

ВАРИАБЕЛЬНОСТЬ СТРУКТУРЫ ЦЕНТРОМЕРНОГО ГИСТОНА H3 И ФИЛОГЕНИЯ РОДА *SECALE*

© Е. В. Евтушенко,¹ Д. Ю. Гаськов, С. С. Гацкая, А. В. Вершинин

Институт молекулярной и клеточной биологии СО РАН, Новосибирск, 630090;

¹электронный адрес: evt@mcb.nsc.ru

Анализ варибельности кодирующих последовательностей центромерного варианта гистона H3 (CENH3) проведен среди 15 видов и подвидов ржи (род *Secale*) с использованием значения нуклеотидного разнообразия (π) и оценки количества полиморфных сайтов (S) для N-терминального района (NTT) и консервативного домена (HFD) гена. Значения π для NTT превышают соответствующие значения для района HFD. Не обнаружено зависимости — между значениями π и длительностью онтогенеза и системой размножения. У многолетнего вида *Secale strictum* C. Presl и однолетнего *S. cereale* L. значения π для NTT достоверно не различаются, превышая соответствующие значения, полученные для вида *S. silvestre*. Оценки π для самоопыляющихся подвидов сравнимы со значениями для перекрестноопыляющихся подвидов. Результаты проведенного анализа дают основания утверждать, что существование видов рода *Secale* в условиях узкой экологической ниши, эволюционный возраст видов и влияние селекции может приводить к уменьшению нуклеотидного разнообразия в структуре гена CENH3.

Ключевые слова: *Secale* L., центромерный гистон H3, нуклеотидный полиморфизм.

Принятые сокращения: CENH3 — центромерный гистон H3, HFD — C-концевой домен CENH3 (histone fold domain), кДНК — комплементарная ДНК, NTT — N-терминальный район CENH3, п. о. — пары оснований, ОТ-ПЦР — полимеразная реакция с обратной транскрипцией.

Одной из эпигенетических характеристик центромер в эукариотической клетке является наличие центромер-специфического варианта гистона H3, CENH3, первоначально описанного как CENP-A у человека (Earnshaw, Rothfield, 1985). Присутствие CENH3 является необходимым условием для сборки всех остальных компонентов кинетохора, к которому прикрепляются нити веретена деления (Warburton et al., 1997). В отличие от канонических гистонов H3, которые весьма консервативны и почти идентичны по последовательности от *Entamoeba* до человека, CENH3 является более варибельным: полагают, что центромерный гистон изменяется вместе с быстро эволюционирующими центромерными повторами, поддерживая консервативные функции центромеры (Roach et al., 2012). Структурно CENH3 состоит из двух частей — варибельного (по длине и последовательности) NH2-терминального конца и более консервативного C-терминального домена (histone fold domain, HFD). Для *Arabidopsis thaliana* показано, что N-терминальная часть CENH3 участвует в сегрегации хромосом в митозе и мейозе (Lermontova et al., 2011). В составе HFD CENH3 находится функционально важный район, состоящий из петли 1 и α 2-спирали, который необходим для локализации CENH3 на центромере (Vermaak et al., 2002). Считается, что у диплоидных организмов присутствует единственная копия гена CENH3, но для ряда диплоидных видов растений показано существование двух и более копий CENH3 (Ishii et al., 2015).

Род *Secale*, относящийся к семейству злаков *Poaceae*, согласно классификации, предложенной Фредериксеном и Петерсеном (Frederiksen, Petersen, 1998), состоит из трех биологических видов: перекрестноопыляющегося многолетнего *S. strictum* Presl., перекрестноопыляющегося однолетнего *S. cereale* L. и клейстогамного однолетнего *S. silvestre* Host. Вид *S. silvestre* считается наиболее древним, тогда как вид *S. cereale* является эволюционно наиболее молодым. Исследований структуры функциональных CENH3 у диплоидных видов дикой ржи ($2n = 14$, RR) (Lundqvist, 1956) ранее не проводили. Такое исследование представляет несомненный интерес вследствие высокой изменчивости систем размножения ржи. Кроме того, виды и подвиды ржи широко используют в скрещиваниях с различными видами пшеницы и в создании ржано-пшеничных дополненных и замещенных линий, которые находят дальнейшее применение в селекции пшеницы (Feuillet et al., 2007).

