

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ



УДК 577.22

Разнообразие гистонов H2A и их влияние на структурные свойства нуклеосомы

Л. Сингх-Пальчевская , А.К. Шайтан* 

Кафедра биоинженерии, биологический факультет, Московский государственный университет имени М.В. Ломоносова,
Россия, 119234, г. Москва, Ленинские горы, д. 1, стр. 12

*e-mail: shaytan_ak@mail.bio.msu.ru

Ключевым эпигенетическим фактором являются гистоновые белки, которые играют важную роль в динамике хроматина и регуляции активности генов. Они делятся на два обширных класса: канонические гистоны и их варианты. Канонические гистоны экспрессируются в основном в ходе S-фазы клеточного цикла, так как участвуют в упаковке ДНК в процессе деления клетки. Гистоновые варианты — это гены гистонов, которые экспрессируются и регулируют динамику хроматина в ходе всего клеточного цикла. В силу функционального и видового разнообразия выделяют различные семейства вариантов гистонов. Некоторые белки характеризуются незначительными отличиями от канонических гистонов, другие же наоборот могут иметь множество важных структурных и функциональных особенностей, влияющих на стабильность нуклеосомы и динамику хроматина. Для того чтобы оценить вариативность гистонов семейства H2A и их влияние на структуру нуклеосомы, мы провели биоинформатический анализ аминокислотных последовательностей гистонов семейства H2A. Проведенная кластеризация методом UPGMA позволила выделить два основных подсемейства белков H2A: «короткие» H2A (short H2A) и другие варианты H2A, демонстрирующие более высокую консервативность аминокислотных последовательностей. Также мы построили и проанализировали множественные выравнивания для различных подсемейств гистонов H2A. Важно отметить, что белки подсемейства «коротких» H2A являются не только самыми низко консервативными внутри своего семейства, но и имеют особенности, оказывающие существенное влияние на структурные свойства нуклеосомы. Кроме того, мы провели филогенетический анализ «коротких» гистонов H2A, в результате которого были более детально охарактеризованы подсемейства, соответствующие вариантам H2A.B, H2A.P, H2A.Q, H2A.L.

Ключевые слова: гистон, H2A, гистоновые варианты, нуклеосома, хроматин, анализ последовательностей, биоинформатика, эпигенетика

DOI: 10.55959/MSU0137-0952-16-78-4-4

Введение

Изучение механизмов функционирования генома является одной из важнейших задач молекулярной биологии на сегодняшний день. Известно, что существует множество различных эпигенетических факторов, влияющих на активность генов и позволяющих организму функционировать и развиваться. Ключевую роль среди них занимают белки гистоны, которые отвечают за структуру и динамику хроматина.

Большая часть гистонов обеспечивает упаковку новой ДНК, синтезированной в ходе репликации. Они называются каноническими и экспрессируются в основном во время S-фазы клеточного цикла [1]. Экспрессия других гистонов, именуе-

мых вариантами, не зависит от репликации. Они встраиваются в существующие нуклеосомы и регулируют активность и функционирование различных участков генома [2].

Выделяют пять основных типов гистонов: коровые H2A, H2B, H3, H4 и линкерный H1/H5. Каждый из них представлен как канонической формой, так и различными вариантами (за исключением H4, который во многих видах представлен лишь канонической изоформой). Аминокислотные последовательности некоторых гистонов имеют отличия всего в нескольких аминокислотных остатках [2]. Например, вариант H3.3 и канонический H3 [3]. При этом они имеют схожие структурные особенности и схожую функциональную роль.

Другие гистоновые варианты могут характеризоваться значительным количеством различий (идентичность последовательностей менее 25%) [2], а также иметь значимые вариации, влияющие на структуру нуклеосомы. Например, вариант гистона *масгоH2A* обладает дополнительным негистоновым макромономером размером около 30 кДа, благодаря которому он способен связывать производные метаболиты НАД⁺, такие как АДФ-рибоза и поли-АДФ-рибоза [4]. Разнообразие гистонов различимо не только на структурном и функциональном уровнях. Некоторые варианты могут быть специфичны для отдельных таксономических групп или клеточных линий. Например, варианты H2A.W и OO H1.8 присутствуют исключительно в растениях и в ооцитах млекопитающих соответственно [5–10].

