0

У образцов *Cotoneaster lucidus* и *Chaenomeles japonica* в домене I выявлена делеция 11 п. н. (CTTTGATAGAA) и специфичная замена C276T. По одной специфичной замене идентифицировано для *Cydonia oblonga* (G138A), *Crataegus monogyna* (T285G) и *Pyrus elata* (C331T).

Длина домена II у большинства изученных видов Rosaceae составила 79 п. н., за исключением *Sorbus aucuparia* (80 п. н.) и представителей рода *Prunus* (81 п. н.). У видов рода *Malus* секции Chloromeles и *M. florentina* (секция Sorbomalus) в домене II была детектирована инсерция размером 24 п. н. (AGTAAGAATCAAAATAGATTTCGAA), в данной инсерции выявлены замены у *M. ioensis* (A579T) и у *M. florentina* (G589T). Интересно отметить, что эта вставка, отсутствующая у других анализируемых видов *Malus*, была обнаружена и в последовательности домена II у представителя другого рода – айвы обыкновенной *C. oblonga*. Всего у образцов *Malus* в данном домене выявлено три замены, одна из которых (G513A) характерна для видов *M. arnoldiana*, *M. floribunda* и одного образца *M. purpurea* (2392).

В последовательности домена II у видов Rosaceae, взятых в качестве внешней группы, выявлено 10 переменных сайтов, из которых два специфичны для *P. elata* (C535T, A541G), по одному для *C. monogyna* (G537T) и *Prunus yedoensis* (C531T), а одна замена была общей для образцов рода *Crataegus* и *Mespilus germanica* (G600A). У видов *Prunus* в последовательности домена выявлено пять SNPs, характерных только для образцов этого рода.

Домен III у анализируемых образцов *Malus* был инвариабельным по длине (63 п. н.), в его последовательности идентифицирован только один SNP у *M. transitoria* (A629T). Также единичные замены детектированы у *Prunus padus* (C638A), *P. persica* (C679A) и *P. elata* (G662A), а у *C. lucidus* выявлены четыре специфичные замены (T640A, A643G, C655T, T657C). Последовательность домена всех образцов *Prunus* характеризовалась присутствием двух SNPs (T632G, T633A), при этом у *P. kansuensis* и *P. persica* также выявлена дупликация TAAAA.

Различие в длине домена IV у исследуемых образцов *Malus* обусловлено наличием поли(T)- и поли(G)-районов, которые имеют разную длину. В последовательности до-

мена выявлено по одной специфичной замене у *M. florentina* (A767C) и *M. prunifolia* (T828G). У образцов из других родов Rosaceae в домене IV было идентифицировано 36 SNPs, из которых 12 специфичны для видов *Prunus*. Также у образцов *Prunus* детектированы три специфичные индели. У четырех видов *Prunus* (*P. kansuensis*, *P. mume*, *P. padus*, *P. persica*) в последовательности 5'-границы домена присутствуют две транзиции A/G, а еще у трех видов (*P. maximowiczii*, *P. pseudocerasus*, *P. yedoensis*) – дополнительная транзиция T/C (GAGG A/G T/C A/G CAAA). Интересно отметить, что у последних трех видов обнаружен ряд специфичных замен (C729A, A741G, T783C) и 9-нуклеотидная дупликация ATARACATA.

Домены V и VI у всех анализируемых образцов семейства Rosaceae были инвариабельными по длине и нуклеотидному составу. Сравнение изученных последовательностей этих доменов у видов Rosaceae с последовательностями, представленными в базе данных NCBI, показало, что домен V полностью идентичен у различных групп растений, в то время как в домене VI могут быть полиморфными только три последних нуклеотида. Повидимому, высокая консервативность доменов V и VI свидетельствует об их исключительной роли в сплайсинге данного интрона, так как именно эти последовательности могут узнаваться факторами сплайсинга и мРНК, как, например, это было показано для ряда интронов II группы в митохондриальном геноме, потерявших способность к автономному сплайсингу (Ostheimer et al., 2003; Brown et al., 2014; Zimmerly, Semper, 2015).