В настоящей работе мы провели сравнительный анализ нуклеотидной изменчивости кодирующих последовательностей гена CENH3 у широкого набора видов и подвидов рода *Secale* с целью определить, в какой степени консервативность структуры CENH3 зависит от происхождения, эволюционного возраста, системы размножения, среды обитания, продолжительности онтогенеза, действия селекционного отбора.

Материал и методика

В исследовании использовали образцы дикой, сорной или культивируемой ржи. *S. cereale* subsp. *cerale* сорта Otello (R1264), *S. cereale* subsp. *vavilovii* (R1027), *S. cereale* subsp. *dighoricum* (R803), *S. strictum* subsp. *kuprijanovii* (R549), *S. strictum* subsp. *ciliatoglume* (R548), *Secale silvestre* (R1116), *S. cereale* subsp. *ancestrale* (R62), *Secale x derzhavinii* Tzvelev (R630) и *S. cereale* subsp. *afghanicum* (HR566/86) были получены из Института растительных культур (IPK, Германия). *S. cereale* L. subsp. *segetale* (PI 102) и *S. strictum* subsp. *anatolicum* (PI 206992) предоставлены USDA-ARS. *S. strictum* subsp. *africanum* (10289), *S. strictum* subsp. *strictum* (10736), *S. cereale* subsp. *cerale* сорта Black Winter (9395) и *S. cereale* subsp. *iranicum* (10431) любезно предоставлены Отделом генетических ресурсов овса, ржи, ячменя Всероссийского института растениеводства им. Н. И. Вавилова (Россия).

Общую РНК экстрагировали из листьев молодых растений с применением TRI Reagent (MRC, Inc., США) согласно руководству производителя. Для синтеза кДНК использовали Reverse Aid Minus First Strand cDNA Synthesis kit (Fermentas, Thermo Fisher Scientific), кДНК использовали в ОТ-ПЦР. Для амплификации N-терминального района использовали праймеры 5'-ATGGCCCGCACCAAGCAC (F) и 5'-GAAACTCGACCGACTTCTG (R). Для амплификации HFD — 5'-GTRGCRCTGCGGGA-GATCAGGA (F) и 5'-CTBGCRAGYTGYATGTCSTTTT

(R) (Nagaki et al., 2015). Обе цепи клонированных ПЦР-продуктов секвенировали согласно стандартному протоколу в ЦКП «Геномика» СО РАН (Новосибирск), поиск гомологии нуклеотидных последовательностей проводили с использованием алгоритма BLAST (Altschul et al., 1990) в базе данных NCBI (<http://blast.ncbi.nlm.nih.gov/Database/>). Программу DnaSP версии 5.10 (Librado, Rozas, 2009) применяли для подсчета уровня генетической вариации внутри видов, используя значение π как оценку внутривидового полиморфизма.

Результаты и обсуждение

Последовательности ДНК центрального варианта гистона H3 выделены и секвенированы из 15 образцов ржи, принадлежащих трем видам *Secale*. Полноразмерный вариант последовательности *CENH3* ржи (*ScCENH3*) составляет 501 п. о., полная длина последовательности N-терминального района (NTT — N-terminal tail) — 216 п. о., размер последовательностей HFD — 211 п. о. (после удаления области праймеров, включая маловариабельные у ржи участки α N-спирали и α 3-спирали). При этом анализировали функционально важные районы HFD — α 1-спираль, петлю 1, α 2-спираль, часть α 3-спирали.

