

## КОНСЕРВАТИВНЫЕ УЧАСТКИ РЕГУЛЯТОРНЫХ МОДУЛЕЙ РАННИХ ГЕНОВ ДРОЗОФИЛЫ, ВКЛЮЧАЮЩИЕ ГОМОТИПИЧЕСКИЕ САЙТЫ СВЯЗЫВАНИЯ ФАКТОРОВ ТРАНСКРИПЦИИ, РАСПОЛОЖЕНЫ С ПЕРИОДОМ 84 НУКЛЕОТИДА, СООТВЕТСТВУЮЩИМ ДЛИНЕ ВИТКА СУПЕРСПИРАЛИ ДНК НУКЛЕОСОМЫ

© 2016 г. А.П. Лифанов, В.Ю. Макеев\*, Н.Г. Есипова

*Институт молекулярной биологии им. В.А. Энгельгардта РАН, 119991, Москва, ул. Вавилова, 32;*

*\*Институт общей генетики им. Н.И. Вавилова РАН, 119991, Москва, ул. Губкина, 3*

*E-mail: johnnie\_me@list.ru*

Поступила в редакцию 24.09.15 г.

Изучены периодичности расположения нуклеотидов на консервативных участках регуляторных модулей 16 генов раннего развития двух видов дрозофилы. Установлено, что взаимное расположение консервативных участков кратно периоду 84 нуклеотида, т.е. соответствует длине витка суперспирали ДНК нуклеосомы. В составе консервативных областей найдены входящие в гомо- и гетеротипические кластеры сайты связывания факторов транскрипции, как правило, находящиеся на расстоянии, соответствующем длине витка суперспирали нуклеосомной ДНК. Высказана гипотеза о возможной роли сайтов связывания факторов транскрипции в регуляции компактизации ДНК по нуклеосомному типу.

*Ключевые слова: энхансер, регуляция транскрипции, сайт связывания регуляторных факторов, гомотипические и гетеротипические кластеры, нуклеосома.*

Исследование строения регуляторных сегментов гена (энхансеров) и определение их местоположения в геномах различных видов живых организмов продолжает оставаться актуальной задачей [1].

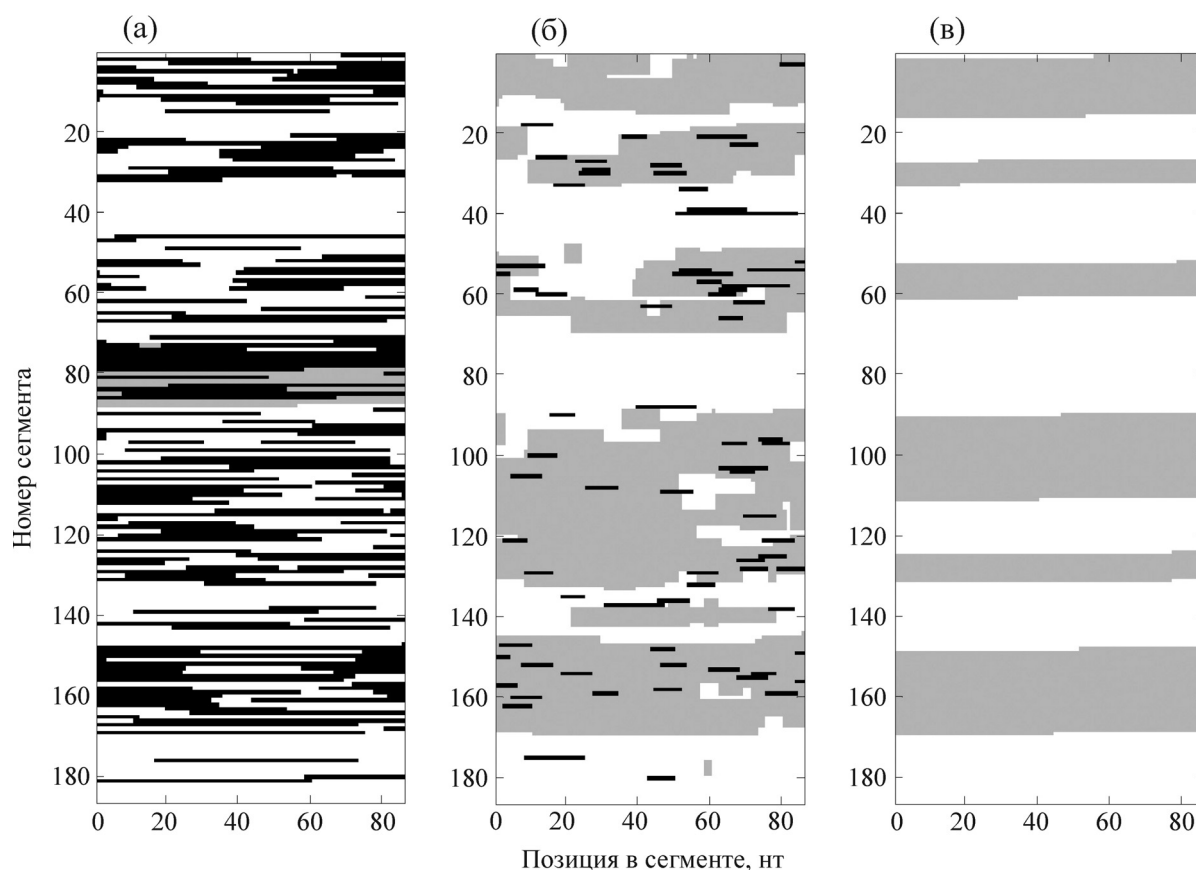
Для локализации энхансеров в геноме широко применяются следующие методы: выделение участков повышенной консервативности – консервативных блоков, определяемых при сравнении геномов достаточно близких видов; поиск мест сгущения (кластеров) сайтов связывания факторов регуляции транскрипции (ССФТ); выявление мест химической модификации гистонов [2]. Для генома дрозофилы места посадки нуклеосом приведены в работе [3], локализация химически модифицированных гистонов определена в работе [4]. Ранее авторами данной работы было показано, что для генома дрозофилы во многих случаях число сайтов связывания одного фактора транскрипции в энхансерах статистически значимо превосходит число таковых в остальной нуклеотидной последовательности локуса, т.е. ССФТ

