

ПОЛИМОРФИЗМ ГЕНА *SQS*, КОДИРУЮЩЕГО СКВАЛЕН-СИНТАЗУ У ВИДОВ АМАРАНТА (*Amaranthus* L.)

© 2020 г. А. Б. Щербань¹, *, А. И. Стасюк¹

¹Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

*e-mail: atos@bionet.nsc.ru

Поступила в редакцию 11.03.2019 г.

После доработки 07.05.2019 г.

Принята к публикации 24.06.2019 г.

Статья посвящена анализу полиморфизма гена, кодирующего сквален-синтазу (*SQS*) у видов амаранта зернового и овощного направлений селекции, а также у их дикорастущих предшественников. Структурный анализ кодирующей части гена у различных видов амаранта выявил низкий уровень полиморфизма и консерватизм основных функциональных доменов. Особое внимание уделено структурному полиморфизму промоторного района гена как наиболее вариабельному. Выявлено два основных гаплотипа гена по структуре этого района: С и Р; при этом вариант С оказался тесно сцепленным с признаком желтой окраски зерна у образцов зернового амаранта. Обсуждается возможное влияние структуры промотора *SQS*-гена на уровень его экспрессии и концентрацию сквалена в зерне.

Ключевые слова: амарант, сквален, сквален-синтаза, ген, первичная структура, полиморфизм.

DOI: 10.31857/S0016675820030145

Амарант (род *Amaranthus*) относится к двудольным травянистым растениям семейства Амарантовые (*Amaranthaceae*) и включает около 75 видов, произрастающих в теплых и умеренных областях земного шара [1]. В настоящее время наиболее общепринято деление рода *Amaranthus* на три подрода: *Amaranthus*, *Albersia* и *Acnida* [2]. Подрод *Amaranthus* содержит три вида: *A. caudatus* L., *A. cruentus* L. и *A. hypochondriacus* L., происходящих из Южной и Центральной Америки. В странах указанного региона эти виды использовались с древнейших времен, в основном как зерновые культуры, лишь немного уступая по своему значению кукурузе и бобовым. В России амарант долгое время не считали культурным растением, так как у нас в основном были распространены его дикие формы. Тем не менее, на необходимость применения амаранта в сельском хозяйстве (как новой кормовой культуры) указывал академик Н.И. Вавилов еще в 1932 г. и им же во Всесоюзном институте растениеводства была собрана первая коллекция амаранта. После ареста Вавилова, начатая по его инициативе исследовательская работа с амарантом была прекращена. Возрождение интереса к амаранту в конце XX в. связано с изучением механизма C4-фотосинтеза, который присущ амаранту как представителю “аспартатной” подгруппы C4-растений, а также с его уникальными пищевыми и биохимическими

свойствами и поистине универсальным потенциалом в качестве зерновой, овощной и кормовой культуры [3, 4]. Масло семян амаранта является богатым источником сквалена — сложного углеводорода, предшественника стероидных соединений. Сквален имеет широкое применение в медицине: в качестве адьюванта в вакцинах, иммуномодулятора и антиоксиданта в комплексной терапии ряда заболеваний: таких как диабет, ишемическая болезнь и др., а также в составе косметических средств [5, 6]. Его содержание в масле амаранта составляет 2,2–10%, что значительно превышает содержание сквалена в оливковом и других растительных маслах [1, 7]. До сих пор сквален получают главным образом из печени акул и китов, поэтому масло амаранта может служить альтернативным сырьевым источником этого ценного вещества.

Повышенная продукция сквалена у амарантовых, по сравнению с другими растениями, может быть следствием мутационных изменений в кодирующих или регуляторных районах генов, контролирующих биосинтез этого вещества. В настоящее время установлена ключевая роль *SQS*-гена в биосинтезе сквалена [8]. Этот ген кодирует фермент сквален-синтазу, который катализирует последний этап синтеза — образование молекулы сквалена из двух молекул фарнезил-пирофосфата [9].

