

УДК 577.213.3+576.316.352+595.773.4

АРХИТЕКТУРА ПРОМОТОРОВ ГЕНОВ “ДОМАШНЕГО ХОЗЯЙСТВА” В МЕЖДИСКАХ ПОЛИТЕННЫХ ХРОМОСОМ

Drosophila melanogaster

Т. Ю. Зыкова^{1,*}, В. Г. Левицкий^{2,3}, академик РАН И. Ф. Жимулёв^{1,3}

Поступило 23.11.2018 г.

Впервые изучили молекулярно-генетическую организацию междисков политенных хромосом, локализованных одновременно на молекулярной и цитологической картах генома *Drosophila melanogaster*. Большая часть изученных междисков содержала один ген, имеющий единственный старт транскрипции, остальные междиски — по одному гену с несколькими альтернативными промоторами, два и более однонаправленных гена и гены с расположением “голова-к-голове”. Также обнаружили сложноорганизованные междиски, в которых локализовались три и более гена, имеющие как однонаправленную, так и двунаправленную ориентацию. Инсуляторные белки, ORC, *P*-инсерции, гиперчувствительные к ДНКазе I сайты и другие характеристики открытого хроматина располагались в промоторной области генов, находившихся в междисках. Эта область — критическая для образования междиска, участка открытого хроматина, в котором совмещаются процессы инициации транскрипции и репликации генов.

Ключевые слова: дрозофила, организация генома, политенные хромосомы, междиски, промоторы.

DOI: <https://doi.org/10.31857/S0869-56524851104-109>

В результате последних крупномасштабных исследований состояний хроматина в объёме всего генома *D. melanogaster* [1] было сформировано представление о локализации генов с повсеместной активностью в междисках политенных хромосом [2]. Гены, которые специфически активны на определённых стадиях и в индивидуальных тканях (гены развития), как оказалось, расположены в плотно упакованных дисках, которые ранее были названы дисками интеркалярного гетерохроматина [3]. Ранее нами было показано, что на уровне полного генома хроматин, соответствующий междискам, имеет сложную структуру в плане генетической организации [2]. В частности, в междисках были найдены одно- и двунаправленные промоторы, гены, целиком расположенные в междисках, и гены с разным числом стартов инициации транскрипции, поэтому встал вопрос о полном исследовании всех междисков, локализованных на молекулярной и цитологической картах генома *D. melanogaster*, на предмет организации типов промоторов в них.

В настоящей работе получены новые данные о топологии или расположении белков и различных характеристик хроматина в пределах последовательностей ДНК междисков. Мы установили, что междиски, содержащие дву- и однонаправленные гены, имеют различную организацию промоторной области, которая, по всей видимости, является критической для образования участка открытого хроматина, который мы описываем как междиск в политенных хромосомах *D. melanogaster*.

На основе данных международного проекта modENCODE и наших исследований [4] была разработана модель четырёх состояний хроматина [2]. Используя сведения о локализации белков, характерных для генетически активного хроматина, мы разработали стереотип его классификации по принципу степени компактизации от минимального (aquamarine, соответствует междискам) до максимального (ruby, соответствует чёрным плотным дискам). Ещё два состояния lazurite и malachite в основном соответствуют кодирующим частям генов [2, 4] и краям блоков генов развития [5].

Согласно расчётам существует пять тысяч междисков в хромосомах *D. melanogaster*, которые содержат примерно 5% ДНК эухроматиновой части генома [6, 7]. Из них только 33 были точно картированы на молекулярной и цитологической картах генома *D. melanogaster*. Поскольку междиски обладают общими свойствами, такими как локализация 5'-UTR генов с повсеместной активностью, обога-

¹ Институт молекулярной и клеточной биологии
Сибирского отделения Российской Академии наук,
Новосибирск

² Институт цитологии и генетики
Сибирского отделения Российской Академии наук,
Новосибирск

³ Новосибирский национальный исследовательский
государственный университет

*E-mail: vatolina@mcb.nsc.ru

шение генами с двунаправленной ориентацией регуляторных областей, локализацией инсуляторных белков (CHRIZ, BEAF-32 и других) и соответствуют хроматину aquamarine [2], анализ их генетической структуры и архитектуры представляет первостепенный интерес.

Ранее мы сравнили организацию и функции генов, полностью локализованных в участке открытого хроматина — междиске, и генов, занимающих две структуры — междиск и серый диск [8].