образуют кластеры. Размер кластера такого типа, названного нами гомотипическим [5], примерно соответствует минимальной наблюдаемой длине энхансера – около 600 нуклеотидов (нт); при этом никаких дополнительных предположений о позиции ССФТ внутри кластера не делается [5]. Были также проанализированы расстояния между соседними ССФТ (как однотипными, так и перекрывающимися сайтами связывания белков разных типов – активаторов и репрессоров транскрипции). При этом наиболее значимыми результатами стали обнаружение периодов в размещении ССФТ, кратных 10,5 нт, и сгущение ССФТ в блоки длиной 50–70 нт, названных композитными элементами [6].

Анализ расположения консервативных блоков в последовательностях энхансеров, проведенный в настоящей работе, имеет целью восполнение пробелов в картине организации хроматина на суб- и нуклеосомных масштабах длин [7,8].

Определение консервативных блоков проводилось с использованием методических приемов, разработанных нами в работе [9]. В данной работе степень консервативности выбиралась

Сокращение: ССФТ – сайты связывания факторов регуляции транскрипции.



**Рис. 1.** Расположение консервативных блоков и сайтов связывания белков – факторов регуляции транскрипции на локусе гена *even-skipped* генома *D. melanogaster*. Карты энхансеров и консервативных блоков локуса представлены в виде сегментов длиной 86 нт, отображенных один под другим (начало локуса – сверху). (а) – Карта консервативных блоков (черные отрезки), серым выделен кодирующий сегмент, (б) – области сгущения консервативных блоков выделены серым цветом, черные отрезки – сайты связывания пяти факторов регуляции транскрипции, (в) – карта энхансеров локуса.

такой, чтобы в найденные блоки входило бы большинство известных ССФТ.

Для генома *D. melanogaster* положение консервативных участков определено для 16 генов раннего развития эмбриона дрозофилы [10]: *bid*, *ems*, *en*, *eve*, *ftz*, *gsb*, *gt*, *hairy*, *hb*, *kni*, *kr*, *otd*, *prd*, *runt*, *sal*, *tll*. Наборы нуклеотидных последовательностей для поиска ССФТ и положение известных регуляторных и кодирующих сегментов в локусах перечисленных генов *Drosophila* взяты из работы [5].

Для графического представления найденных консервативных участков для каждого из перечисленных генов была построена карта локуса – последовательность чисел, содержащая единичные значения в позициях, лежащих внутри участков, и нули – в остальных местах. На карту локуса также могут наноситься положения ССФТ (рис. 1).