Таблица 1. Распределение P- и C-гаплотипов промоторного района *SQS*-гена среди проанализированных образцов амаранта

Вид	Каталожный номер образца*	Происхождение	Тип промотора**	Окраска зерна
<i>A. caudatus</i> L.	11001 (к-40198)	Таджикистан, местный	Pm	Розовый
	11002 (к-вр.191)	Бурунди, местный	Pm	Розовый
	11003 (к-27152)	Россия, Красноярский р-н	Pm	Розовый
	11015 (к-67)	Гана, местный	P	Коричневый
	11023 (к-65)	Польша, местный	Pm	Черный
	11047	Украина	Pm	Розовый
	11049	Россия, Башкортарстан	P	Светло-коричневый
	11064 (к-вр.231)	Венгрия	P	Розовый
	11079	Россия	Pm	Светло-коричневый
<i>A. edulis</i> Spegazzini	11030 (к-70)	Гана, местный	P	Черный
<i>A. edulis</i> L.	11070 (к-36)	Аргентина	C	Желтый
	11072 (к-45913)	Россия, сорт "Эльбрус"	C	Желтый
<i>A. cruentus</i> L.	11006 (к-50)	Румыния	P	Черный
	11008 (к-71)	Непал	C	Желтый
	11011 (к-64)	Канада	C	Желтый
	11014 (к-26)	Казахстан, местный	P	Черный
	11017 (к-78)	Западный Камерун	P	Черный
	11018	Китай	C	Желтый
	11019 (к-21)	Китай	P	Черный
	11073 (к-9)	Индия	P	Черный
	11074 (к-12)	Индия	P	Черный
	11075	Россия, Томск	P	Черный
	11080	Россия	P	Черный
<i>A. hypochondriacus</i> L.	11010 (к-61)	США	C	Желтый
	11022 (к-51)	Германия, местный	P	Черный
	11024 (к-вр.127)	Ямайка	C	Желтый
	11071 (к-4)	Китай	C	Желтый
<i>A. hybridus</i> L.	11083 (к-28024)	Германия	P	Черный
	11088	Россия, Томск	P	Черный
<i>A. leucospermus</i> L.	11058	Россия, Новосибирск, ВАСХНИЛ	C	Желтый
	11059	Россия, Новосибирск	P	Желтый
<i>A. mangostanus</i> L.	11009 (к-28)	Китай	P	Коричневый
<i>A. spinosus</i> L.	11013 (к-77)	Германия, дикораст.	P	Черный
<i>A. lividus</i> L.	11036 (к-57)	Румыния, местный	P	Черный
	11037	Россия, Томск	P	Черный
<i>A. mantegazzianus</i> Passerini	11028 (к-69)	Австралия, дикораст.	P	Черный

Таблица 1. Окончание

Вид	Каталожный номер образца*	Происхождение	Тип промотора**	Окраска зерна
<i>A. aureus</i> L.	11055 (к-вр.160)	Германия	P	Черный
	11093	Молдавия	P	Черный
<i>A. deflexus</i> L.	11057 (к-вр.106)	Германия	P	Черный
<i>A. erythrostachys</i> L.	11060	Россия, Москва	P	Черный
<i>A. nobilis</i> L.	11062	Россия, Томск	P	Черный
<i>A. powellis</i> L.	11065 (к-62)	Греция	P	Черный
	11066 (к-59)	Чехия	P	Черный
<i>A. tricolor</i> L.	11089	Румыния	P	Черный
Не определен	11007 (к-53)	Западный Камерун	P	Черный
	11029 (к-вр.190)	Италия	P	Темно-коричневый
	11034 (к-вр.170)	Мексика, местный	C	Желтый
	11035 (к-вр.116)	Гана, местный	P	Черный
	11046 (к-10)	Индия, местный	P	Черный
	11050 (к-42)	Танзания, местный	P	Темно-коричневый
	11051 (к-43)	Танзания, местный	P	Черный
	11052 (к-46)	Мали, местный	C	Черный
	11054 (к-44)	Индия, местный	P	Черный
	11067 (к-вр.72)	Узбекистан, местный	P	Черный
	11068 (к-вр.155)	Заир, местный	P	Черный
	11077 (к-вр.29)	Индия, местный	P	Черный
	11081 (к-14)	Индия, местный	P	Черный
	11082 (к-17)	Индия, местный	P	Черный
	11084 (к-вр.147)	Конго, местный	P	Черный
11090	США	P	Черный	

Примечание. Данные по окраске зерна взяты из монографии [1]. Случаи ассоциации желтой окраски зерна с С-вариантом промотора *SQS*-гена выделены серым фоном. Случаи несоответствия между черной и желтой окрасками зерна и двумя вариантами промотора выделены жирным шрифтом.

* Номера образцов коллекции ИЦиГ СО РАН (в скобках соответствующие номера из каталога ВИР).

