

ГЕНОМИКА.
ТРАНСКРИПТОМИКА

УДК 575.174.015.3

ИДЕНТИФИКАЦИЯ И АНАЛИЗ НУКЛЕОТИДНОГО ПОЛИМОРФИЗМА
ГЕНОВ *Brassica rapa* (РЕПА), КОДИРУЮЩИХ БЕЛКИ С ДОМЕНОМ
ХОЛОДОВОГО ШОКА (CSDP)

© 2013 г. Н. Н. Рыжова^{1*}, М. А. Филюшин¹, А. М. Артемьева², М. В. Бердникова³,
В. В. Таранов³, А. В. Бабаков³, Е. З. Кочиева¹

¹Центр “Биоинженерия” Российской академии наук, Москва, 117312

²Всероссийский научно-исследовательский институт растениеводства им. Н.И. Вавилова, Санкт-Петербург, 190000

³Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии
Российской академии сельскохозяйственных наук, Москва, 127550

Поступила в редакцию 17.04.2012 г.

Принята к печати 13.06.2012 г.

Получены полные нуклеотидные последовательности генов холодого шока *BrCSDP2* и *BrCSDP4* *Brassica rapa*. Установлено, что выделенные гены принадлежат к группе генов *AtCSP2/AtCSP4 Arabidopsis thaliana* и *TsCSDP2/TsCSDP4 Thellungiella salsuginea*, кодирующих белки с доменом холодого шока (CSD) и двумя мотивами так называемых цинковых пальцев. Описана структура и охарактеризованы аллельные варианты этих генов. Показано, что аллельный полиморфизм этих генов обусловлен как точковыми нуклеотидными заменами, так и небольшими инделями. Уровни генетического сходства варьировали от 1.0 (между *BrCSDP* из отдельных растений образцов *B. rapa*) до 0.53 (между *AtCSDP1* и *AtCSDP4*). В свою очередь, коэффициенты сходства нуклеотидных последовательностей *BrCSDP2* и *AtCSDP2* составили 0.89, а *BrCSDP4* и *AtCSDP4* – 0.85. В продуктах *in silico* трансляции выделенных генов выявлены аминокислотные замены. При этом, однако, не обнаружено корреляции между полиморфизмом полученных последовательностей и признаками устойчивости к холодовому стрессу.

Ключевые слова: домен холодого шока, структура и аллельный полиморфизм, репа, семейство Brassicaceae, нуклеотидные и аминокислотные замены, индели, эволюция и филогения.

IDENTIFICATION AND NUCLEOTIDE POLYMORPHISMS IN *BRASSICA RAPA* GENES CODING COLD SHOCK DOMAIN PROTEINS (CSDP), by N. N. Ryzhova^{1*}, M.A. Filiushin¹, A.M. Artemjeva², M.V. Berdnikova³, V.V. Taranov³, A. V. Babakov³, E. Z. Kochieva¹ (¹Centre “Bioengineering”, Russian Academy of Sciences, Moscow, 117312 Russia; ²All Russia Vavilov Research Institute of Plant Industry, St. Petersburg, 190000, Russia; ³All Russia Research Institute of Agricultural Biotechnology, Russian Academy of Agricultural Sciences, Moscow, 127550 Russia). Full-length *BrCSDP2* and *BrCSDP4* cold shock gene sequences of *Brassica rapa* are obtained. It is shown that the isolated genes belong to a group *AtCSP2/AtCSP4* of *Arabidopsis thaliana* and *TsCSDP2/TsCSDP4* of *Thellungiella salsuginea* genes encoding proteins with a cold shock domain (CSD) and two zinc finger motives. The structure and the allelic variants of these genes are described and characterized. It is shown that the identified allelic polymorphism is due to both of point substitutions and small indels. Coefficients of total genetic similarity ranged from 1.0 to 0.53. In tern the genetic similarity coefficient for *BrCSDP2* and *AtCSDP2* was 0.89, and for *BrCSDP4* and *AtCSDP4* was 0.85. Translation *in silico* of gene sequences has revealed amino acid substitutions in the protein sequence, but no significant correlation between the detected polymorphism and signs of resistance to cold stress were found.

Key words: cold shock genes, structure and allelic polymorphism, turnips, Brassicaceae, nucleotide and amino acid substitutions, indels, evolution and phylogeny.

DOI: 10.7868/S0026898412060158

Белки с доменом холодого шока (Cold Shock Domain, CSD) образуют консервативное семейство, широко представленное как у про-, так и у

эукариот. Все белки этого семейства содержат CSD-домен с двумя консервативными мотивами RNP1 (KGFGFI)/RNP2 (LFVHQ), обогащенны-

Принятые сокращения: CSD – домен холодого шока (cold shock domain); CSDP – белки с доменом холодого шока; CSP – белки холодого шока; ZnF – цинковый палец (zinc finger); а.о. – аминокислотный остаток (при числе). GS – коэффициент генетического сходства (genetic similarity).

* Эл. почта: rynatalia@yandex.ru

ми остатками ароматических аминокислот и не-обходимыми, по всей видимости, для связывания с нуклеиновыми кислотами [1, 2]. Одна из основных функций белков этого семейства – участие в холодовой адаптации, поскольку CSD-домен обладает способностью связываться и дестабилизировать вторичные структуры РНК, которые образуются при пониженных температурах, и восстанавливать тем самым их функции при холодовом стрессе [2, 3]. Сравнительно недавно показали, что эти белки участвуют не только в обеспечении устойчивости растений к пониженным температурам, но в нормальных условиях они могут вовлекаться в другие процессы, например, в такие, как регуляция развития зародыша, времени цветения и формирования плодов у растений [4–6]. Предполагается, что подобная множественность функций белков с CSD (CSDP) определяется присутствием в их составе дополнительных доменов и регуляторных мотивов [7–9].