В силу своего многообразия гистоновые белки делятся на различные семейства. Однако до сих пор отсутствует систематизированное представление о значимых вариациях аминокислотных последовательностей гистонов и их структурных особенностях. Для того чтобы более детально изучить семейство гистоновых вариантов H2A, мы провели биоинформатический анализ их аминокислотных последовательностей и оценили влияние вариативности вариантов H2A на структуру нуклеосомы.

Материалы и методы

Для проведения комплексного анализа разнообразия гистоновых вариантов H2A были собраны их аминокислотные последовательности. В исследуемый массив данных были отобраны последовательности из курируемого набора базы данных HistoneDB (<https://histdb.intbio.org>) [2]. На основе анализа литературы [5, 9–13], данный набор был расширен последовательностями гистонов, которые принадлежат недавно обнаруженным семействам H2A. В результате было собрано более 200 аминокислотных последовательностей гистонов H2A.

Методы биоинформатического анализа. Чтобы оценить общую вариативность всех аминокислотных последовательностей семейства H2A, была произведена их кластеризация. Для этого было построено филогенетическое дерево с использованием программы ClustalW2 [14] для создания множественного выравнивания и алгоритма UPGMA [15] в качестве простого метода иерархической кластеризации.

Для проведения эволюционного анализа с целью сократить влияние разнообразия неупорядоченных хвостов гистонов были извлечены центральные области глобулярных доменов гистонов, гистоновые фолды (histone fold domain, HFD), являющиеся структурными мотивами, характерными для всех гистонов и состоящими из трех альфа-спиралей, соединенных двумя петлями. Множественные выравнивания полученных последова-

тельностей строились с применением программы MUSCLE [16]. Филогенетические деревья были построены с помощью алгоритмов PhyML [17], в основе которых лежат методы максимального правдоподобия.

Для обнаружения значимых вариаций и структурных особенностей были построены множественные выравнивания с использованием программы MUSCLE [16]. Визуальное представление результатов было получено с применением программного обеспечения texshade [18], которое позволяет не только обрабатывать несколько выравниваний последовательностей, но и имеет специальные режимы для подсвечивания функционально значимых остатков.

Результаты и обсуждение

Кластеризация последовательностей гистоновых белков семейства H2A. В ходе кластеризации были построены и проанализированы филогенетическое дерево и матрица попарной идентичности между аминокислотными последовательностями семейства H2A, представленные на рис. 1. Полученные результаты позволяют сделать заключение о том, что гистоны семейства H2A можно разделить на два обширных подсемейства. Состав одного из них представляет собой группу «коротких» гистонов H2A (short H2A), в которую входят подгруппы гистонов H2A.P, H2A.B, H2A.L и H2A.Q, известные из литературы [11, 20, 21]. Другое подсемейство состоит из последовательностей канонических белков и различных вариантов H2A, таких как H2A.X, H2A.Z, *масгоH2A*, H2A.W и H2A.R [1, 4, 5, 9, 11, 22]. Из матрицы попарной идентичности (рис. 1) можно сделать заключение о том, что данные подсемейства имеют разную степень консервативности аминокислотных последовательностей. Так, варианты, относящиеся к «коротким» H2A, демонстрируют очень низкую консервативность по сравнению с остальными вариантами семейства H2A. К тому же, отчетливо выделяются наименее (H2A.P) и наиболее (канонические H2A) консервативные варианты. Опираясь на результаты кластеризации (рис. 1), можно также отметить, что, несмотря на низкую степень сходства аминокислотных последовательностей, в каждом варианте из подсемейства «коротких» H2A выделяются отдельные кластеры последовательностей.

Филогенетический анализ и классификация «коротких» гистоновых белков семейства H2A. Опираясь на филогенетическое дерево (рис. 2А), построенное с помощью алгоритмов максимального правдоподобия с использованием исследуемых последовательностей без неупорядоченных С- и N-хвостов, можно сделать заключение о том, что самыми дивергентными гистонами семейства являются варианты «коротких» H2A, а самыми близкими к ним ортологами являются гистоны недавно обнаруженного варианта H2A.R [7]. Интересно