Для оценки внутривидового полиморфизма доменов *CENH3* определяли генетическое разнообразие на уровне

Нуклеотидное разнообразие последовательностей *CENH3* у видов и подвидов *Secale*

Виды и подвиды <i>Secale</i>	Способ размножения ^a	Кодирующая область <i>CENH3</i>			
		NTT		HFD	
		$\pi_{tot} \pm SD (\times 10^{-2})^b$	S_n^b	$\pi_{tot} \pm SD (\times 10^{-2})$	S_n
Вид <i>S. strictum</i> Presl.		1.39 ± 0.15 (n = 51) ^c		0.83 ± 0.09 (n = 92)	
Многолетние подвиды					
<i>S. anatolicum</i>	п	1.42 ± 0.26	15 (6; 9)	1.22 ± 0.24	10 (4; 6)
<i>S. africanum</i>	с	1.41 ± 0.24	8 (5; 3)	1.02 ± 0.16	13 (4; 9)
<i>S. strictum</i>	п	1.22 ± 0.19	9 (2; 7)	1.13 ± 0.33	13 (4; 9)
<i>S. ciliatoglume</i>	»	1.25 ± 0.23	6 (4; 2)	0.84 ± 0.10	14 (6; 8)
<i>S. kuprijanovii</i>	»	0.65 ± 0.26	3 (2; 1)	0.200 ± 0.073	5 (2; 3)
<i>Secale</i> × <i>derzhavinii</i>	»	0.23 ± 0.12	1 (0; 1)	2.12 ± 0.51	11 (2; 9)
Вид <i>S. cereale</i> L.		1.22 ± 0.15 (n = 57)		0.88 ± 0.11 (n = 106)	
Однолетние подвиды					
<i>S. ancestrale</i>	»	1.38 ± 0.19	8 (3; 5)	1.13 ± 0.19	15 (5; 10)
<i>S. iranicum</i>	с	1.37 ± 0.22	6 (3; 3)	0.95 ± 0.27	5 (1; 4)
<i>S. afghanicum</i>	п	1.24 ± 0.12	5 (5; 2)	0.81 ± 0.25	11 (7; 4)
<i>S. vavilovii</i>	с	1.21 ± 0.40	6 (5; 2)	0.71 ± 0.21	6 (5; 1)
<i>S. segetale</i>	п	0.94 ± 0.32	4 (4; 0)	0.95 ± 0.18	13 (7; 6)
<i>S. cereale</i> , Otello	»	0.770 ± 0.099	7 (4; 3)	0.170 ± 0.092	2 (2; 0)
<i>S. cereale</i> , B. Winter	»	0.90 ± 0.26	7 (4; 3)	1.47 ± 0.49	15 (8; 5)
<i>S. dighoricum</i>	»	0.39 ± 0.19	3 (1; 2)	0.053 ± 0.046	1 (0; 1)
Однолетний вид					
Вид <i>S. silvestre</i> Host.	с	0.91 ± 0.16	9 (4; 5)	0.87 ± 0.16	15 (8; 8)

Примечание. ^a п — перекрестное опыление, с — самоопыление. ^b Приведены общее разнообразие для синонимичных и несинонимичных замов и его стандартное отклонение (SD). ^c Количество полиморфных сайтов, в скобках указано число синонимичных и несинонимичных замов. ^d Общее количество последовательностей вида для домена.

нуклеотидных последовательностей (π) и количество полиморфных сайтов (S). Нуклеотидное разнообразие π — среднее число нуклеотидных различий на сайт между двумя последовательностями (Nei, Li, 1979), анализировали отдельно для NTT и HFD, принимая во внимание различные функции доменов *CENH3* (см. таблицу).