В настоящее время гены CSDP идентифицированы у 19 родов растений [10]. Их число варьирует от двух в геноме *Oryza sativa* и *Zea mays*, до семи у *Arabidopsis thaliana* и *Glycine max*, причем отдельные виды могут содержать различные типы генов/белков CSD [9, 10].

CSDP растений отличаются от всех подобных белков бактерий и животных тем, что в их С-концевой части локализованы так называемые мотивы цинковых пальцев (ZnF) ретровирусного типа (CX₂CX₄HX₄C, (CCHC)), разделенные Gly-богатыми участками [3]. У однодольных растений найдены белки, содержащие от двух до четырех мотивов CCHC (“цинковые пальцы”), в то время как у двудольных обнаружены CSDP с двумя–семью такими мотивами. Предполагается, что различные комбинации ZnF определяют специфичность действия CSDP в растениях [5, 9, 11, 12].

В геноме арабидопсиса найдены четыре гена CSDP, продукты которых можно разделить на две группы: с двумя ZnF – AtCSP2 и AtCSP4 и с семью ZnF – AtCSP1 и AtCSP3. Известно, что экспрессия генов белков этих групп дифференциально регулируется в ответ на различные внешние сигналы [4, 7, 9].

У *Thellungiella salsuginea* (= *Eutrema salsugineum*) – растения, обладающего повышенной устойчивостью к низким температурам, также идентифицированы четыре гена белков с CSD-доменом (TsCSDP1–4) [13]. Все они, как и аналогичные белки других растений, содержат высококонсервативный N-концевой CSD-домен. У TsCSDP2 и TsCSDP4 в С-концевом домене локализованы по два ZnF, а у TsCSDP1 и TsCSDP3 – шесть и семь ZnF соответственно.

В настоящее время нуклеотидные последовательности всех генов CSDP определены только у 11 видов растений [10]. Несмотря на определен-

ные успехи в изучении структуры CSDP, до сих пор неясными остаются происхождение и эволюция генов этого семейства у растений, отсутствуют данные о генетическом разнообразии и аллельном полиморфизме этих генов и о их роли в формировании признаков холодоустойчивости. Взаимосвязь функций генов этого семейства с устойчивостью к холоду и органогенезом делает изучение генов семейства CSDP весьма актуальным для прикладной генетики и селекции культурных растений.

Таким образом, целью данной работы стала идентификация генов *Brassica rapa* (репа), кодирующих CSDP, анализ их структуры и полиморфизма. Известно, что культурный вид *B. rapa*, который характеризуется внутривидовыми различиями в устойчивости к холодovому стрессу, является одним из видов, наиболее родственных *A. thaliana* (сем. Brassicaceae).

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Из коллекции ГНУ ВИР были отобраны образцы *B. rapa*, характеризующиеся различной устойчивостью/восприимчивостью к холодovому стрессу. В качестве контроля использовали образцы *A. thaliana* и *T. salsuginea* (= *Eutrema salsugineum*) (сем. Brassicaceae, табл. 1).

ДНК выделяли из стерильных тканей проростков отобранных образцов *B. rapa*, *A. thaliana* и *T. salsuginea* согласно стандартному протоколу [14].

Для ПЦР-амплификации генов CSDP *B. rapa* использовали пары праймеров: **FB2** 5'-GAGAA-GATGAGTGGMGAMAAC-3'/**RB2** 5'-GTTTCGW-TTYAACGTCCACC-3' и **FB4** 5'-GGAGAGAGA-GDTTYGATCAG-3'/**RB4** 5'-GTTTTKCTTYARC-GGSCGCC-3'. Праймеры были подобраны на основе нуклеотидных последовательностей, кодирующих CSD-домены у различных видов растений, в том числе *AtCSDP1* (At4g36020/NM_119769.2), *AtCSDP3* (At2g17870/NM_127341.3), *AtCSDP2* (At4g38680/NM_120029.2) и *AtCSDP4* (At2g21060/NM_127676.2) *A. thaliana*, *TsCSDP1* (GQ227856), *TsCSDP2* (GQ227854), *TsCSDP3* (GQ227855) *T. salsuginea*, а также последовательности *B. rapa* L47853, представленной в базе данных NCBI.

Смесь для ПЦР (15 мкл) готовили на льду. Смесь содержала 160 мкМ каждого dNTP, 1.6 мМ MgCl₂, 0.3 мкМ каждого праймера, 0.3 ед. Taq-полимеразы, 1.5 мкл 10× буфера (“Диалат ЛТД”, Москва) и 50–100 нг ДНК. Реакцию амплификации проводили в амплификаторе “Applied Biosystems” в следующем режиме: денатурация – 94°C, 30 с; отжиг праймеров при соответствующей температуре – 30 с; синтез – 72°C, 1 мин, 30–35 циклов; завершающая элонгация 72°C, 10 мин.

Результаты амплификации проверяли путем электрофореза в 1.5%-ном агарозном геле в 1×

TBE. Положительным результатом считали получение фрагмента ДНК длиной примерно 600 п.н. Размер фрагментов оценивали с использованием маркера молекулярных масс с шагом 100 п.н. GeneRuler™100bp DNA Ladder Plus (“Fermentas”, Литва).

ПЦР-фрагменты клонировали в векторе pDrive Cloning Vector с использованием набора QIAGEN PCR Cloning Kit и компетентных клеток QIAGEN EZ Competent Cells (“QIAGEN”, Нидерланды). Амплифицированные фрагменты секвенировали как с 5'-, так и с 3'-концов при помощи системы Big dye (“Applied Biosystems”) на ABI 310 capillary DNA Analyzer (Центр “Биоинженерия”, Москва). Соответствие полученных нуклеотидных последовательностей гену *CSDP* проверяли при помощи программы BLAST 2.2.26+ [15]. Трансляцию и анализ полиморфизма проводили в программе MEGA5 [16]. Расчет коэффициентов генетического сходства (GS, genetic similarity)/генетических расстояний (GD, genetic distance) и реконструкцию эволюционно-филогенетических деревьев выполняли в PAUP 4.0b10 [17].