** С – “сruentus-подобный” тип; P – преобладающий тип; Pm – модифицированный P-гаплотип содержит трек (TAA)₃ вместо (TAA)₂ вблизи предполагаемого TATA-бокса.

Анализ первичной структуры

Аmplифицированные фрагменты ДНК секвенировали с использованием ABI PRISM Dye Terminator Cycle Sequencing ready reaction kit (Perkin Elmer Cetus, USA) и соответствующих специфических праймеров. Секвенирование проводили на базе Центра коллективного пользования “Геномика” СО РАН (Новосибирск, <http://sequenst.niboch.nsc.ru>). Для сравнения полученных нуклеотидных и аминокислотных последовательностей использовали программу Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Поиск предполагаемых цис-регуляторных элемен-

тов в составе промоторных районов проводили с помощью базы данных PlantPAN 2.0 (<http://plantpan2.itps.ncku.edu.tw>). Полученные промоторные и кодирующие последовательности депонированы в ГенБанке NCBI (<https://www.ncbi.nlm.nih.gov/>) под номерами: МК587531–МК587574 и МК598767–МК598768 соответственно.

Маркирование *SQS*-гена по структуре промоторного района

Для идентификации двух различных типов промоторного района *SQS*-гена в образцах амаранта были использованы две комбинации праймеров.