Сравнение значений π для трех видов *Secale* показало, что для видов *S. cereale* и *S. strictum* C. Presl. они примерно равны и несколько ниже для вида *S. silvestre*, при этом значения π ниже для консервативного домена по сравнению с NTT (значения π для NTT перечисленных видов были равны 1.22, 1.39 и 0.91, для HFD — 0.88, 0.83 и 0.87). Оценки нуклеотидного разнообразия кодирующей структуры *CENH3* видов *Secale* близки к оценкам, полученным для *HTR12* (аналог *CENH3*) подвидов *Arabidopsis lyrata* и *A. lyrata* ssp. *petraea* (Kawabe et al., 2006).

Для последовательностей *CENH3* некоторых подвидов ржи оценки π были намного ниже, чем значения для видов. Среди многолетних подвидов наиболее низкие значения π были у дикой ржи *S. kuprijanovii* ($\pi_{\text{tot}} = 0.65 \pm 0.26$ для NTT и $\pi_{\text{tot}} = 0.200 \pm 0.073$ для HFD). Среди однолетних подвидов низкий полиморфизм *CENH3* наблюдали у культивируемой ржи *S. cereale* сорта Отелло ($\pi_{\text{tot}} = 0.770 \pm 0.099$ для NTT и $\pi_{\text{tot}} = 0.170 \pm 0.092$ для HFD) и сорной ржи *S. dighoricum* ($\pi_{\text{tot}} = 0.39 \pm 0.19$ для NTT и $\pi_{\text{tot}} = 0.053 \pm 0.040$ для HFD). Способ размножения, по-видимому, не влияет на нуклеотидное разнообразие в локусе *CENH3*: значения π для самоопыляющихся подвидов *S. africanum* (1.41 ± 0.24 и 1.02 ± 0.16), *S. vavilovii* (1.21 ± 0.40 и 0.71 ± 0.21) и самофертильной ржи *S. iranicum* (1.37 ± 0.22 и 0.95 ± 0.27) сравнимы со значениями, полученными для перекрестноопыляющихся подвидов.

Можно предположить, что наблюдаемый у некоторых подвидов низкий паттерн нуклеотидного разнообразия в локусе *CENH3* обусловлен узким географическим ареалом распространения данного подвида. Например, дикий подвид *S. kuprijanovii* обитает в природных, не потревоженных деятельностью человека ландшафтах (Tang et al., 2011). Отсутствие полиморфных сайтов и низкое нуклеотидное разнообразие последовательностей *CENH3* у перекрестноопыляющейся ржи *S. dighoricum* можно объяснить тем, что изученный нами образец, вероятно, происходит из популяции, произрастающей в очень ограниченном ареале в долине Большого Кавказа (Цвелев, 1976). Невысокий уровень нуклеотидного разнообразия в NTT у культивируемой ржи сорта Otello и искусственно фертильного гибрида *Secale* × *derzhavinii* возник, по-видимому, вследствие уменьшения разнообразия при селекции. Такие подвиды однолетней и многолетней ржи, как *S. afghanicum*, *S. ancestrale*, *S. vavilovii*, *S. segetale*, *S. anatolicum* и *S. ciliatoglume*, являются сорными, встречаются в посевах пшеницы, на обочинах дорог рядом с полями (Tang et al., 2011). Повсеместное распространение в сочетании с перекрестным способом размножения может поддерживать относительно высокий уровень нуклеотидного разнообразия *CENH3* у этих подвидов.

Высокий уровень нуклеотидного разнообразия последовательностей *CENH3* в NTT самофертильного образца однолетней ржи *S. iranicum* обусловлен наличием трех несинонимичных и трех синонимичных замен в позициях 50, 66, 99, 126, 130 и 144 ORF. Большая часть этих замен встречается как у однолетних, так и многолетних видов, за исключением замены в позиции 50. В полученном