Филогенетические деревья строили с использованием метода максимальной экономии (Maximum Parsimony, MP) в программе PAUP 4.0b10 (Swofford, 2002). Устойчивость (статистическую поддержку) филогенетических деревьев в MP-

анализе оценивали с помощью 1000 реплик бутстрепа [18]. Значения процента бутстрепа (Bootstrap Percentage, BP) менее 50 не рассматривали.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для идентификации генов *CSDP* и анализа их нуклеотидного полиморфизма использовали шесть образцов *B. rapa* с различной способностью к холодовой акклиматизации. Каждый образец был представлен одним–четырьмя генотипами индивидуальных растений.

Использование разработанных праймеров позволило у всех образцов *B. rapa* амплифицировать ПЦР-фрагменты ДНК длиной около 600 п.н., 27 из которых клонировали и секвенировали. Последующий анализ подтвердил специфичность амплификации и природу амплифицированных последовательностей.

Все нуклеотидные последовательности содержали открытую рамку считывания, старт- и стоп-кодона. Полная длина генов варьировала от 549 до 603 п.н. (табл. 2). Анализ продуктов трансляции этих последовательностей подтвердил присутствие в них всех ключевых доменов и мотивов, характерных для белков семейства с CSD. Все они содержали CSD с каноническими инвариантными мотивами RNP1 (KGFGFI)/RNP2 (LFVHQ)

Таблица 1. Образцы видов семейства Brassicaceae и их устойчивость к холодовому стрессу (данные ГНУ ВИР)

№	Образец	Каталожный номер (ГНУВИР)	Сорт	Происхождение	Устойчивость к холодовому стрессу
1	<i>B. rapa</i> 6/1	381	—	Вьетнам	Неустойчивый
2	<i>B. rapa</i> 6/2	381	—	Вьетнам	Неустойчивый
3	<i>B. rapa</i> 6/3	381	—	Вьетнам	“_”
4	<i>B. rapa</i> 6/4	381	—	Вьетнам	“_”
5	<i>B. rapa</i> 8/1	391	Xing Jang	Китай	Устойчивый
6	<i>B. rapa</i> 8/2	391	Xing Jang	Китай	Устойчивый
7	<i>B. rapa</i> 8/3	391	Xing Jang	Китай	“_”
8	<i>B. rapa</i> 8/4	391	Xing Jang	Китай	“_”
9	<i>B. rapa</i> 12/1	155	Майцай	Китай	Устойчивый
10	<i>B. rapa</i> 12/2	155	Майцай	Китай	Устойчивый
11	<i>B. rapa</i> 12/3	155	Майцай	Китай	“_”
12	<i>B. rapa</i> 12/4	155	Майцай	Китай	“_”
13	<i>B. rapa</i> 14/1	157	Аншун	Китай	Неустойчивый
14	<i>B. rapa</i> 14/2	157	Аншун	Китай	Неустойчивый
15	<i>B. rapa</i> 14/3	157	Аншун	Китай	“_”
16	<i>B. rapa</i> 14/4	157	Аншун	Китай	“_”
17	<i>B. rapa</i> CR	—		Россия	Устойчивый
18	<i>B. rapa</i> НТ	—		Россия	Неустойчивый
19	<i>A. thaliana</i>	—		Россия	Не определен
20	<i>T. salsuginea</i>	—		Россия	Не определен

Таблица 2. Сравнительная характеристика генов и белков CSD *B. rapa*, *A. thaliana*, *T. salsuginea*

Вид, ген	ДНК CSDP, п.н.	Белок CSDP, а.о.	Домен CSD, а.о.	Gly1, а.о.	ZnF1	Gly2, а.о.	ZnF2
<i>A. thaliana</i> AtCSP2	612	204	77	53	CYKCGEPGHMARDC	40	CYSCGESGHFARDC
<i>A. thaliana</i> AtCSP4	606	202	81	56	CFKCGEPGHMAREC	30	CYSCGESGHFARDC
<i>T. salsuginea</i> TsCSDP2	606	202	78	47	CYKCGEPGHMARDC	44	CYSCGESGHFARDC
<i>T. salsuginea</i> TsCSDP4	—	—	—	47	CYKCGEPGHMARDC	31	CYSCGESGHFARDC
<i>B. rapa</i> BrCSP2	549	183	79	50	CYKCGEPGHMAREC	20	CYSCGESGHFSRDC
<i>B. rapa</i> BrCSP4	600/603	200/201	79	60/62	CYKCGEPGHMARDC	26/27	CYSCGESGHFARDC

(рис. 1), а также по два мотива ССНС, чередующихся с Gly-богатыми повторами. Длина продуктов транскрипции генов *B. rapa* составила 183 и 200/201 а.о. (табл. 2).

Полиморфизм нуклеотидных последовательностей генов *BrCSDP2* и *BrCSDP4*

Сравнительный анализ полиморфизма полученных нами генов *CSDP B. rapa* и всех четырех генов *A. thaliana* показал, что у *B. rapa* эти гены представлены двумя типами последовательностей, гомологичных последовательностям коротких генов *AtCSDP2* и *AtCSDP4 A. thaliana*, кодирующих белки с двумя мотивами ZnF (рис. 1).