Общее число найденных консервативных блоков равнялось 662; гистограмма длин кон-

сервативных блоков приведена на рис. 2. Длины консервативных блоков лежат в диапазоне от 18 до 200 нт; максимум на гистограмме достигается для длин ~55–60 нт.

Для установления периодичности расположения консервативных участков карты локусов разделяются на последовательные сегменты длиной 86 нт; эта величина близка к 84 нт – длине одного витка двойной спирали ДНК, окружающей гистоновое ядро нуклеосомы [7]. Выбор периода 86 нт определяется как половина от 172 нт – периода нуклеосомного повтора, характерного для данного набора генов. Полученные сегменты располагаются один над другим. Результат такой сегментации для локуса гена *even-skipped* генома *D. melanogaster* приведен на рис. 1а.

Протяженные области достаточно большого числа найденных консервативных блоков, располагающихся в соседних сегментах, лежат друг над другом. Число входящих в пары консер-

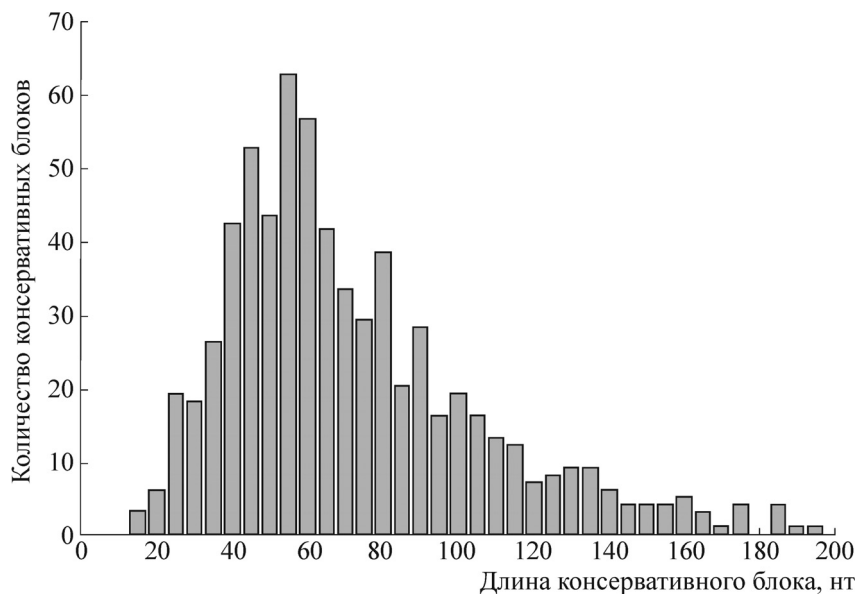


Рис. 2. Гистограмма длин консервативных блоков.

вативных блоков, содержащих такие блоки длиной более 18 нт и, следовательно, имеющих области, расположенные на расстоянии 84 нт, составляет около 60% от общего числа консервативных блоков, найденных в перечисленных генах. В районе расположения энхансеров (см. карту энхансеров, рис. 1в) обычно находятся места сгущения как одиночных консервативных блоков, так и консервативных блоков, образующих пары. Чтобы подчеркнуть эффект коррелированности расположения энхансеров и таких мест сгущения, массив сегментов (рис. 1а) выбранного локуса суммировался по шести соседним строкам; выделялись позиции в сегменте, для которых величина суммы составляла три и более. Такой выбор параметров фильтрации оказался оптимальным. Кодирующий сегмент локуса, содержащий протяженные консервативные участки (отмечен серым цветом на рис. 1а), не рассматривался.

Полученная в результате карта областей сгущения консервативных блоков приведена на рис. 1б. На ней также отмечены позиции сайтов связывания нескольких факторов регуляции транскрипции, определенных по методике, описанной в работе [5].

Положение позиций сгущения консервативных блоков (рис. 1б) достаточно хорошо коррелирует с картой энхансеров (рис. 1в) в случаях подробно изученных генов, для которых набор энхансеров, по-видимому, является практически полным и границы энхансеров определены достаточно точно. Однако даже в случае наиболее изученного гена *even-skipped* сравнение двух

карт показывает, что хотя иногда сгущение консервативных блоков и наблюдается, однако энхансер отсутствует: например, над второй полосой карты энхансеров (рис. 1в), под третьей и между четвертой и пятой полосами. Можно высказать предположение, что означенные области также могут являться энхансерами и иметь отношение к регуляции экспрессии до сих пор не описанными факторами транскрипции. Коэффициент корреляции двух карт (рис. 1б и 1в) составляет около 0,5; при удалении означенных областей энхансеров он приближается к единице. Для других генов набора совпадение карт не такое хорошее, что может быть связано с их недостаточной изученностью.