нами наборе последовательностей *CENH3* это замещение, приводящее к замене аминокислоты (K17R), было найдено в последовательностях многолетних *S. anatolicum*, *S. africanum* и однолетней *S. iranicum*. Присутствие общих вариантов N-терминальной части центромерного гистона H3 у этих образцов может объяснять наличие относительно небольшого репродуктивного барьера между *S. iranicum* и *S. strictum*. При скрещивании *S. iranicum* и *S. strictum* стерильность пыльцы гибридов составляет около 50 %, завязываемость семян — 40 % (Кобылянский, 1983). В используемой нами филогенетической классификации (Frederiksen, Petersen, 1998) *S. iranicum* не является самостоятельным подвидом *S. cereale* (ряд авторов полагают его образцом *S. vavilovii*), но по биологическим, цитологическим и генетическим особенностям этот подвид стоит ближе к *S. strictum*, чем к *S. cereale*. Согласно систематике ржи, предложенной Кобылянским (1983), *S. iranicum* является отдельным видом. У подвида *S. anatolicum* помимо полноразмерного варианта *CENH3* обнаружен вариант, имеющий делецию длиной 21 п. о., в районе NTT. Аналогичная последовательность *CENH3* есть и у подвида *S. africanum*, что говорит в пользу гипотезы об обособлении данного подвида от подвида *S. anatolicum*.

В исследованиях на *A. thaliana* показано влияние единичных аминокислотных замен в последовательности *CENH3* (как в NTT, так в HFD) на прохождение мейоза. Растения с единичной заменой аминокислоты в *CENH3* дают потомство при самоопылении, но при скрещивании с растениями дикого типа наблюдаются нарушения в распределении хромосом, анеуплоидия и гибель семян, следовательно, природные различия в структуре *CENH3* даже близкородственных видов могут способствовать генетическому разнообразию и репродуктивной изоляции популяций (Kurpu et al., 2015).

Полученные результаты позволяют заключить, что на уровень нуклеотидного разнообразия последовательностей *CENH3* у однолетних и многолетних подвидов ржи не влияет способ размножения (перекрестное опыление или самоопыление). К уменьшению разнообразия может приводить существование вида в условиях узкой экологической ниши или селекция. Полученные результаты указывают на важность изучения вариативности нуклеотидных последовательностей, кодирующих *CENH3*. Оценки нуклеотидного разнообразия *CENH3*, анализ распространения несинонимичных замен дают возможность проследить историю возникновения отдельных подвидов в роде *Secale*. С учетом важной роли *CENH3* в функционировании центромеры и сегрегации хромосом исследование структуры функциональных *CENH3* может служить одним из критериев при подборе вариантов скрещивания диких и культивируемых видов с целью переноса экономически важных признаков.

Работа выполнена при финансовой поддержке бюджетного проекта 0310-2014-0002 ИМКБ СО РАН.

Список литературы

- Кобылянский В. Д. 1983. Филогения рода *Secale* L. Тр. по прикл. ботанике, генетике и селекции. 79 : 38—45. (*Kobylyanskij V. D. 1983. Phylogeny of the genus Secale. Trudy Prikl. Bot. 79 : 38—45.*)
- Цвелев Н. Н. 1976. Систематический обзор злаков флоры СССР. В кн.: Злаки СССР. Л.: Наука. 173 с. (*Tzvelev N. N. 1976. Poaceae URSS. L.: Nauka. 173 p.*)