Из 23 последовательностей *B. rapa* 11 характеризовались наибольшим сходством с геном *AtCSDP2* (0.89 GS) и были названы *BrCSDP2*. Оставшиеся 12 были наиболее сходны с геном *AtCSDP4* (0.85 GS), их обозначили как *BrCSDP4*. В целом уровни генетического сходства варьировали от 1.00 (между *BrCSDP* отдельных растений в образцах *B. rapa*) до 0.53 (между *AtCSDP1* и *AtCSDP4*).

Сравнение нуклеотидных последовательностей *BrCSDP2/BrCSDP4* и генов другого охарактеризованного представителя Brassicaceae — *T. salsuginea* — выявило также несколько большее сходство между генами *BrCSDP4* и *TsCSDP2* (0.86 GS), чем между *BrCSDP4* и *AtCSDP4* (0.85 GS), что согласуется с филогенетическими данными. Виды *T. salsuginea* и *B. rapa* относят к одной эволюционной линии, в то время как *A. thaliana* по ряду молекулярных данных принадлежит к другой эволюционной линии Brassicaceae [19].

Интересно, что уровни генетического сходства между *AtCSDP2* и *AtCSDP4* (0.91) были больше, чем между *BrCSDP2* и *BrCSDP4* (0.88), *BrCSDP2* и *AtCSDP2* (0.89), а также *BrCSDP4* и *AtCSDP4* (0.85). Это может говорить о том, что *CSDP4* произошел от *CSDP2* уже после разделения эволюционных линий *A. thaliana* и *B. rapa*. Кроме того, у *B. rapa* гены *BrCSDP2* и *BrCSDP4* разделились, по всей видимости, до того, как произошла дивергенция генов *AtCSDP2* и *AtCSDP4* у *A. thaliana*.

В генах *BrCSDP2* и *BrCSDP4* замены локализовались в области, кодирующей как CSD и ZnF, так и Gly-богатые повторы. В целом нуклеотидные последовательности генов *BrCSDP2* и *BrCSDP4* различались 138 заменами (22% от выровненной последовательности длиной 621 п.н.). Из них 30 (14.5% из 207 а.о.) приводили к аминокислотным заменам, специфичным для каждого из белков *BrCSDP2* и *BrCSDP4*. Нуклеотидные последовательности *BrCSDP2* и *BrCSDP4* различались не только SNP, но и рядом инделей, локализованных, главным образом, в областях, кодирующих Gly-богатые участки и CSD.

Анализ полиморфизма отдельных структурных элементов генов *BrCSP2* и *BrCSP4 B. rapa*

Домен холодого шока. Как у *BrCSDP2*, так и у *BrCSDP4* CSD состоял из 79 а.о., что сравнимо с длиной CSD у *A. thaliana* (*AtCSDP2* — 77 а.о., *AtCSDP4* — 81 а.о.) и *T. salsuginea* (*TsCSDP2* — 78 а.о.) (рис. 1, табл. 2).

Как и ожидалось, аминокислотные/нуклеотидные последовательности консервативных мотивов RNP1 (KGFQFI) и RNP2 (LFVHQ) в CSD-домене были идентичными как в *BrCSDP2/BrCSDP4*, так и в *AtCSDP2/AtCSDP4* у *A. thaliana* и в *TsCSDP2/TsCSDP4* у *T. salsuginea*. Другие участки CSD в исследованных генах/белках были гораздо более полиморфными. В области CSD-домена гены *BrCSDP2* и *BrCSDP4* различались 50 SNP, 15 из которых приводили к аминокислотным заменам. Помимо точковых замен, кодирующие CSD последовательности генов *BrCSDP2* и *BrCSP4* различались двумя трехнуклеотидными инделями и, соответственно, присутствием дополнительных аминокислотных остатков — Gly в *BrCSDP4* и Ser в *BrCSDP2* (рис. 1).

Мотивы ZnF. Как упоминалось выше, С-концевые области белков *BrCSDP2* и *BrCSDP4* содержат по два консервативных мотива ZnF ретровирусного типа (С2ХС4ХН4ХС). Аминокислотная последовательность второго ZnF характеризовалась наибольшей консервативностью. Единственная аминокислотная замена Ser → Ala была специфич-