отметить, что структура дерева указывает на то, что белки канонического H2A и варианта H2A.X разошлись в ходе эволюции не один раз, а в ходе нескольких независимых эволюционных событий. При этом, вероятно, что вариант H2A.X является предшественником канонической формы, так как он играет важную роль в репарации повреждений ДНК [1]. Более того, можно заметить, что некоторые последовательности вариантов H2A.B, H2A.P, H2A.Q, H2A.L, относящихся к подсемейству «коротких» гистонов H2A, образуют отдельные клады (рис. 2Б). Чтобы сопоставить результаты эволюционного анализа с оценкой идентичности последовательностей, мы провели иерархическую кластеризацию гистоновых вариантов «коротких» H2A, используя те же методы, что и на первом этапе нашего исследования. В результате были получены и проанализированы филогенетическое дерево, по-

строенное с помощью метода UPGMA, и матрица попарной идентичности между аминокислотными последовательностями «коротких» H2A (рис. 2В). Мы видим, что варианты семейства «коротких» гистонов H2A могут быть разделены на несколько кластеров. При этом последовательности некоторых кластеров характеризуются достаточно высокой консервативностью (более 70% идентичности). Важно отметить, что в каждый кластер попали последовательности, принадлежащие к одной или нескольким кладам филогенетического дерева. Таким образом, разбиение последовательностей на клады, полученное в результате филогенетического анализа центральных частей гистоновых доменов (гистоновых складок), согласуется с разбиением на основе проведенной нами кластеризации исходных последовательностей, включающих неупорядоченные хвосты (рис. 2Б).