- Altschul S. F., Gish W., Miller W., Myers E. W., Lipman D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215 (3) : 403.
- Earnshaw W. C., Rothfield N. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma.* 91 : 313—321.
- Feuillet C., Langridge P., Waugh R. 2007. Cereal breeding takes a walk on the wild side. *Trends Genet.* 24 : 24—32.
- Frederiksen S., Petersen G. 1998. A taxonomic revision of *Secale*. *Nord J. Bot.* 18 : 399—420.
- Ishii T., Karimi-Ashtiyani R., Banaei-Moghaddam A. M., Schubert V., Fuchs J., Houben A. 2015. The differential loading of two barley CENH3 variants into distinct centromeric substructures is cell type and development-specific. *Chromosome Res.* 23 : 277—284.
- Kawabe A., Nasuda S., Charlesworth D. 2006. Duplication of centromeric histone H3 (*HTR12*) gene in *Arabidopsis halleri* and *A. lyrata*, plant species with multiple centromeric satellite sequences. *Genetics.* 174 : 2021—2032.
- Kuppu S., Tan E. H., Nguyen H., Rodgers A., Comai L., Chan S. W. L., Britt A. B. 2015. Point mutations in centromeric histone induce post-zygotic incompatibility and uniparental inheritance. *PLoS Genet.* 11 : e1005494.
- Lermontova I., Koroleva O., Rutten T., Fuchs J., Schubert V., Moraes I., Koszegi D., Schubert I. 2011. Knockdown of *CenH3* in *Arabidopsis* reduces mitotic divisions and causes sterility by disturbed meiotic chromosome segregation. *Plant J.* 68 : 40—50.
- Librado P., Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 25 : 1451—1452.
- Lundqvist A. 1956. Self-incompatibility in rye. I. Genetic control in the diploid. *Hereditas.* 42 : 293—348.
- Nei M., Li W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Nat. Acad. Sci. USA.* 76 : 5269—5273.
- Nagaki K., Kashihara K., Murata M. 2005. Visualization of diffuse centromeres with centromere-specific histone H3 in the holocentric plant *Nuzula nivea*. *The Plant Cell.* 17 : 1886—1893.
- Roach K. C., Ross B. D., Malik H. S. 2012. Rapid evolution of centromeres and centromeric/kinetochore proteins. In: *Evolution in the fast lane: rapidly evolving genes and genetic systems.* Oxford Univ. Press. 83—93.
- Tang Z. X., Ross K., Ren Z. L., Yang Z. J., Zhang H. Y., Chikmawati T., Miftahudin, Gustafson J. P. 2011. *Secale*. In: *Wild crop relatives: genomic and breeding resources, cereals.* Berlin; Heidelberg: Springer-Verlag. 367—396.
- Vermaak D., Hayden H. S., Henikoff S. 2002. Centromere targeting element within the histone fold domain of Cid. *Mol. Cell Biol.* 22 : 753—756.
- Warburton P. E., Cooke C. A., Bourassa S., Vafa O., Sullivan B. A., Stetten G., Gimelli G., Warburton D., Tyler-Smith C., Sullivan K. F., Poirier G. G., Earnshaw W. C. 1997. Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Curr. Biol.* 7 : 901—904.

Поступила 1 XII 2015

VARIABILITY OF CENTROMERIC HISTONE H3 VARIANTS
AND PHYLOGENETIC RELATIONSHIPS IN *SECALE*

E. V. Evtushenko,¹ D. Yu. Gaskov, S. S. Gatskaya, A. V. Vershinin

Institute of Molecular and Cellular Biology, SB RAS, Novosibirsk, 630090;

¹ e-mail: evt@mcb.nsc.ru

We examined nucleotide sequence variation in centromeric histone H3 (*CENH3*) in fifteen rye species and subspecies, including annuals, perennials, self-pollinating and cross-pollinated plants. Levels of genetic variation within N-terminal tail (NTT) and histone fold domain (HFD) were estimated as average per site pairwise nucleotide diversity (π) and as the number of segregating polymorphic sites (*S*). A comparison of nucleotide diversity (π_{tot}) for NTT and HFD showed that estimates of diversity are consistently greater in NTT. Our estimates of NTT nucleotide diversity for perennial *Secale strictum* C. Presl were similar to those observed in annual *S. cereale* L., elevating those estimates for ancient species *S. silvestre*. A similar pattern of nucleotide polymorphism was observed between the self-pollinating and cross-pollinated subspecies. These results suggest that lower nucleotide diversity within *CENH3* for some wild and cultivated subspecies of *Secale* may be attributable to narrow ecological niche, either evolutionary age or breeding.

Key words: *Secale* L., centromeric histone H3, nucleotide polymorphism.