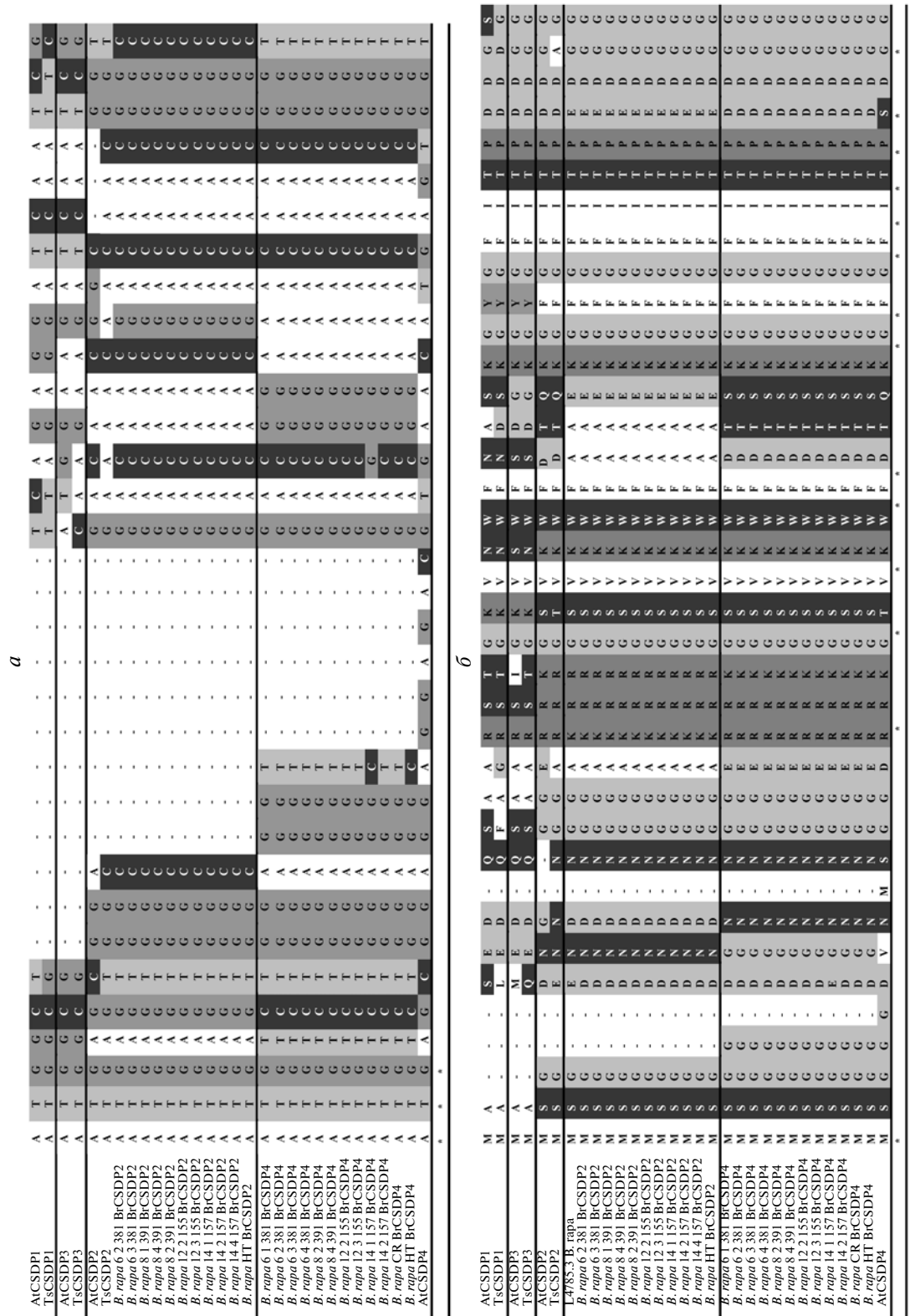


Рис. 1. Фрагмент выравнивания нуклеотидных (а) и аминокислотных (б) последовательностей CSD-домена генов/белков *B. rapa*, *A. thaliana*, *T. salsuginea*.

ной для мотива второго ZnF в BrCSDP2. Интересно, что этот же белок BrCSDP2 содержал специфический мотив первого ZnF с заменой Asp → Glu.

Gly-богатые повторы. Наиболее полиморфными в генах *B. rapa*, *A. thaliana* и *T. salsuginea* были участки, кодирующие Gly-богатые повторы, следующие за CSD и первым ZnF, при этом значительные различия обнаружены как на уровне ДНК, так и белка. Так длина Gly-участка, фланкирующего CSD-домен в BrCSDP4, варьировала от 60 до 62 а.о., а в BrCSDP2 была равна 50 а.о. В свою очередь, у арабидопсиса этот участок состоит из 53 и 56 а.о., а у *T. salsuginea* – 47 а.о. (табл. 2).

Еще большие различия обнаружены во втором Gly-богатом повторе, следующем за первым ZnF. У *B. rapa* его длина составила 20 а.о. в BrCSDP2 и 26/27 а.о. в BrCSDP4, у *A. thaliana* и *T. salsuginea* – 40 и 30 а.о. в AtCSDP2 и AtCSDP4 соответственно, 44 и 31 а.о. – в TsCSDP4 и TsCSDP2 соответственно (табл. 2). Помимо инделей Gly-богатые повторы содержали множественные SNP.

Внутривидовой полиморфизм

Помимо идентификации и характеристики двух генов CSDP с двумя мотивами ZnF особый интерес представлял анализ внутривидового аллельного полиморфизма каждого из этих генов с поиском возможной корреляции между аминокислотными заменами и известными признаками, в том числе со способностью *B. rapa* к быстрой холодовой акклиматизации. Всего в растениях из шести образцов *B. rapa* с разным уровнем холодоустойчивости выявлено 11 различных последовательностей гена *BrCSDP2* и 12 – *BrCSDP4* (табл. 1). Так как *B. rapa* относится к перекрестно опыляемым растениям, каждый образец был представлен несколькими (1–4) генотипами.

В данном наборе генотипов найдено всего три аллеля гена *BrCSDP2* и пять аллелей гена *BrCSDP4*. При этом ни один из трех генотип-специфичных SNP в гене *BrCSDP2* не приводил к аминокислотным заменам в соответствующих белках. В свою очередь, два из пяти генотип-специфичных SNP гена *BrCSDP4* приводили к аминокислотным заменам в белке, что указывает на несколько больший уровень полиморфизма гена *BrCSDP4*.

Обнаруженные генотип-специфичные аллели гена *BrCSDP4* *B. rapa* также были связаны с инделями в Gly-богатых повторах. Эти индели соответствовали inserции 9 п.н., GGAGGAGGA₃₇₃₋₃₈₁, и двум трехнуклеотидным делециям GGC₃₆₁₋₃₆₃ и GGT₄₈₇₋₄₈₉ в образцах *B. rapa* 14/1 (#157) и *B. rapa* НТ.

Таким образом, по числу нуклеотидных/аминокислотных замен, а также инделей внутривидовой полиморфизм гена *BrCSDP4* был выше, чем у гена *BrCSDP2*, а аминокислотная последова-

тельность белка BrCSDP2 оставалась полностью консервативной во всех анализируемых образцах. В этой связи следует отметить, что хотя функциональные особенности генов *BrCSDP2* и *BrCSDP4* не установлены, их ортологи *AtCSDP2* и *AtCSDP4* в геноме *A. thaliana* имеют сходный паттерн экспрессии в процессе развития растения [5]. Кроме того, было установлено, что содержание белка AtCSDP2 выше, чем AtCSDP4, что указывает на меньшее значение белка AtCSDP4, однако существует возможность функциональной взаимозаменяемости этих двух генов [7].