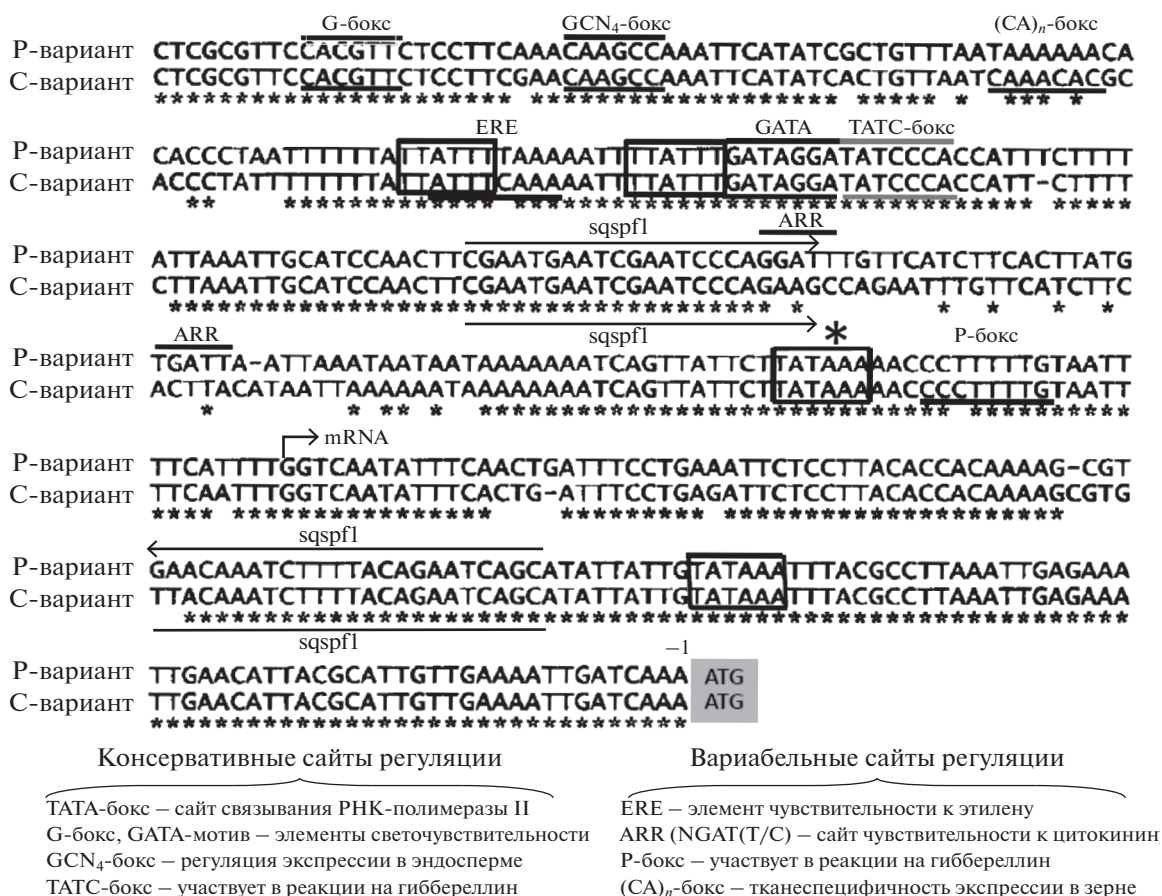


Рис. 2. Сравнение нуклеотидных последовательностей промоторного района *SQS*-гена *A. cruentus* (11019; Р-вариант) и *A. hypochondriacus* (11010; С-вариант). Линиями подчеркнуты предполагаемые цис-регуляторные сайты (внизу дана их краткая характеристика). АТГ-кодон, или сайт инициации трансляции выделен серым фоном. Предполагаемые ТАТА-боксы выделены рамками (звездочкой отмечен наиболее вероятный). Места отжига праймеров для ПЦР-маркирования вариантов промотора обозначены прямыми стрелками.

Р-тип: sqspfl (5'-cgaatgaatcgaatcccaggat-3')/sqsprl (5'-gctgattctgtaaagattgttc-3'). С-тип: sqscfl (5'-cgaatgaatcgaatcccagaag-3')/sqscr1 (5'-gctgattctgtaaagattgtaac-3'). Размер ПЦР-продукта в обоих случаях ~180 пн (рис. 2). Для амплификации использована методика "Touch Down" ПЦР: 2 мин при 94°C; 15 циклов (94°C, 30 с; 65°C, 30 с; 72°C, 30 с); 25 циклов (94°C, 30 с; 55°C, 30 с; 72°C, 30 с); 2 мин при 72°C.

РЕЗУЛЬТАТЫ

Ранее на основе анализа кДНК была проанализирована первичная структура *SQS*-гена в геноме *A. cruentus* и показано, что этот ген присутствует в одной копии [10]. Его кодирующая часть вместе с интронами имеет длину 6490 пн и включает 13 экзонов (рис. 1). С помощью BLAST-поиска в составе референсной полной геномной последовательности вида *A. hypochondriacus* (см. Методы) нами была установлена аналогичная последовательность *SQS*-гена вместе с 5'-прилегающим протя-

женным районом. Степень ее гомологии с предыдущей по аминокислотной последовательности составляет 98%. Для анализа полиморфизма *SQS*-гена у различных видов амаранта, нами были разработаны специфические праймеры к промоторному (~0.5 тпн от АТГ-кодона) и кодирующему районам. Для частичного анализа последнего был выбран участок между экзонами 5 и 9, содержащий три важнейших функциональных домена (рис. 1).

Промоторный район

С использованием специфических праймеров sqsf2/sqsr3 к промоторному району *SQS*-гена у всех проанализированных образцов амаранта был обнаружен ПЦР-фрагмент длиной примерно 600 пн (результат не представлен). Секвенирование выделенных ПЦР-продуктов показало, что полиморфизм в данном районе гена ограничен двумя основными структурными вариантами, обозначенными как: С-гаплотип (или *cruentus*-

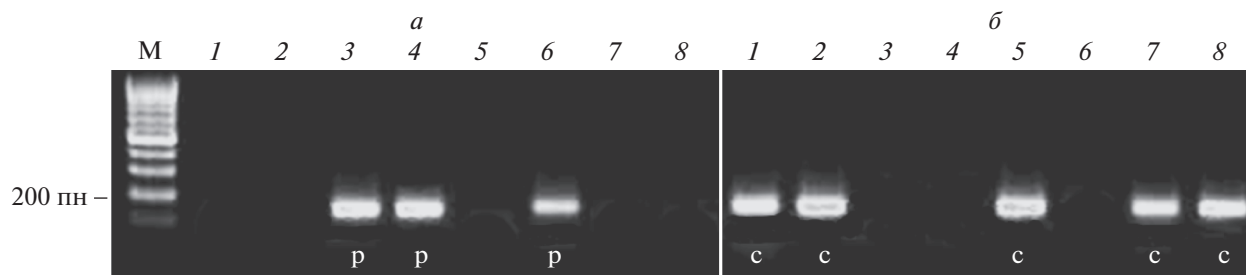


Рис. 3. ПЦР с праймерами sqspfl/sqspr1 (а) и sqscf1/sqscr1 (б). 1–3 – образцы *A. cruentus* 11008, 11011, 11019; 4–5 – *A. hypochondriacus* 11022, 11010; б – *A. caudatus* 11001; 7 – *A. edulis* 11072; 8 – *A. leucospermus* 11058. Внизу указан тип промотора.