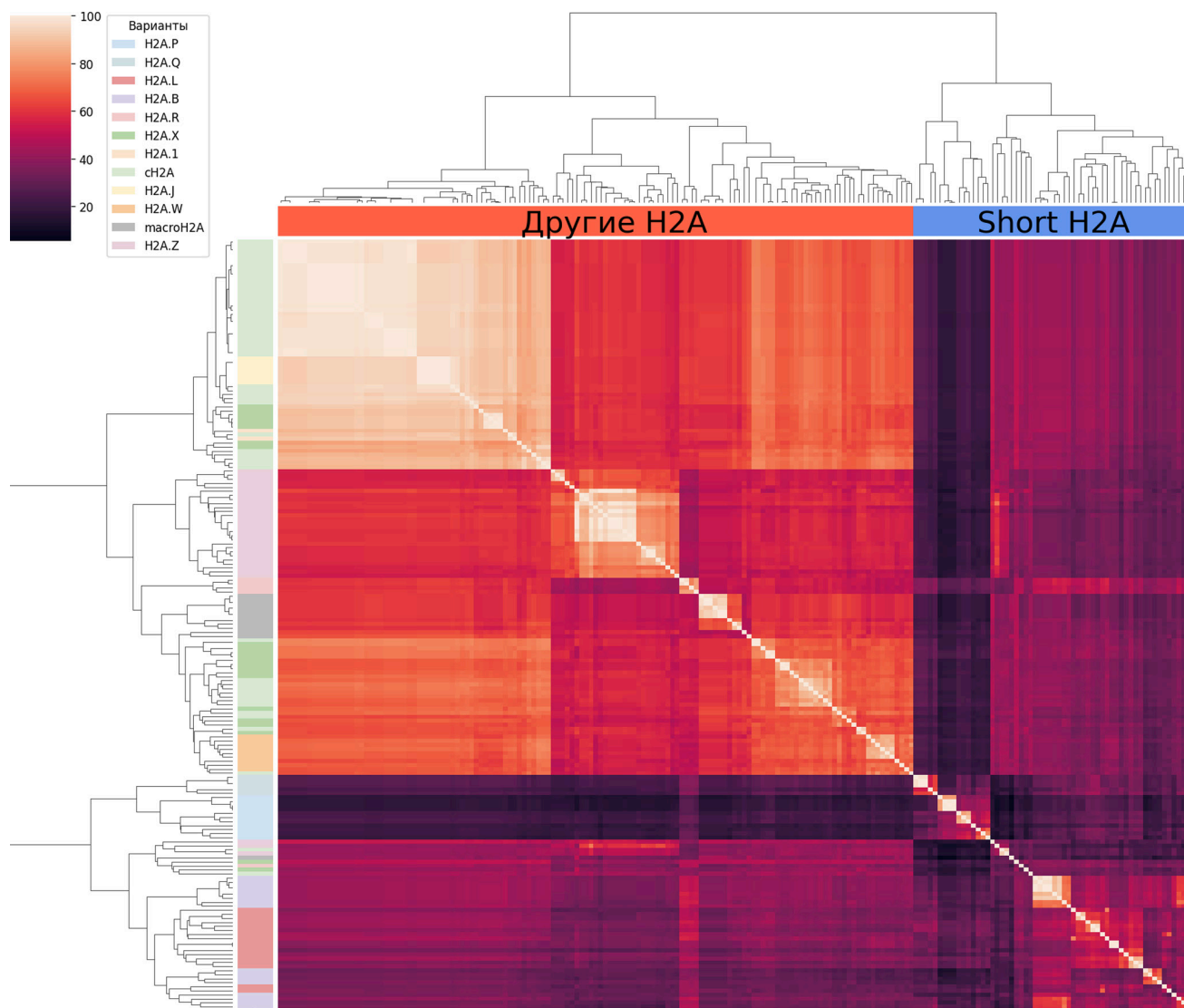


Рис. 1. Матрица попарной идентичности между аминокислотными последовательностями гистонов семейства H2A, полученная в ходе кластеризации гистоновых белков H2A. Цветовая шкала отражает степень идентичности аминокислотных последовательностей. Чем ближе значение к нулю (более темный оттенок), тем ниже идентичность. Слева и сверху представлено дерево иерархической кластеризации гистоновых белков H2A. Цветовая шкала слева определяет название варианта семейства H2A, к которому относится последовательность. Цветовая шкала сверху определяет два кластера гистоновых белков семейства H2A: «короткие» H2A и другие варианты семейства H2A.