В целом интрон гена *rps16* у 38 представителей 32 видов рода *Malus* оказался низкополиморфным. Всего было выявлено 14 SNPs, из которых 10 локализованы в доменах, а четыре – в междоменных областях. Для сравнения: у семи анализируемых видов *Prunus* было идентифицировано 23 SNPs, а также ряд специфичных инделей. Проведенные ранее исследования выявили значительный уровень вариабельности интрона гена *rps16* у различных родов растений. Например, у 39 образцов 7 видов рода *Allium* в интроне было идентифицировано 57 SNPs и 20 инделей (Рыжова и др., 2009), а у 18 видов рода *Morus* выявлено 113 переменных сайтов (Yuhua et al., 2011).

Более того, различия, детектированные в последовательности интрона гена *rps16* у *Malus* и других взятых в анализ представителей восьми родов Rosaceae, за исключением рода *Prunus*, также были невелики. Возможно, это связано с тем, что род *Prunus* относится к надтрибе Kerriodae, а все другие анализируемые виды – к надтрибе Pyrodae семейства Rosaceae (Potter et al., 2007).

Таким образом, впервые был изучен полиморфизм интрона хлоропластного гена *rps16* у видов *Malus* и родственных видов Rosaceae. Показано, что у 32 рассмотренных представителей рода *Malus* данный интрон имеет крайне низкий уровень вариабельности (1.5 %), в то время как у семи представителей рода *Prunus* полиморфизм интрона гена *rps16* был несколько выше (2.25 %). Выявленный сравнительно низкий полиморфизм интрона гена *rps16* не позволяет использовать его для изучения филогении рода *Malus*, однако данная последовательность может быть рекомендована для выяснения филогенетических отношений между родами семейства Rosaceae. В последовательности интрона были определены границы всех шести доменов, характерных для интронов данного типа, и охарактеризована их вариабельность. Идентифицированные SNPs локализовались в доменах I–IV, при этом в домене II у пяти видов *Malus* выявлена 24-нуклеотидная вставка, характерная также для айвы обыкновенной (*Cydonia oblonga*). Домены V и VI были инвариабельными у всех анализируемых образцов *Malus* и родственных видов Rosaceae, что, по-видимому, связано с их исключительной ролью в вырезании интрона.

Acknowledgments

This work was supported by State Budgeted Project 0104-2014-0210 and the Russian Foundation for Basic Research, project 16-34-60220.

Conflict of interest

The authors declare no conflict of interest.

