

УДК 575.113.1

**ИДЕНТИФИКАЦИЯ И ХАРАКТЕРИСТИКА СЕМЕЙСТВА ГЕНОВ MADS-BOX ПЕРЦА
*CAPSICUM ANNUUM L.*****А.В. Нежданова, А.М. Камионская, А.В. Щенникова**

Институт Биоинженерии им. К.Г. Скрябина, ФИЦ Биотехнологии РАН, Москва, Россия

Перец (род *Capsicum*, семейство Solanaceae) является одной из важнейших сельскохозяйственных культур (Naves et al., 2019). Известно, по меньшей мере, 32 вида перца, включая пять культивируемых *C. annuum*, *C. frutescens*, *C. chinense*, *C. baccatum* и *C. pubescens* (Qin et al. 2014). Современная селекция перцев направлена на улучшение характеристик урожая и все чаще применяет знания о молекулярно-генетических сетях, определяющих ценные признаки.

Контроль репродуктивного развития считается консервативным среди высших растений. Одними из ключевых регуляторов данного процесса является семейство MADS-доменных транскрипционных факторов (ТФ), которые, согласно моделям ABCDE и Квартет, в различных комбинациях и в составе мультимерных комплексов определяют идентичность органов цветка. (Parenicova et al., 2003; Theißen and Saedler, 2001). Семейство MADS объединяет белки типов I (ARG80/SRF-подобные M_α , M_β и M_γ) и II (MEF2-подобные $MIKC^C$ и $MIKC^*$) (Parenicova et al., 2003).

На сегодняшний день секвенированы геномы трёх культивируемых видов перца *C. annuum*, *C. chinense* и *C. baccatum*, а также одного дикорастущего образца *C. annuum* var. *glabriusculum* (Qin et al., 2014; Kim et al., 2017). Идентификация и сравнительный анализ ключевых генов полезных признаков в геномах видов перца даст возможность селекционерам не только предсказывать характеристики плодов, но и вести направленную селекцию сортов с улучшенными хозяйственно значимыми показателями.

Целью работы стала идентификация и структурно-филогенетический анализ семейства MADS-box генов перца *C. annuum*.

Семейство генов MADS-box перца было идентифицировано с помощью поиска в геноме *C. annuum* (PRJNA186921; ID: 1681888) базе данных NCBI (<https://www.ncbi.nlm.nih.gov>). С использованием программы NCBI-BLASTx и консервативного 5'-участка (180 п. н.) MADS-box генов, кодирующего домен MADS, в базе данных NR (Non-redundant protein sequences) были выявлены 234 кодирующие последовательности *C. annuum*. Выравнивание данных последовательностей между собой с помощью программы ClustalX позволило определить и удалить идентичные и неполные сиквенсы. В результате было идентифицировано 194 уникальные последовательности кДНК.

Для изучения экзон-интронной структуры каждого гена *C. annuum* gene ID кодирующей последовательности использовали для поиска в базе данных NCBI (All Databases). Далее, извлекали полногеномную последовательность (раздел Gene), включая данные о ее локализации. Координаты экзонов устанавливали с помощью MEGA 6.06 (<https://www.megasoftware.net/home>).

Было обнаружено, что, как и у модельного вида *Arabidopsis thaliana*, гены типа II ($MIKC^C$ и $MIKC^*$) перца содержат несколько интронов, в то время как гены типа I (M_α , M_β и M_γ) – только 1, редко 2 интрона, либо не имеют их вовсе. Исключение составили два гена M_α с числом интронов 7 и 11, соответственно. Известно, что уровень транскрипции генов типа II значительно превышает таковой у генов типа I. Также, гены типа II имеют более выраженные функции в сравнении с генами типа I. Поэтому мы предполагаем, что данный факт может быть связан с количеством интронов, которые могут содержать различные регуляторные элементы. В отличие от генов M_α , M_β и M_γ , гены $MIKC^C$ намного длиннее и содержат от трех до десяти экзонов. Гены $MIKC^*$ сходны по структуре с $MIKC^C$ (от 5 до 12 экзонов), однако не содержат последовательности, кодирующей домен K.

Все предполагаемые аминокислотные последовательности содержали MADS-домен.