Таким образом, если белки BrCSP2 и BrCSP4 выполняют такие же функции, как и AtCSDP2 и AtCSDP4 [4, 5, 8], можно предположить, что ген *BrCSP4* представляет собой дивергентную копию *BrCSP2* и не оказывает критического влияния на признак. Белковые продукты этих генов могут быть взаимозаменяемыми, как AtCSDP2 и AtCSDP4, а основное значение *BrCSP4* связано, предположительно, с поддержанием геномного полиморфизма генов этого семейства и его эволюцией [20].

Отдельный интерес представлял анализ возможных корреляций между аминокислотным полиморфизмом белков BrCSDP2/BrCSDP4 и признаком холодовой устойчивости. В настоящее время скрининг популяций и образцов, представленных в банках данных, довольно успешно используется для поиска аллельных вариантов, ассоциированных с искомыми признаками растений [20–23].

Нами показано, что в целом на основе идентифицированных инделей и замен в гене *BrCSDP4* *B. rapa* можно выделить две основные группы генотипов. При этом, однако, не выявлено корреляции между обнаруженными аминокислотными заменами и признаками устойчивости к холодовому стрессу. Кроме того, идентифицированные аллели не были специфичными для образца. Согласно полученным данным, генотипы отдельных растений сорта Аншун #157 характеризовались присутствием двух различных аллелей гена *BrCSDP4* (табл. 3), что указывает на генетическую гетерогенность сорта и возможное существование не только двух типов гомозиготных, но и гетерозиготных генотипов. Аналогичные ситуации достаточно часто встречаются у перекрестно опыляемых и полиплоидных видов [24–26].

В данном случае отсутствие корреляции между выявленными заменами и признаками устойчивости к холодовому стрессу может говорить о более сложном механизме формирования холодовой устойчивости у *B. rapa*, которая определяется не только идентифицированными генами, но и целым рядом других генов и ферментов. Кроме того, подобно генам *AtCSDP2/AtCSDP4* *A. thaliana* выделенные гены *B. rapa* могут участвовать в дру-

гих регуляторных реакциях в клетке, в частности, в регуляции эмбриогенеза и прорастания семян [6, 7].

Филогенетический анализ нуклеотидных последовательностей генов

На основе установленного полиморфизма рассчитаны генетические расстояния и проведен филогенетический анализ генов *CSDP B. rapa*, а также ранее известных генов *CSDP A. thaliana* и *T. salsuginea*, входящих в семейство Brassicaceae.

Анализ проводили с использованием как последовательностей, кодирующих CSD-домен (данные не приведены), так и полных генов *CSDP*.

Все построенные деревья имели в целом сходную топологию. Основное различие состояло в том, что использование полных нуклеотидных последовательностей генов давало более достоверную поддержку клад, по всей видимости, связанную с увеличением числа филогенетически информативных признаков при увеличении длины анализируемой последовательности и включением в анализ 3'-концевых областей, кодирующих ZnF. Аналогичное повышение достоверности клад наблюдалось также при анализе генов, кодирующих белки с CSD у гребешка *Chlamys farreri* [27]. По всей видимости, несмотря на сравнительно высокий полиморфизм области, кодирующей мотивы ZnF и Gly-богатые повторы, эта часть гена также несет важную эволюционно-филогенетическую информацию.

В остальном анализируемые нуклеотидные последовательности формировали на дендрограммах три основные эволюционные ветви — гены *CSDP1* (BP 91%), гены *CSDP3* (BP 98%), кодирующие белки с шестью—семью мотивами ZnF, и гены *CSDP2/CSDP4 B. rapa*, *A. thaliana* и *T. salsuginea*, кодирующие белки с двумя мотивами ZnF (BP 100%) (рис. 2). Полученная топология строго подтверждает принадлежность идентифицированных *BrCSDP2/BrCSDP4 B. rapa* к группе генов Brassicaceae, кодирующих два мотива ZnF.

Как видно на дендрограмме (рис. 2), все 11 нуклеотидных последовательностей гена *BrCSDP2 B. rapa* формируют единую группу (BP 100%), входящую в один кластер с генами *AtCSDP2/TsCSDP2* (BP 87%), который, в свою очередь, объединяется с *AtCSDP4*. В то же время, нуклеотидные последовательности гена *BrCSDP4 B. rapa* образуют отдельный удаленный кластер, показывая более высокую дивергенцию генов *BrCSDP2* и *BrCSDP4 B. rapa* по сравнению с генами *AtCSDP2* и *AtCSDP4* арабидопсиса.

На обеих дендрограммах прослеживается явная дифференциация образцов в соответствии с аллельным полиморфизмом генов *BrCSDP2* и *BrCSDP4*.