подобный) и Р-гаплотип (преобладающий тип). Уровень гомологии между ними составляет 92%. Внутри каждого типа полиморфизм незначительный и как правило не превышает 1%. С-тип распространен в ограниченном числе видов, главным образом среди культурных образцов зернового направления: *A. cruentus*, *A. hypochondriacus*, *A. leucospermus* L., *A. edulis* L. (син. *caudatus*). Р-гаплотип является наиболее распространенным и встречается как среди образцов зерновых видов: *A. caudatus*, *A. cruentus* (часть образцов) и др., у представителей овощного направления: *A. lividus*, *A. tricolor*, а также у дикорастущих видов: *A. hybridus*, *A. powellii*, *A. nobilis*, *A. spinosus* (табл. 1).

Сравнение первичной структуры С- и Р-типов промотора *SQS*-гена выявило наряду с предполагаемыми консервативными сайтами, такими как: ТАТА-бокс, G-бокс и GATA-мотив (чувствительность к свету), GCN₄-бокс (экспрессия в эндосперме), также и ряд сайтов, варьирующих по структуре между этими типами и связанных с действием гормонов: этилена, цитокинина и гиббериллина (рис. 2). Данные варьирующие сайты относятся к потенциальным цис-регуляторным элементам, способным обуславливать межвидовые различия в уровне транскрипции *SQS*-гена в зависимости от ткани и стадии развития растения.

Выявленные структурные особенности С- и Р-гаплотипов промоторного района позволили разработать специфические праймеры для их маркирования и быстрого ПЦР-скрининга образцов амаранта (см. Методы). Комбинации праймеров, несущих SNP на 3'-концах ограничивают один и тот же участок промоторного района длиной около 180 пн (рис. 2). Обе комбинации были успешно апробированы на всей выборке образцов, что подтвердило предыдущие данные секвенирования (на рис. 3 представлен результат для восьми образцов).

Кодирующий район

Для анализа первичной структуры кодирующего района *SQS*-гена, расположенного между экзонами 5–9 (см. Методы) были взяты два образца амаранта: *A. cruentus* (Асс № 11019) и *A. hypochondriacus* (Асс № 11010), имеющие Р- и С-гаплотипы промоторного района соответственно. Сравнение полученных нуклеотидных последовательностей с известными последовательностями *A. cruentus* (С-гаплотип; АВ691229) и *A. hypochondriacus* (Р-гаплотип; id40120) показало, что последовательности одного типа, но относящиеся к различным видам, имеют высокую гомологию – 99.7%, тогда как последовательности разных типов одного и того же вида характеризуются меньшей степенью гомологии – 96.1%. Почти все различия находятся в составе интронов; при этом преобладают SNP (рис. 4). Самые крупные делеции выявлены в 5-м и 7-м интронах (8 и 12 пн соответственно). Кодирующие последовательности у проанализированных С- и Р-вариантов *SQS*-гена содержат шесть синонимичных замен и две несинонимичные; при этом последние не затрагивают консервативные функциональные домены II–IV (рис. 4).

ОБСУЖДЕНИЕ

Амарант является богатым источником многих ценных питательных веществ, в том числе сквалена [4]. Важную роль в регуляции биосинтеза этого вещества выполняет ген сквален-синтазы, который был достаточно подробно изучен у целого ряда растительных объектов [8, 11–16]. Недавно у амаранта *A. cruentus* была проанализирована структура *SQS*-гена и его экспрессия в различных тканях [10]. С помощью BLAST-поиска нами была установлена другая последовательность этого гена в составе референсного генома *A. hypochondriacus* [17, 18]. Эти последовательности оказались высокоомологичными в пределах всего кодирующего района, включающего 13 эк-

sqsl1019	60	Ex5	60	839
sqsl1010	60		60	832
sqsl1019	120		120	887
sqsl1010	112		112	892
sqsl1019	180		180	947
sqsl1010	172		172	952
sqsl1019	240		240	1007
sqsl1010	232		232	1012
sqsl1019	300		300	1067
sqsl1010	292	Ex6	292	1072
sqsl1019	360		360	1127
sqsl1010	352		352	1132
sqsl1019	420		420	1187
sqsl1010	412		412	1192
sqsl1019	480		480	1247
sqsl1010	472		472	1252
sqsl1019	540		540	1307
sqsl1010	532	Ex7	532	1312
sqsl1019	600		600	1366
sqsl1010	592		592	1372
sqsl1019	660		660	1426
sqsl1010	652		652	1432
sqsl1019	720		720	Ex9
sqsl1010	712		712	1479
sqsl1019	779		779	1484
sqsl1010	772		772	