Число консервативных блоков значительно превосходит как число ССФТ локуса, так и число мест посадки нуклеосом, и привлечение к рассмотрению консервативных блоков, занимающих по длине промежуточное положение между ССФТ и ДНК нуклеосом, позволяет уточнить локализацию регуляторных элементов в составе энхансера. Большинство ССФТ располагается в местах сгущения консервативных блоков, следующих с периодом 84 нт (рис. 1б); при этом места размещения блоков нескольких соседних ССФТ, входящих в композитные элементы, также согласуются с периодом 84 нт. Композитные элементы, таким образом, оказываются подмножеством консервативных блоков – это консервативные блоки, содержащие несколько ССФТ. Период 10,5 нт и длина композитных элементов коррелируют с параметрами нуклеосомной ДНК.

Если на основе обсуждаемых последовательностей сформированы нуклеосомы, то при наличии в расположении ССФТ периода 84 (что видно по расположению ССФТ в гомотипических кластерах, см. рис. 1), благодаря компактизации ДНК локуса при формировании нуклеосом, произойдет сближение ССФТ в пространстве нуклеосомной частицы. Также при формировании нуклеосом имеет место сближение ССФТ всего энхансера, в том числе в гомотипических кластерах, благодаря компактизации ДНК локуса. Полученные нами результаты позволяют высказать предположение, что функциональная роль гомотипических кластеров состоит в регуляции компактизации на нуклеосомном уровне организации ДНК.

Можно также полагать, что кооперативное действие таких наборов ССФТ способствует образованию нуклеосом.

Можно высказать гипотезу, что функциональная роль гомотипических кластеров состоит в регуляции компактизации ДНК в форме нуклеосом.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных

исследований (грант № 15-04-99605а) и Программы Президиума РАН «Молекулярная и клеточная биология».

#### СПИСОК ЛИТЕРАТУРЫ

1. R. C. Hardison and J. Taylor, *Nat. Rev. Genet.* **13** (7), 469 (2012).
2. J. Touchman, *Nature Education Knowledge* **3** (10), 13 (2010).
3. T. N. Mavrich, et al., *Nature* **453**, 358 (2008).
4. P. V. Kharchenko, et al., *Nature* **471** (7339), 480 (2011).
5. A. P. Lifanov, V. J. Makeev, A. G. Nazina, and D. A. Papatsenko, *Genome Res.* **13**, 579 (2003).
6. V. J. Makeev, A. P. Lifanov, A. G. Nazina, and D. A. Papatsenko, *Nucl. Acids Res.* **31** (20), 6016 (2003).
7. G. Felsenfeld and M. Groudine, *Nature* **421** (6921), 448 (2003).
8. L. Marino-Ramirez, M. G. Kann, B. A. Shoemaker, and D. Landsman, *Expert Rev. Proteomics* **2**, 719 (2005).
9. А. П. Лифанов, В. Ю. Makeev и Н. Г. Есипова, *Биофизика* **60** (1), 15 (2015).
10. L. Wolpert and C. Tickle, *Principles of Development* (Oxford University Press, 2002).

## Conservative Sections of Transcription Regulatory Modules of Early *Drosophila* Genes Including Homotypic Transcription Factor Binding Sites are Situated with 84 Nucleotide Period Corresponding to Nucleosomal DNA Superhelix Turn Length

A.P. Lifanov\*, V.J. Makeev\*\*, and N.G. Esipova\*

\**Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, Moscow, 119991 Russia*

\*\**Vavilov Institute of General Genetics, Russian Academy of Sciences, ul. Gubkina 3, Moscow, 119991 Russia*

The periodicity in arrangement of nucleotides has been studied in conservative regions in regulatory modules of 16 early development genes of two *Drosophila* species. It was found that mutual disposition of conservative segments is a multiple of the length of one coil of nucleosomal DNA superhelix. Homotypic and heterotypic clusters of transcription factor binding sites were observed in the conservative regions. These sites are as a rule arranged at a distance corresponding to the length of the coil of superhelical DNA of nucleosome. It was suggested that these transcription factor binding sites can be involved in regulation of DNA compacting according to nucleosome type.

*Key words: enhancer, transcription regulation, transcription factor binding site, homotypic and heterotypic clusters of sites, nucleosome*