Мы использовали также последовательность L47853 *B. rapa*, полученную ранее [28] и проанализированную [10]. Результаты проведенного нами сравнительного исследования показывают, что эта последовательность кодирует только CSD и не содержит 3'-концевой области, соответству-

Таблица 3. Идентифицированные SNP гена *BrCSP4* и аминокислотные замены в соответствующем белке у проанализированных генотипов *B. rapa*

Вид/образец	SNP						Аминокислотные замены		
	C ₁₂ T	G ₁₅ C	C ₁₂₀ T	C ₁₇₇ T	G ₂₁₇ T	G ₅₀₂ A	E/D ₅ синонимичная замена	S/A ₇₃ несинонимичная замена	S/G несинонимичная замена
<i>B. rapa</i> 6/1	T	C	T	T	K	A	D	S/A	S
<i>B. rapa</i> 6/2	T	C	T	T	G	A	D	A	S
<i>B. rapa</i> 6/3	T	C	T	T	T	A	D	S	S
<i>B. rapa</i> 6/4	T	C	T	T	T	A	D	S	S
<i>B. rapa</i> 8/2	T	C	T	T	T	A	D	S	S
<i>B. rapa</i> 8/4	T	C	T	T	T	A	D	S	S
<i>B. rapa</i> 12/2	T	C	T	T	G	A	D	A	S
<i>B. rapa</i> 12/3	T	C	T	T	T	A	D	S	S
<i>B. rapa</i> 14/1	C	G	C	T	G	G	E	A	G
<i>B. rapa</i> 14/2	T	C	T	T	T	A	D	S	S
<i>B. rapa</i> CR	T	C	T	T	C	A	D	S	S
<i>B. rapa</i> HT	C	C	T	C	G	G	D	A	G

Примечание. Подчеркнуты SNP, приводящие к замене аминокислотных остатков.

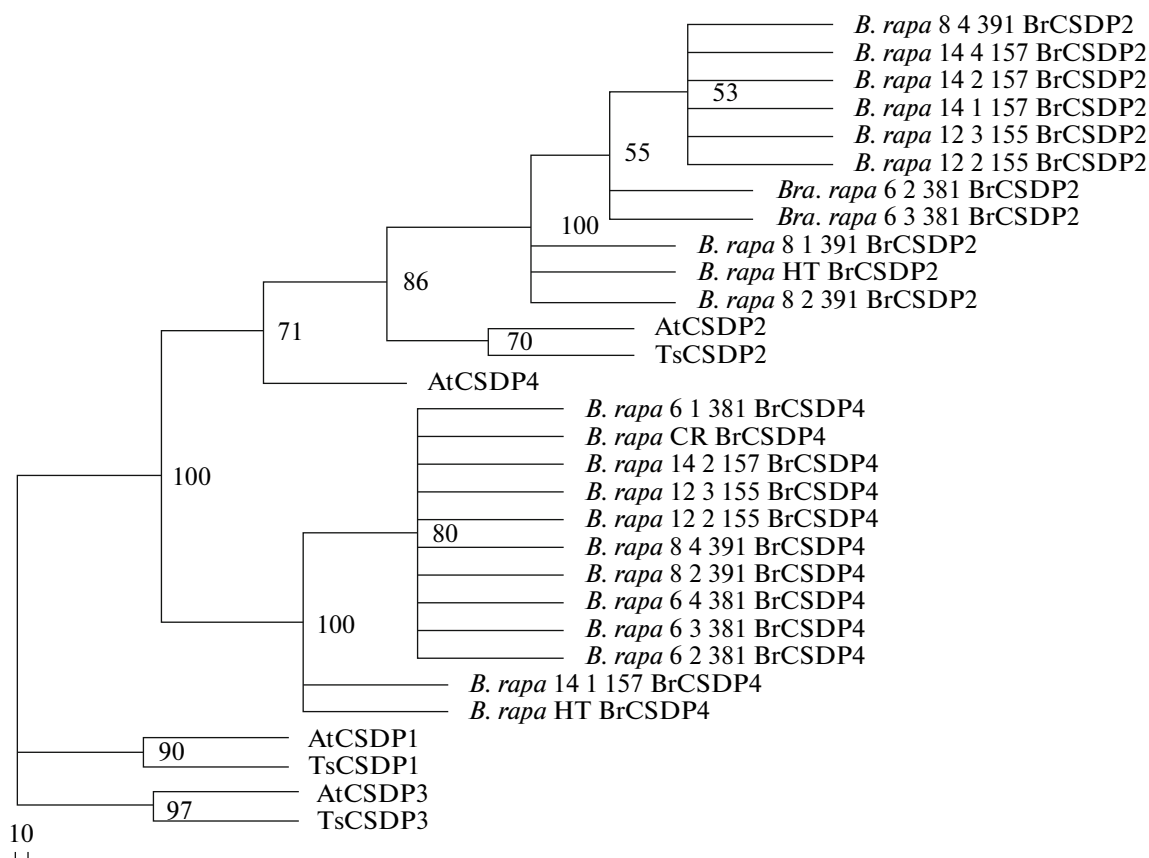


Рис. 2. Дендрограмма, отражающая сходство нуклеотидных последовательностей генов *CSDP* *B. rapa*, *A. thaliana* и *T. salsuginea*, полученная с использованием метода максимальной экономии (MP, Maximum Parsimony). В узлах ветвей указаны значения процента бутстрепа (Bootstrap Percentage, BP).

ющей мотивам ZnF. Однако при включении L47853 в филогенетический анализ аминокислотных последовательностей CSD отмечено наибольшее сходство этого участка именно с последовательностями гена *BrCSDP2*. Помимо нескольких специфических замен L47853 содержала все SNP/аминокислотные замены, специфичные для группы генов *CSDP2* Brassicaceae (рис. 1).

Таким образом, в результате проведенной работы были разработаны праймеры для дифференциальной амплификации генов *B. rapa*, кодирующих белки с CSD-доменом. С использованием этих праймеров определены полные нуклеотидные последовательности двух типов генов (*BrCSDP2* и *BrCSDP4*) из шести образцов *B. rapa*. Идентифицированные гены принадлежат к группе коротких генов *AtCSP2/AtCSP4* *A. thaliana* и *TsCSDP2/TsCSDP4* *T. salsuginea*, кодирующих CSD-домен и два ZnF-мотива. Охарактеризована структура и полиморфизм идентифицированных генов, выявлены и описаны их аллели. Не выявлено корреляции этих аллелей с признаками устойчивости к холодному стрессу, что может указывать на сложную генетическую природу

формирования адаптивности к холодному стрессу у *B. rapa*.