Рис. 4. Сравнение нуклеотидных последовательностей части кодирующего района *SQS*-гена *A. cruentus* (sqsl1019) и *A. hypochondriacus* (sqsl1010). Экзоны 5–9 выделены серым фоном. Соответствующие экзонам аминокислотные последовательности даны под нуклеотидными. Основные функциональные домены II–IV выделены серым фоном. Подчеркнуты кодоны, в которых произошли замены. Прямым шрифтом обозначены кодоны с синонимичными заменами, курсивом – с несинонимичными (соответствующая замена аминокислоты указана сверху).

зонов (рис. 1). Однако достаточно протяженная структура промоторного района была известна только у последнего вида. Поэтому мы поставили цель – изучить полиморфизм *SQS*-гена у различных видов амаранта, уделяя особое внимание промотору, как основному регуляторному району, от которого может зависеть уровень экспрессии данного гена. В качестве материала была использована коллекция образцов, большая часть которых была предоставлена ВИР. На протяжении ряда лет эту коллекцию размножали на опытных полях ИЦиГ СО РАН самоопылением, чтобы исключить перекрестную гибридизацию, характерную для амаранта [1].

Вначале, с использованием разработанных праймеров нами было проведено выделение из геномной ДНК образцов и последующее секвенирование фрагментов ДНК в пределах ~0.5 тпн выше старта трансляции *SQS*-гена (рис. 1). Выявленный полиморфизм этого района был ограничен двумя основными вариантами, или гаплотипами, которые были обозначены как С и Р (рис. 2). Последний тип преобладает среди образцов амаранта различного происхождения (как культурных, так и дикорастущих), тогда как С-гаплотип встречается, главным образом, у группы видов зернового амаранта: *A. hypochondriacus*, *A. cruentus*, *A. edulis*, *A. leucospermus* (табл. 1). При этом Р-гаплотип у этих видов также присутствует (у местных дикорастущих форм). Следует отметить, что два последних вида близки к *A. caudatus* и иногда рассматриваются как его подвиды (<http://www.theplantlist.org/tpl1.1/record/kew-2632762>).

Исследованные образцы собственно вида *A. caudatus* представлены либо Р-гаплотипом, либо его слабо модифицированной формой Р_м, отличающейся вставкой ТАА-тринуклеотида в промоторном районе. На основании ранее проведенной оценки филогении рода [2, 19] виды *A. hypochondriacus* и *A. cruentus*, с одной стороны, и *A. caudatus*, с другой, имеют разное происхождение и различающиеся первичные ареалы распространения (Центральная и Южная Америка соответственно). Следовательно можно предположить, что С-гаплотип впервые появился до расхождения обеих географических групп зернового амаранта. В этой связи интересен факт отсутствия переходных структурных вариантов между Р-типом, как более древним, и типом С, при наличии большого количества независимых мутаций в промоторном районе (рис. 2). Возникает вопрос: имели ли эти мутации какое-либо селективное значение? На связь с процессом доместикации указывает установленный нами факт тесного сцепления С-гаплотипа с желтой окрас-

кой зерна (табл. 1). Только два образца являются исключением: у первого (11059) желтая окраска связана с Р-гаплотипом, у второго (11052) наблюдалось сочетание черной окраски с типом С. Как известно, пригодное в пищу зерно амаранта имеет светлую (желтую, белую или розовую, как у *A. caudatus*) окраску, тогда как зерно черного или темно-коричневого цвета не имеет пищевой ценности из-за высокой концентрации токсинов (<http://www.echonet.org>).

Наряду с этим наличие изменений в потенциальных регуляторных сайтах, связанных с чувствительностью к ряду гормонов, таких как цитокинин, этилен, гиббериллин (рис. 2), указывает на то, что эти изменения в совокупности могли привести к существенной вариации уровня экспрессии гена и соответствующей вариации содержания конечного продукта – сквалена, в зерновой ткани амаранта. Ранее было установлено, что культурные формы зернового амаранта имеют повышенное содержание сквалена относительно дикорастущих форм. Так например у образца 11010 *A. hypochondriacus*, из которого в ИЦиГ СО РАН был получен сорт “Янтарь”, содержание сквалена достигает 9.7%, что примерно в два раза превышает среднее содержание этого вещества у амаранта [1, 7]. Связано ли это превышение с С-гаплотипом промотора *SQS*-гена, и каков механизм регуляции его экспрессии – покажут дальнейшие исследования. Практический аспект работы заключается в разработке молекулярных маркеров для быстрой и эффективной селекции образцов амаранта, несущих различные гаплотипы *SQS*-гена (рис. 3).