References

- Brown G.G., Francs-Small C.C., Ostersetter-Biran O. Group II intron splicing factors in plant mitochondria. *Front. Plant Sci.* 2014;5:1-13. DOI 10.3389/fpls.2014.00035.
- Carrillo C., Chapdelaine Y., Bonen L. Variation in sequence and RNA editing within core domains of mitochondrial group II introns among plants. *Mol. Gen. Genet.* 2001;264(5):595-603. DOI 10.1007/s004380000345.
- Dai L., Chai D., Gu S.Q., Gabel J., Noskov S.Y., Blocker F.J., Lambowitz A.M., Zimmerly S. A three-dimensional model of a group II intron RNA and its interaction with the intron-encoded reverse transcriptase. *Mol. Cell.* 2008;30(4):472-485. DOI 10.1016/j.molcel.2008.04.001.
- Jansen R.K., Cai Z., Raubeson L.A., Daniell H., dePamphilis C.W., Leebens-Mack J., Müller K.F., Guisinger-Bellian M., Haberle R.C., Hansen A.K., Chumley T.W., Lee S.B., Peery R., McNeal J.R., Kuehl J.V., Boore J.L. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc. Natl. Acad. Sci. USA.* 2007;104(49):19369-19374. DOI 10.1073/pnas.0709121104.
- Kelchner S.A. Group II introns as phylogenetic tools: structure, function, and evolutionary constraints. *Am. J. Bot.* 2002;89(10):1651-1669. DOI 10.3732/ajb.89.10.1651.
- Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 2016;33:1870-1874. DOI 10.1093/molbev/msw054.
- Lei W., Ni D., Wang Y., Shao J., Wang X., Yang D., Wang J., Chen H., Liu C. Intraspecific and heteroplasmic variations, gene losses and inversions in the chloroplast genome of *Astragalus membranaceus*. *Sci. Rep.* 2016;6:21669. DOI 10.1038/srep21669.
- Lo E.Y.Y., Donoghue M.J. Expanded phylogenetic and dating analyses of the apples and their relatives (Pyraeae, Rosaceae). *Mol. Phylogen. Evol.* 2012;63(2):230-243. DOI 10.1016/j.ympev.2011.10.005.
- Michel F., Umesono K., Ozeki H. Comparative and functional anatomy of group II catalytic introns – a review. *Gene.* 1989;82(1):5-30.
- Ostheimer G., Williams-Carrier R., Belcher S., Osborne E., Gierke J., Barkan A. Group II intron splicing factors derived by diversification of an ancient RNA binding module. *EMBO J.* 2003;22(15):3919-3929. DOI 10.1093/emboj/cdg372.
- Potter D., Eriksson T., Evans R.C., Oh S., Smedmark J., Morgan D.R., Kerr M., Robertson K.R., Arsenault M., Campbell C.S. Phylogeny and classification of Rosaceae. *Plant Syst. Evol.* 2007;266:5-43. DOI 10.1007/s00606-007-0539-9.
- Ryzhova N.N., Slugina M.A., Kochieva E.Z., Skryabin K.G. Polymorphism and structural variation of *rps16* group-II intron in the *Solanum* species. *Genetika = Genetics (Moscow).* 2013;49(7):717-721. DOI 10.1134/S1022795413070120.
- Ryzhova N.N., Kholda O.A., Kochieva E.Z. Structural characteristics of the chloroplast *rps16* intron in *Allium sativum* and related *Allium* species. *Molekulyarnaya biologiya = Molecular Biology (Moscow).* 2009;43(5):766-775. DOI 10.1134/S0026893309050082.
- Shaw J., Lickey E.B., Beck J.T., Farmer S.B., Liu W., Miller J., Siripun K.C., Winder C.T., Schilling E.E., Small R.L. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 2005;92(1):142-166. DOI 10.3732/ajb.92.1.142.
- Smedmark J.E.E., Rydin C., Razafimandimbison S.G., Khan S.A., Liede-Schumann S., Bremer B. A phylogeny of Urophylleae (Rubiaceae) based on *rps16* intron data. *Taxon.* 2008;57(1):24-32. DOI 10.5167/uzh-11754.
- Toor N., Hausner G., Zimmerly S. Coevolution of group II intron RNA structures with their intron encoded reverse transcriptases. *RNA.* 2001;7(8):1142-1152. DOI 10.1017/S1355838201010251.
- Vogel J., Hubschmann T., Borner T., Hess W.R. Splicing and intron-internal RNA editing of *trnK-matK* transcripts in barley plastids: support for MatK as an essential splice factor. *J. Mol. Biol.* 1997;270(2):179-187. DOI 10.1006/jmbi.1997.1115.
- Yuhua W., Wei W., Wei T., Weiguo Z. Analysis of chloroplast ribosomal subunit S16 (*rpS16*) intron sequences in *Morus* (Urticales: Moraceae). *Afric. J. Biotechnol.* 2011;10(77):17695-17699. DOI 10.5897/AJB11.2218.
- Zimmerly S., Semper C. Evolution of group II introns. *Mobile DNA.* 2015;6:7. DOI 10.1186/s13100-015-0037-5.