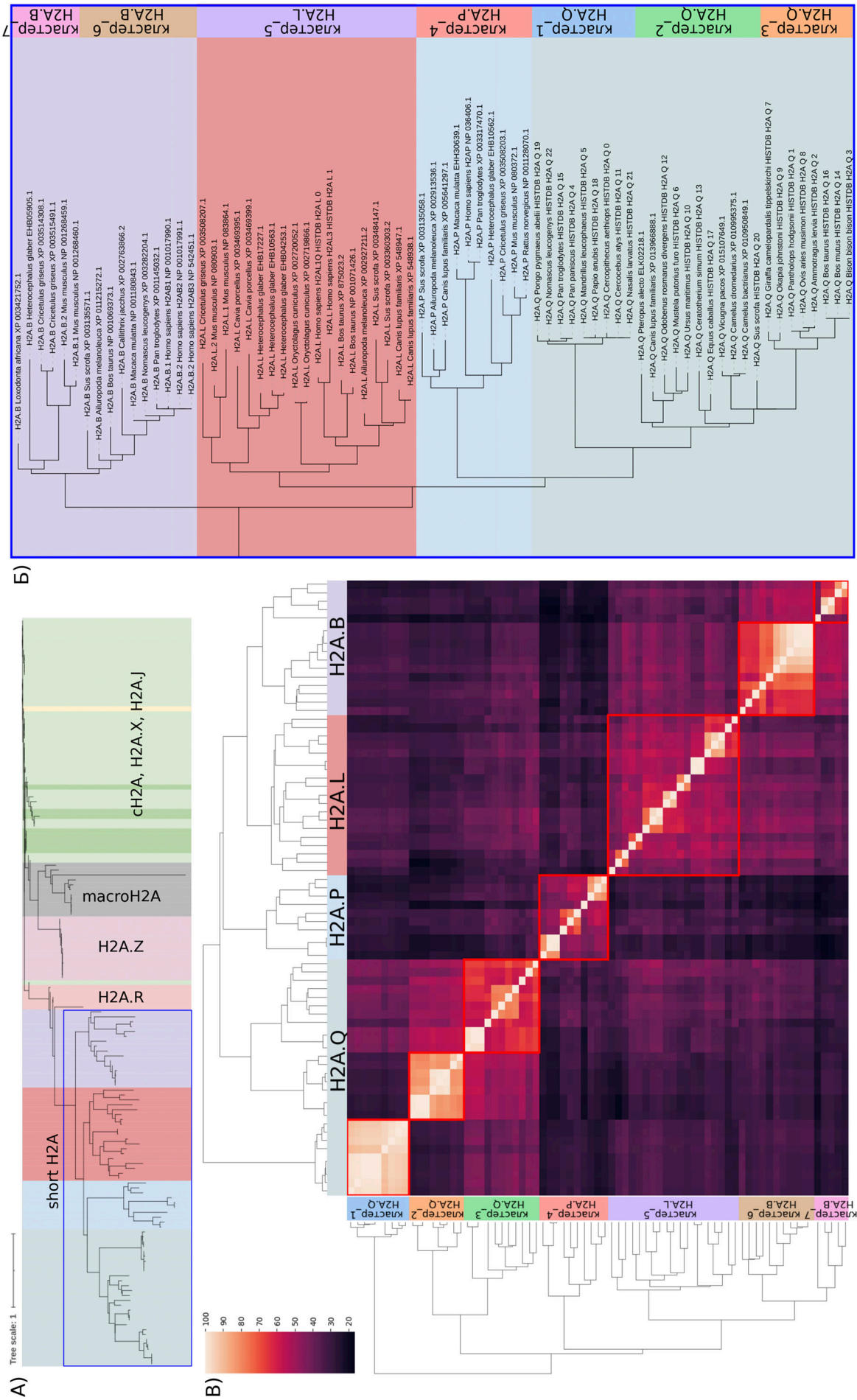


Рис. 2. Филогенетический анализ аминокислотных последовательностей «коротких» гистонов семейства H2A. **А.** Филогенетическое дерево, построенное с помощью методов максимального правдоподобия, для аминокислотных последовательностей центральных частей гистоновых доменов (HFD) «коротких» гистонов H2A. **Б.** Визуализация крупным планом ветвей филогенетического дерева с последовательностями вариантов H2A.В, H2A.Р, H2A.Л, H2A.Х, H2A.Д. Справа на дереве отмечены названия кластеров, полученные в результате кластеризации исходных последовательностей «коротких» гистонов семейства H2A, включающих неупорядоченные хвосты. **В.** Матрица попарной идентичности между аминокислотными последовательностями «коротких» гистонов семейства H2A, полученная в ходе кластеризации «коротких» гистоновых белков H2A. Цветовая шкала слева определяет кластеры «коротких» гистоновых белков H2A. Цветовая шкала сверху определяет название варианта семейства H2A, к которому относится последовательность.

Вариации последовательностей гистоновых белков H2A и их влияние на структуру нуклеосомы. Чтобы охарактеризовать особенности каждого из вариантов семейства H2A, были проанализированы множественные выравнивания аминокислотных последовательностей гистоновых белков семейства H2A. В результате были выявлены различные вариации, являющиеся структурно и функционально значимыми. Важно отметить, что, несмотря на высокую консервативность и одинаковую функциональную роль канонических гистонов, среди их последовательностей также были обнаружены некоторые отличия. Например, аминокислотные последовательности канонического H2A у некоторых протистов (*Alveolata*) характеризуются длинным N-концевым хвостом, а у растений – немного удлиненным C-концевым хвостом. Интересно, что гистоны грибов сильно отличаются от гистонов остальных видов живых организмов. При этом у них отсутствуют две аминокислоты, находящиеся в области «acidic patch», что может приводить к снижению стабильности хроматина [22].

Как показывает анализ множественных выравниваний, вариант H2A.X является наиболее близким к канонической форме гистонов (идентичность аминокислотных последовательностей более 71%). Аминокислотная последовательность гистонического фолда варианта H2A.X отличается от канонического H2A всего в пяти позициях, а область структурного домена «docking domain» – в двух позициях. Вариант H2A.X также интересен тем, что у него на C-концевом хвосте имеется мотив SQE/DΦ (SQAY у дрозофилы), где Φ – представляет собой гидрофобный остаток. Данная особенность имеет важную функциональную роль. Известно, что при фосфорилировании остатка серина (S) вариант H2A.X способен служить маркером двухцепочечных разрывов ДНК и тем самым привлекать машинерию, устраняющую повреждение ДНК [23].