Работа выполнена при финансовой поддержке Минобрнауки (ГК №16.512.11.2157) и программы Президиума Российской академии наук «Молекулярная и клеточная биология».

СПИСОК ЛИТЕРАТУРЫ

1. Horn X.G., Hofweber R., Kremer W., Kalbitzer H.R. 2007. Structure and function of bacterial cold shock proteins. *Mol. Cell. Life Sci.* **64**, 1457–1470.
2. Скабкин М.А., Скабкина О.В., Овчинников Л.П. 2004. Мультифункциональные белки с доменом холодного шока в регуляции экспрессии генов. *Успехи биол. химии.* **44**, 3–52.
3. Chaikam V., Karlson D.T. 2010. Comparison of structure, function and regulation of plant cold shock domain proteins to bacterial and animal cold shock domain proteins. *Biochem. Mol. Biol. Reports.* **43**, 1–8.
4. Fusaro A.F., Vocca S.N., Ramos R.L., et al. 2007. AtGRP2, a cold induced nucleocytoplasmic RNA-binding protein, has a role in flower and seed development. *Planta.* **225**, 1339–1351.

5. Sasaki K., Kim M.H., Imai R. 2007. *Arabidopsis* cold shock domain protein 2 is a RNA chaperone that is regulated by cold and developmental signals. *Biochem. Biophys. Res. Commun.* **364**, 633–638.
6. Nakaminami K., Hill K., Perry S.E., et al. 2009. *Arabidopsis* cold shock domain proteins: relationships to floral and silique development. *J. Exp. Bot.* **60**, 1047–1062.
7. Park S.J., Kwak K.J., Oh T.R., et al. 2009. Cold shock domain proteins affect seed germination and growth of *Arabidopsis thaliana* under abiotic stress conditions. *Plant Cell Physiol.* **50**, 869–878.
8. Yang Y., Karlson D.T. 2011. Overexpression of AtCSP4 affects late stages of embryo development in *Arabidopsis*. *J. Exp. Bot.* **62**(6), 2079–2091.
9. Sasaki K., Imai R. 2012. Pleiotropic roles of cold shock domain proteins in plants. *Frontiers Plant Sci./Plant Genet. Genomics.* **2**, 1–6.
10. Karlson D., Imai R. 2003. Conservation of the cold shock domain protein family in plant. *Plant Physiol.* **131**, 12–15.
11. Karlson D., Nakaminami K., Toyomasu T., Imai R. 2002. A cold regulated nucleic acid binding protein of winter wheat shares a domain with bacterial cold shock proteins. *J. Biol. Chem.* **277**, 35248–35356.
12. Nakaminami K., Karlson D., Imai R. 2006. Functional conservation of cold shock domains in bacteria and higher plants. *Proc. Natl. Acad. Sci. USA.* **103**, 10122–10127.
13. Таранов В.В., Бердникова М.В., Носов А.В., Галкин А.В., Бабаков А.В. 2010. Белки с доменом холодового шока в растении – экстреморфите *Thellungiella salsuginea*: структура генов и их дифференцированная экспрессия при холодовой адаптации. *Молекулярная биология.* **44** (5), 889–897.
14. Edwards S.K., Johnston C., Thompson C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analyses. *Nucl. Acids Res.* **19**(6), 1349.
15. Zhang Z., Schwartz S., Wagner L., Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7**(12), 203–214.
16. Tamura K., Dudley J., Nei M., Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599.
17. Swofford D.L. 2002. PAUP* Phylogenetic Analysis Using Parsimony (*and other methods), v. 4.0 beta 10. Sinauer Associates, Sunderland (MA).
18. Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution.* **39**, 783–791.
19. Couvreur T.L. P., Franzke A., Al-Shehbaz I.A., et al. 2010. Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae). *Mol. Biol. Evol.* **27**(1), 55–71.
20. Rose L.E., Bittner-Eddy P.D., Langley C.H., et al. 2004. The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. *Genetics.* **166**, 1517–1527.
21. Nieto C., Morales M., Orjeda G., et al. 2006. An eIF4E allele confers resistance to an uncapped and non-polyadenylated RNA virus in melon. *Plant J.* **48**, 452–462.
22. Kobayashi Y., Kuroda K., Kimura K., et al. 2008. Amino acid polymorphisms in strictly conserved domains of a P-type ATPase HMA5 are involved in the mechanism of copper tolerance variation in *Arabidopsis*. *Plant Physiol.* **148**, 969–980.
23. Wang N., Qian W., Suppanz I., et al. Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the *FRIGIDA* homologue *BnaA.FRI.a*. *J. Exp. Bot.* **62** (15), 5641–5658.
24. Caicedo A.L., Schaal B.A. 2004. Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Mol. Ecol.* **13**, 1871–1882.
25. Traore K., McClung A.M., Chen M.-H., Fjellstrom R. 2011. Inheritance of flour paste viscosity is associated with a rice *Waxy* gene exon 10 SNP marker. *J. Cereal Sci.* **53**, 37–44.
26. Vezzulli S., Micheletti D., Riaz S., et al. 2008. A SNP transferability survey within the genus *Vitis*. *MC Plant Biol.* **8**, 128, doi: 10.1186/1471-2229-8-128.
27. Yang C., Wang L., Siva V.S., et al. 2012. A novel cold-regulated cold shock domain containing protein from scallop *Chlamys farreri* with nucleic acid-binding activity. *PLoS ONE.* **7**(2): e32012. doi:10.1371/journal.pone.0032012.
28. Lim C.O., Kim H.Y., Kim M.G., et al. 1996. Expressed sequence tags of Chinese cabbage flower bud cDNA. *Plant Physiol.* **111**(2), 577–588.