Для сравнительного филогенетического анализа из базы данных NCBI были извлечены все 108 известных MADS-box генов модельного растения *A. thaliana* (Parenicova et al., 2003). Аминокислотные последовательности, кодируемые идентифицированными генами *A. thaliana*, были использованы для классификации MADS-семейства *C. annuum* методом максимального правдоподобия с применением отдельно подобранной модели замен в программе MEGA6.06.

Структурно-филогенетический анализ выявленных последовательностей в сравнении с *A. thaliana* разделил MADS-ТФ перца на два кластера, соответствующих ТФ типа I (SRF-подобные) и типа II

(MEF2-подобные). Кластер I состоял из трёх клад – M_{α} , M_{β} и M_{γ} (рис. 1). Кластер II – из двух основных клад, $MIKC^C$ и $MIKC^*$ (или M_{δ}) (рис. 2). В свою очередь, в $MIKC^C$ группе выделялись подклады APETALA1/FRUITFULL (A-активность; время цветения), APETALA3/PISTILLATA (B-активность), AGAMOUS/SEEDSTICK (C- и D-активности), SEPALLATA (SEP; E-активность), AGL24/SVP (время цветения), SOC1 (время цветения), FLC (время цветения) и AGL15 (развитие семян) (рис. 2). Каждая филогенетическая группа содержала в 3–4 раза больше паралофов, чем в случае *A. thaliana* (рис. 1, 2).

Поиск и сравнение консервативных мотивов в данных белках с помощью программы MEME 5.0.4 (<http://meme-suite.org/tools/meme>) позволил провести дополнительное структурное разделение MADS-гомологов на группы.

Всего было выявлено 84 консервативных мотива. Почти все MADS-гомологи перца содержали мотивы 1 и 2, которые вместе составляют MADS-домен, за исключением клады FLC и ряда белков типа I (отсутствие мотива или его замена другим консервативным мотивом). Мотивы 3–6 (103 а.о.) соответствуют I-области и K-домену и присутствуют практически у всех белков типа II, за некоторым исключением (клады FLC и SOC1). Структура белков типа I более разнообразна, чем у типа II. При этом они не имеют мотивов, соответствующих K-домену, что и отличает их от белков типа II, согласно (Parenicova et al., 2003). Полученные данные четко демонстрируют структурное сходство MADS-белков перца в пределах каждой группы, что полностью соответствует показанному делению белков на клады. Обнаружены уникальные консенсусы, характерные только для отдельных клад: 38 и 39 у AP3-подобных белков, 32 – у SVP/AGL24, 17 или 19 – у FLC, и др. Данные мотивы могут отражать специфичность функции конкретных белков в сравнении с белками других групп.