Кодирующая часть *SQS*-гена включает пять функциональных доменов, структура которых консервативна у разных групп растений [10]. Наибольшая степень консерватизма характерна для доменов II, III и IV, которые, как предполагают, формируют активный центр белка и играют ключевую роль в реакции, катализируемой сквален-синтазой [16]. Поэтому для оценки полиморфизма мы выбрали район длиной около 1.5 тпн, включающий указанные домены (рис. 1). Анализ его первичной структуры у образцов видов *A. cruentus* (11019) и *A. hypochondriacus* (11010), имеющих альтернативные типы промотора, показал низкий уровень полиморфизма, за исключением интронов, в которых был обнаружен ряд SNP и две делеции (рис. 4). Изменения в составе экзонов включают восемь замен, из которых только два относятся к несинонимичным, но эти замены находятся за пределами основных функциональных доменов II–IV.

Таким образом, анализ полиморфизма *SQS*-гена у культурных и дикорастущих форм амаранта выявил поддержание консервативной структуры кодирующей части, отвечающей за функции белка при наличии выраженных изменений в промоторном районе гена (С-гаплотип), которые могли иметь селективное значение при отборе форм зернового амаранта с высокими питательными качествами зерна, в частности с повышенной концентрацией сквалена. Планируемый анализ экспрессии *SQS*-гена и содержания сквалена в зерне у образцов амаранта с различными типами промотора позволит выявить связь между структурой промотора, уровнем экспрессии гена и концентрацией сквалена.

Авторы выражают благодарность Н.Б. Железновой за любезно предоставленные семена амаранта.

Работа финансировалась за счет проекта Минобрнауки РФ № 0324-2019-0039.