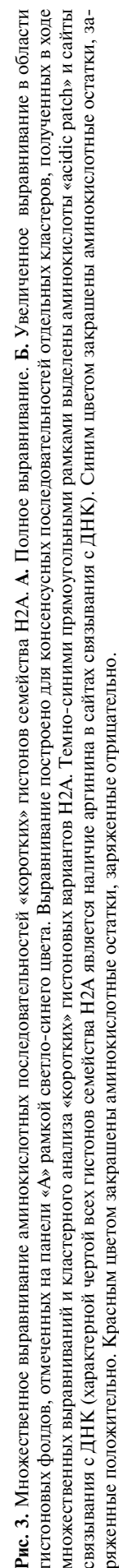
Вариант H2A.Z представляет собой множество последовательностей, обладающих высокой консервативностью между собой (идентичность последовательностей более 72%). При этом анализ выравниваний аминокислотных последовательностей варианта H2A.Z у хордовых позволил выделить две группы белков, которые отличаются всего тремя аминокислотными остатками (данные не представлены). Несмотря на столь незначительные вариации, они кодируются разными генами. Кроме того, некоторые исследования свидетельствуют о том, что данные группы могут быть функционально независимыми. Например, белки варианта H2A.Z, попавшие в одну из групп, кодируются геном, который гомологичен человеческому H2AZ1, а также лучше взаимодействует с белком, содержащим бромодомен BRD2 [24]. Для белков другой группы, кодируемых геном, гомологичным человеческому H2AZ2, было обнаружено, что они предпочтитель-

но связываются с гистоном семейства H3, триметилированным по лизину 4 (H3K4me3) [21].

Самым длинным гистоновым вариантом является macroH2A (в среднем 360 аминокислотных остатков). Такая длина объясняется наличием негистонического макродомена, который соединен с C-концом центральной области глобулярного домена. Аминокислотные последовательности варианта macroH2A демонстрируют высокий уровень консервативности между собой (более 64% идентичности по всей длине). При этом область негистонического макродомена очень разнообразна, что может свидетельствовать о его структурной и функциональной значимости. Более того, на человеческих белках, которые кодируются генами MACRONH2A1 и MACRONH2A2, было показано, что некоторые из них способны связывать производные метаболиты НАД⁺, такие как АДФ-рибоза и поли-АДФ-рибоза, с помощью макродоменов [4].

Аминокислотные последовательности варианта H2A.W, обнаруженного исключительно в растениях, являются представителями одного из самых консервативных вариантов семейства H2A (более 76% идентичности). Наиболее разнообразные вариации встречаются в областях, расположенных между альфа-спиралями. Вариант H2A.W также может быть охарактеризован наличием специфического мотива SPKK, расположенного в C-концевом хвосте. Исходя из литературных данных, можно предположить, что данный мотив играет важную роль в структуре хроматина, так как он может связываться с малой бороздкой ДНК [12].

Отдельного внимания заслуживают «короткие» гистоны семейства H2A, которые преимущественно экспрессируются во время развития мужских половых клеток млекопитающих [19]. Недавно было обнаружено, что они обладают рядом особенностей, схожими с мутациями онкогистонов, которые могут приводить к дестабилизации нуклеосом [20]. Кроме того, что «короткие» H2A обладают самой низкой консервативностью (рис. 3), у них обнаружены значимые вариации в аминокислотных последовательностях. Например, все варианты «коротких» H2A имеют укороченные C-концевой хвост и структурный домен «docking domain», характерный для всех белков семейства H2A. При этом H2A.Q имеет наиболее короткий хвост (рис. 3). Важно отметить, что в аминокислотных последовательностях гистонов группы «коротких» H2A имеются вариации в сайтах связывания с ДНК и в регионе «acidic patch», который отвечает за стабильность хроматина [22]. Например, у варианта H2A.P отсутствуют оба аргинина, участвующие в связи с малой бороздкой ДНК, а также четыре из шести остатков «acidic patch», причем один из них заменен на положительно заряженную аминокислоту – аргинин. Вариант H2A.L лишен всего лишь трех остатков из региона «acidic patch». В аминокислотных после-



довательностях H2A.B обнаружены всего два из шести остатков «acidic patch» и один из аргининов — в сайте связывания с ДНК. Интересно отметить, что наблюдаются отличия между двумя кластерами последовательностей в группе гистоновых вариантов H2A.B, которые были получены в ходе кластерного анализа «коротких» гистонов семейства H2A. Например, можно выделить две позиции, находящиеся в регионе «acidic patch». В одной из них оказалась положительно заряженная аминокислота в обоих кластерах H2A.B (E127K и E127R), в другой — только в одном из кластеров (E159R). Аминокислотные последовательности недавно обнаруженного H2A.Q также имеют различия между выявленными кластерами [6]. В одном из них отсутствуют все кислотные остатки в регионе «acidic patch» (причем в трех позициях встречаются положительно заряженные аминокислоты: E122K, E130K, E158K) и все аргинины в сайтах связывания с ДНК. В двух других кластерах последовательностей у варианта H2A.Q можно увидеть только один из шести остатков «acidic patch».