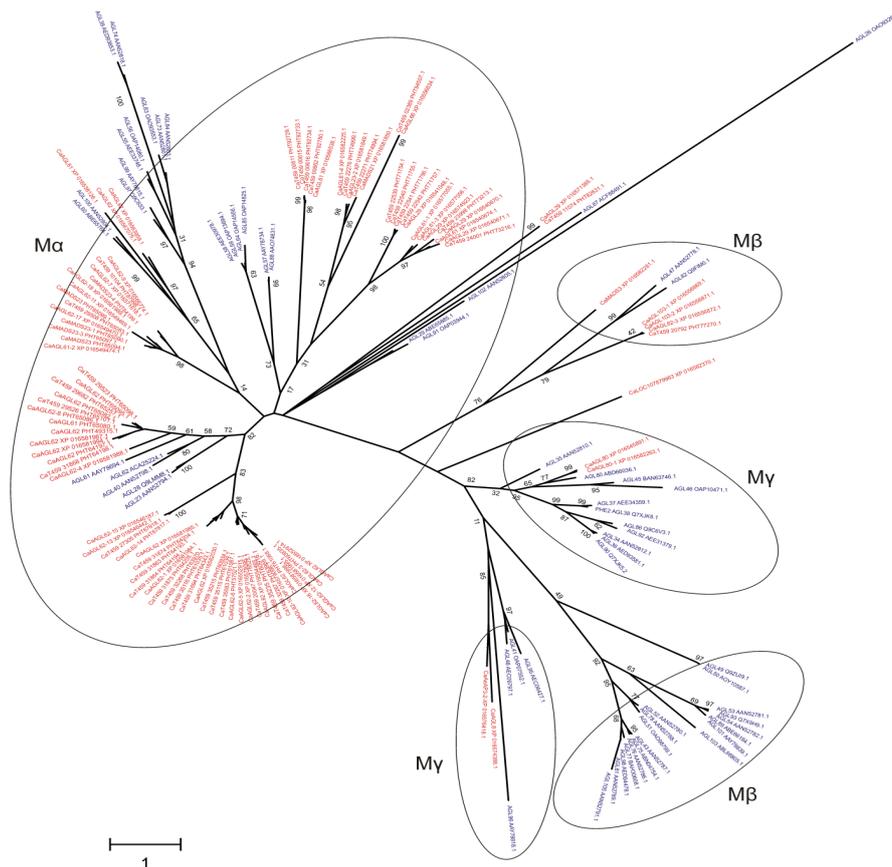


Рисунок 1. Филогения MADS белков типа I *A. thaliana* (выделено синим цветом) и *C. annuum* (выделено красным цветом)

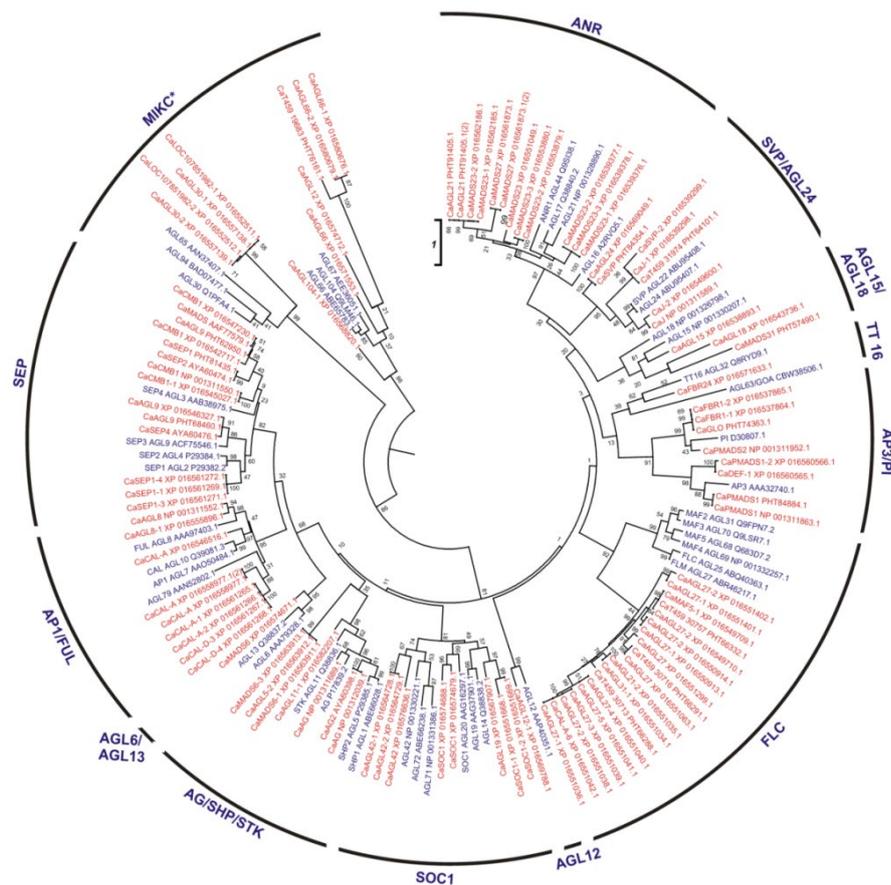


Рисунок 2. Филогения MADS белков типа II *A. thaliana* (выделено синим цветом) и *C. annuum* (выделено красным цветом)