СПИСОК ЛИТЕРАТУРЫ

1. Железнов А.В., Железнова Н.Б., Бурмакина Н.В., Юдина Р.С. Амарант: научные основы интродукции. М.: Академ. изд-во "Гео", 2009. 235 с.
2. Stetter M.G., Schmid K.J. Phylogenetic relationships and genome size evolution within the genus *Amaranthus* indicate the ancestors of an ancient crop // *Mol. Phylogenet. and Evol.* 2017. V. 109. P. 80–92. <https://doi.org/10.1016/j.ympev.2016.12.029>
3. Магомедов И.М., Чиркова Т.В. Амарант – прошлое, настоящее и будущее // *Усп. современ. естествознания (сер. Биол. науки)*. 2015. № 1. С. 1108–1113.
4. Venskutonis P.R., Kraujalis P. Nutritional components of amaranth seeds and vegetables: a review on composition, properties, and uses // *Comprehensive Rev. in Food Sci. and Food Safety*. 2013. V. 12. P. 381–412. <https://doi.org/10.1111/1541-4337.12021>
5. Huang Z.R., Lin Y.K., Fang J.Y. Biological and pharmacological activities of squalene and related compounds: Potential uses in cosmetic dermatology // *Molecules*. 2009. V. 14. P. 540–554. <https://doi.org/10.3390/molecules14010540>
6. Caselato-Sousa V.M., Amaya-Farf'an J. State of knowledge on amaranth grain: a comprehensive review // *J. of Food Sci.* 2012. V. 77. № 4. P. 93–104. <https://doi.org/10.1111/j.1750-3841.2012.02645.x>
7. He H.P., Cai Y.Z., Sun M., Corke H. Extraction and purification of squalene from *Amaranthus* grain // *J. Agric. Food Chem.* 2002. V. 50. P. 368–372. <https://doi.org/10.1021/jf010918p>
8. Nakashima T., Inoue T., Oka A. et al. Cloning, expression, and characterization of cDNAs encoding *Arabidopsis thaliana* squalene synthase // *Proc. Natl Acad. Sci. USA*. 1995. V. 92(6). P. 2328–2332. <https://doi.org/10.1073/pnas.92.6.2328>
9. Zurer P. Biosynthesis of squalene // *Chem. Engineering News*. 1997. V. 75(1). P. 6–7. <https://doi.org/10.1021/cen-v075n001.p006>
10. Park Y.-J., Nemoto K., Minami M., Matsushima K. Molecular cloning, expression and characterization of a squalene synthase gene from grain amaranth (*Amaranthus cruentus* L.) // *JARQ*. 2016. V. 50(4). P. 307–317. <https://doi.org/10.6090/jarq.50.307>
11. Hata S., Sanmiya K., Kouchi H. et al. cDNA cloning of squalene synthase genes from mono- and dicotyledonous plants, and expression of the gene in rice // *Plant and Cell Physiol.* 1997. V. 38(12). P. 1409–14013. <https://doi.org/10.1093/oxfordjournals.pcp.a029137>
12. Devarenne T.P., Shin D.H., Back K. et al. Molecular characterization of tobacco squalene synthase and regulation in response to fungal elicitor // *Arch. Biochem. Biophys.* 1998. V. 349(2). P. 205–215. <https://doi.org/10.1006/abbi.1997.0463>
13. Akamine S., Nakamori K., Chechetka S.A. et al. cDNA cloning, mRNA expression, and mutational analysis of the squalene synthase gene of *Lotus japonicas* // *Biochimica and Biophysica Acta*. 2003. V. 1626. P. 97–101. [https://doi.org/10.1016/S0167-4781\(03\)00042-3](https://doi.org/10.1016/S0167-4781(03)00042-3)
14. Huang Z., Jiang K., Pi Y. et al. Molecular cloning and characterization of the yew gene encoding squalene synthase from *Taxus cuspidate* // *J. Biochem. Mol. Biol.* 2007. V. 40. P. 625–635. <https://doi.org/10.5483/BMBRep.2007.40.5.625>
15. Uchida H., Yamashita H., Kajikawa M. et al. Cloning and characterization of a squalene synthase gene from a petroleum plant, *Euphorbia tirucalli* L. // *Planta*. 2009. V. 229. P. 1243–1252. <https://doi.org/10.1007/s00425-009-0906-6>
16. Kim T.D., Han J.-E., Huh G.H. et al. Expression and functional characterization of three squalene synthase genes associated with saponin biosynthesis in *Panax ginseng* // *Plant Cell Physiol.* 2011. V. 52. P. 125–137. <https://doi.org/10.1093/pcp/pcq179>
17. Clouse J.W., Adhikary D., Page J.T. et al. The amaranth genome: genome, transcriptome, and physical map assembly // *Plant Genome*. 2016. V. 9(1). P. 1–14. <https://doi.org/10.3835/plantgenome2015.07.0062>
18. Lightfoot D.J., Jarvis D.E., Ramaraj T. et al. Single-molecule sequencing and Hi-C-based proximity-guided assembly of amaranth (*Amaranthus hypochondriacus*) chromosomes provide insights into genome evolution // *BMC Biol.* 2017. V. 15:74. P. 1–15. <https://doi.org/10.1186/s12915-017-0412-4>
19. Wu X., Blair M.W. Diversity in grain amaranths and relatives distinguished by genotyping by sequencing (GBS) // *Frontiers in Plant Sci.* 2017. V. 8. P. 1–12. <https://doi.org/10.3389/fpls.2017.01960>

Polymorphism of the Squalene Synthase Gene (*SQS*) in Various Species of Amaranth (*Amaranthus* L.)

A. B. Shcherban^{a,*} and A. I. Stasyuk^a

^a*Federal Research Center Institute of Cytology and Genetics of Siberian Branch
of the Russian Academy of Sciences, Novosibirsk, 630090 Russia*

**e-mail: atos@bionet.nsc.ru*

The article is devoted to the analysis of polymorphism of the gene encoding squalene synthase (*SQS*) in species of amaranth of grain and vegetable directions of breeding, and their wild predecessors. Structural analysis of the coding part of the gene in different amaranth species revealed a low level of polymorphism and conservation of the main functional domains. Particular attention is paid to the structural polymorphism of the promoter region of the gene, as the most variable. According to the structure of this region, two main gene haplotypes were identified: C and P; while variant C was closely associated with yellow color of the grain in the accessions of grain amaranth. The possible influence of the structure of the *SQS*-gene promoter on the level of its expression and the concentration of squalene in the grain is discussed.

Keywords: amaranth, squalene, squalene synthase, gene, primary structure, polymorphism.