Заключение

В результате анализа разнообразия гистоновых белков семейства H2A мы показали, что гистоны семейства H2A обладают значительным разнообразием первичных последовательностей. Охарактеризованные нами различия, которые встречаются

у гистоновых вариантов H2A, влияют на структурные, физико-химические и функциональные свойства гистонов и нуклеосом. Проведенный в нашей работе анализ также показал, что каждое отдельное подсемейство обладает специфичным набором характеристик, некоторые из которых находятся в участках взаимодействия с другими гистонами и ДНК, что может оказывать влияние на стабильность нуклеосомы. При этом важную роль в динамике хроматина могут играть не только значительные отличия (например, наличие доменов или мотивов), но и отдельные небольшие вариации в аминокислотных последовательностях (например, вариации, расположенные в «acidic patch»). Проведенный в данной работе анализ группы «коротких» гистонов H2A, продемонстрировал, что некоторые последовательности вариантов подсемейства (H2A.B, H2A.P, H2A.Q, H2A.L) образуют отдельные клады на филогенетическом дереве, возникшие, вероятно, в ходе различных эволюционных событий и обладающие характерными отличиями между первичными последовательностями.

Работа выполнена при поддержке гранта Министерства науки и высшего образования № 075-15-2021-1062. Исследования проводили без использования животных и без привлечения людей в качестве испытуемых. Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Talbert P.B., Henikoff S. Histone variants at a glance. *J. Cell Sci.* 2021;134(6):jcs244749.
2. Draizen E.J., Shaytan A.K., Mariño-Ramírez L., Talbert P.B., Landsman D., Panchenko A.R. HistoneDB 2.0: a histone database with variants—an integrated resource to explore histones and their variants. *Database.* 2016;2016:baw014.
3. Malik H.S., Henikoff S. Phylogenomics of the nucleosome. *Nat. Struct. Mol. Biol.* 2003;10(11):882–891.
4. Kustatscher G., Hothorn M., Pugieux C., Schefzek K., Ladurner A.G. Splicing regulates NAD metabolite binding to histone macroH2A. *Nat. Struct. Mol. Biol.* 2005;12(7):624–625.
5. Yelagandula R., Stroud H., Holec S., Zhou K., Feng S., Zhong X., Muthurajan M., Nie X., Kawashima T., Groth M., Luger K., Jacobsen S., Berger F. The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis*. *Cell.* 2014;158(1):98–109.
6. Tanaka M., Hennebold J.D., Macfarlane J., Adashi E.Y. A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog. *Development.* 2001;128(5):655–664.
7. Jiang D., Borg M., Lorković Z.J., Montgomery S.A., Osakabe A., Yelagandula R., Axelsson E., Berger F. The evolution and functional divergence of the histone H2B family in plants. *PLOS Genetics.* 2020;16(7):e1008964.
8. Strickland M., Strickland W.N., Brandt W.F., Von Holt C., Wittmann-Liebold B. The complete amino-acid sequence of histone H2B(3) from sperm of the sea urchin *Parechinus angulosus*. *Eur. J. Biochem.* 1978;89(2):443–452.
9. Kawashima T., Lorković Z.J., Nishihama R., Ishizaki K., Axelsson E., Yelagandula R., Yelagandula R., Kohchi T., Berger F. Diversification of histone H2A variants during plant evolution. *Trends Plant Sci.* 2015;20(7):419–425.
10. Ueda K., Suzuki M., Ono M., Ide N., Tanaka I., Inoue M. Male gametic cell-specific histone gH2A gene of *Lilium longiflorum*: genomic structure and promoter activity in the generative cell. *Plant Mol. Biol.* 2005;59(2):229–238.
11. Molaro A., Young J.M., Malik H.S. Evolutionary origins and diversification of testis-specific short histone H2A variants in mammals. *Genome Res.* 2018;28(4):460–473.
12. Churchill M.E., Suzuki M. “SPKK” motifs prefer to bind to DNA at A/T-rich sites. *EMBO J.* 1989;8(13):4189–4195.
13. Zambrano-Mila M.S., Aldaz-Villao M.J., Casas-Mollano J.A. Canonical histones and their variants in plants: evolution and functions. *Epigenetics in plants of agronomic importance: fundamentals and applications*. Eds. R. Alvarez-Venegas, C. De-la-Peña, J. Casas-Mollano. Cham: Springer; 2019:185–222.
14. Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007;23(21):2947–2948.
15. Wheeler T.J., Kececioglu J.D. Multiple alignment by aligning alignments. *Bioinformatics.* 2007;23(13):i559–i568.
16. Edgar R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–1797.

17. Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 2010;59(3):307–321.

18. Beitz E. TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinformatics.* 2000;16(2):135–139.

19. Jiang X., Soboleva T.A., Tremethick D.J. Short histone H2A variants: small in stature but not in function. *Cells.* 2020;9(4):867.

20. Chew G.L., Bleakley M., Bradley R.K., Malik H.S., Henikoff S., Molaro A., Sarthy J. Short H2A histone variants are expressed in cancer. *Nat. Commun.* 2021;12(1):490.

21. Dryhurst D., Ishibashi T., Rose K.L., Eirín-López J.M., McDonald D., Silva-Moreno B., Veldhoen N., Helbing C.C., Hendzel M.J., Shabanowitz J., Hunt D.F.,

Ausió, J. Characterization of the histone H2A.Z-1 and H2A.Z-2 isoforms in vertebrates. *BMC Biology.* 2009;7(1):86.

22. Shaytan A.K., Landsman D., Panchenko A.R. Nucleosome adaptability conferred by sequence and structural variations in histone H2A-H2B dimers. *Curr. Opin. Struct. Biol.* 2015;32:48–57.

23. Millar C.B. Organizing the genome with H2A histone variants. *Biochem. J.* 2013;449(3):567–579.

24. Draker R., Ng M.K., Sarcinella E., Ignatchenko V., Kislinger T., Cheung P. A Combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation. *PLoS Genetics.* 2012;8(11):e1003047.

Поступила в редакцию 18.07.2023

После доработки 17.11.2023

Принята в печать 26.11.2023

RESEARCH ARTICLE

Diversity of H2A histones and their implications for nucleosome structural properties

L. Singh-Palchevskaia , A.K. Shaytan* 

¹*Institute of Bioengineering, Department of Biology, Lomonosov Moscow State University, Leninskie gory 1-12, Moscow, 119234, Russia*

*e-mail: shaytan_ak@mail.bio.msu.ru

Histone proteins are key epigenetic factors, which play an important role in chromatin dynamics and gene activity regulation. They are divided into two broad classes: canonical histones and their variants. Canonical histones are expressed mainly during the S-phase of the cell cycle, as they are involved in DNA packaging during cell division. Histone variants are histone genes that are expressed and regulate chromatin dynamics throughout the cell cycle. Due to the functional and species diversity, various families of histone variants are distinguished. Some proteins may differ slightly from canonical histones, while others, on the contrary, may have many important structural and functional features that affect nucleosome stability and chromatin dynamics. In order to assess the variability of the H2A histone family and their role in nucleosome structure, we performed a bioinformatic analysis of the amino acid sequences of the H2A histone family. The clustering performed by the UPGMA method made it possible to reveal two main subfamilies of H2A proteins: short H2A and other H2A variants demonstrating highly conserved amino acid sequences. We also constructed and analyzed multiple alignments for various H2A histone subfamilies. It is important to note that the proteins of the short H2A subfamily are not only the least conserved within the H2A family, but also have features that significantly affect the structural properties of the nucleosome. In addition, we performed a phylogenetic analysis of short H2A, which resulted in the identification and characterization of individual clades on the phylogenetic tree for the variants H2A.B, H2A.P, H2A.Q, H2A.L.

Keywords: *histone, H2A, histone variants, nucleosome, chromatin, sequence analysis, bioinformatics, epigenetics*

Funding: The research was supported by the Russian Ministry of Science and Higher Education grant No. 075-15-2021-1062.

Сведения об авторах

Сингх-Пальчевская Лаврит — мл. науч. сотр. кафедры биоинженерии биологического факультета МГУ. Тел.: 8-495-939-57-38; e-mail: l.singh@intbio.org; ORCID: <https://orcid.org/0009-0009-1351-6675>

Шайтан Алексей Константинович — докт. физ.-мат. наук, доцент кафедры биоинженерии биологического факультета МГУ. Тел.: 8-495-939-57-38; e-mail: shaytan_ak@mail.bio.msu.ru; ORCID: <https://orcid.org/0000-0003-0312-938X>