Учитывая консервативность механизма цветения, можно предположить, что выявленные у перца *C. annuum* структурные гомологи охарактеризованных генов MADS-box *A. thaliana* имеют сходные с ними функции. Анализ имеющихся данных о функциях MADS-box генов *A. thaliana*, *C. annuum* и других видов позволил сделать предварительные выводы о возможной роли генов *CanMADS* в онтогенезе *C. annuum* (рис. 3). Так, белки типа I могут играть важные роли в репродуктивном развитии перца, как было показано для их гомологов у *A. thaliana* (Masiero et al., 2011). Функции белков типа II более разнообразны. Члены клады AP1/FUL могут выполнять А-активность при закладке меристемы цветка и околоцветника перца, что подтверждается более ранними исследованиями (Sung et al., 2001). Гомологи AP3/PI вероятно осуществляют В-активность, то есть определяют идентичность лепестков и тычинок. В подтверждение этому, ранее было показано участие PAP3 в развитии тычинок перца (Li et al., 2013; Ma et al., 2016). Белки группы AG/SHP/STK предположительно отвечают за С- и D-активности, участвуя в дифференцировке репродуктивных органов. SEP-белки (Е-активность) могут определять идентичность всех цветковых органов, как ранее показано для *CanMADS1* (Sung et al., 2001) и *CaMADS-RIN* (Dong et al., 2014, Hassan et al., 2014). Время цветения могут регулировать гомологи AGL24/SVP, что уже было продемонстрировано для генов перца MPF2 и *CaJ* (Hu, Saedler, 2007; Cohen et al., 2012), а также гомологи SOC1 и FLC. Гомологи AGL15 могут участвовать в контроле развития семян, гомологи ANR и AGL12 – корневой системы, гомолог TT16 и *CaMADS-RIN* – синтеза антоцианов и каротиноидов, соответственно, а гомологи FUL, AG, SEP – в регуляции роста и созревания плода. Последнее было ранее продемонстрировано на примере генов-гомологов *RIN*, *FUL1*, *FUL2*, *TAGL1* и *CMB1* перца (Dubey et al., 2019). На основе полученных данных была предложена схема участия выявленных MADS-box генов в генетической регуляции репродуктивного развития перца.



Рисунок 3. Возможные функции MADS-ТФ перца в развитии *C. annuum*

Полученные результаты могут быть использованы как для фундаментальных исследований развития растений, так и в селекции новых сортов перца.

Работа выполнена при поддержке РФФИ № 18–29–07007 и Министерства науки и высшего образования Российской Федерации.

Литература

- Cohen O. et al. CaJOINTLESS is a MADS-box gene involved in suppression of vegetative growth in all shoot meristems in pepper //Journal of experimental botany. – 2012. – Т. 63. – №. 13. – С. 4947–4957.
- Dong T. et al. A non-climacteric fruit gene CaMADS-RIN regulates fruit ripening and ethylene biosynthesis in climacteric fruit //PloS one. – 2014. – Т. 9. – №. 4. – С. e95559.
- Dubey M. et al. Identification of genes involved in fruit development/ripening in Capsicum and development of functional markers //Genomics. – 2019. – Т. 111. – №. 6. – С. 1913–1922.
- Hassan O.S., Badie F., Safwat G. Quantification of the gene expression of bell peppers (*Capsicum annuum*) ripening gene (s) using real-time PCR //African Journal of Biotechnology. – 2014. – Т. 13. – №. 53. – С. 4745–4752.
- Hu J.Y., Saedler H. Evolution of the inflated calyx syndrome in Solanaceae //Molecular biology and evolution. – 2007. – Т. 24. – №. 11. – С. 2443–2453.
- Kim S. et al. New reference genome sequences of hot pepper reveal the massive evolution of plant disease-resistance genes by retroduplication //Genome biology. – 2017. – Т. 18. – №. 1. – С. 1–11.
- Li X. et al. Expression pattern of class B gene PAP3 in flower development of pepper //International journal of molecular sciences. – 2013. – Т. 14. – №. 12. – С. 24643–24655.
- Ma N. et al. PAP3 Regulates Stamen but Not Petal Development in *Capsicum annuum* L //Horticultural Plant Journal. – 2016. – Т. 2. – №. 2. – С. 91–96.
- Masiero S. et al. The emerging importance of type I MADS box transcription factors for plant reproduction //The Plant Cell. – 2011. – Т. 23. – №. 3. – С. 865–872.
- Naves E.R. et al. Capsaicinoids: pungency beyond *Capsicum* //Trends in plant science. – 2019. – Т. 24. – №. 2. – С. 109–120.
- Parenticová L. et al. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world //The Plant Cell. – 2003. – Т. 15. – №. 7. – С. 1538–1551.
- Qin C. et al. Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization //Proceedings of the National Academy of Sciences. – 2014. – Т. 111. – №. 14. – С. 5135–5140.
- Sung S.K. et al. Characterization of MADS box genes from hot pepper //Molecules and cells. – 2001. – Т. 11. – №. 3. – С. 352–359.
- Theissen G., Saedler H. Floral quartets //Nature. – 2001. – Т. 409. – №. 6819. – С. 469–471.