В настоящей работе мы детально исследовали молекулярно-генетическую организацию последовательностей 33 междисков, локализованных на молекулярной карте *D. melanogaster*. В табл. 1 указано число доменов хроматина *aquamarine*, соответствующих междискам. Два междиска из 33 являются сложными и содержат в пределах своей последовательности несколько участков хроматина *aquamarine*. Для анализа мы использовали версию FlyBase r.5.57 *D. melanogaster*, которая содержит 29761 транскрипт 13753 белоккодирующих генов и междиски с 58 генами и 97 транскриптами с уникальными стартами.

В нашем исследовании мы обнаружили, что междиски имели разную генетическую организацию (табл. 1). Большая часть изученных междисков содержала один ген. В основной части этих междисков локализовались гены, имевшие один старт транскрипции, остальные междиски содержали по одному гену с несколькими альтернативными промоторами (табл. 1, группы I и II). Остальные междисковые последовательности содержали два и более гена. Группа III — это междиски с двумя однонаправленными генами и несколькими транскрипционными стартами. Группа IV — междиски с разнонаправленными генами и группа V — сложноорганизованные междиски, в которых находились три и более гена, имевшие как одно-, так и двунаправленную ориентацию.

Ранее мы показали [9], что точки инициации транскрипции генов расположены в основном в районе центра междисковых последовательностей, а граница хроматина *aquamarine*, соответствующего междискам на молекулярной карте, практически всегда локализована в первом интроне междискового гена.

Для анализа топологии мы выбрали междиски из двух контрастных групп I и IV, содержавшие гены с одним стартом и разнонаправленные гены соответственно (табл. 1). Мы оценили размеры междисков, границы которых были локализованы, при помощи модели четырёх состояний хроматина [2]. Длина междисков, содержавших транскрипты с одним стартом, в среднем составила 3,5 т.п.н., а их

Таблица 1. Локализация генов в междисках

Группа	№	Схема генетической организации междиска	Число генов в междиске	Число стартов инициации транскрипции	Число меж-дисков
I	1		1	1	12
II	2		1	2	4
	3		1	3	2
	4		1	5	1
Всего					19
III	1		2	2	1
	2		2	3	2
Всего					3
IV	1		2	2	1
	2		2	2	2
	3		2	3	2
	4		2	4	2
	5		2	5	1
	6		2	6	1
Всего					9
V	1		3	3	3
	2		3	6	1
	3		4	7	1
Всего					5

размеры менялись от 1 до 15,4 т.п.н. (рис. 1а). Размеры междисков, в которых находились разнонаправленные гены, составили в среднем 3,4 т.п.н. (1,6–7 т.п.н.). Расстояние между промоторами в этих междисках — в среднем 0,7 т.п.н. (рис. 1б). Такие большие размеры четырёх междисков из 33 изученных объясняются тем, что используемая нами модель идентифицирует первый длинный интрон активного гена в хроматине *aquamarine*. Однако все элементы генома и белки, характерные для междисков, рас-

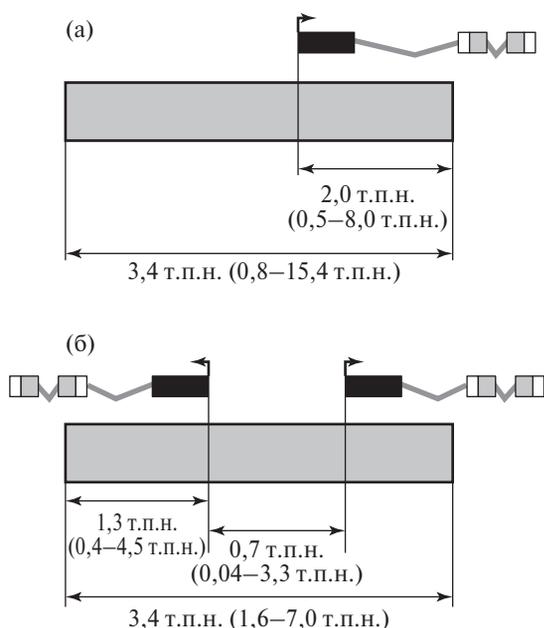


Рис. 1. Расположение генов в междисках. Здесь и на рис. 2 и 3: (а) — междиски, содержащие гены с одним стартом инициации транскрипции (группа I в табл. 1). (б) — междиски, содержащие двунаправленные гены (группа IV).

полагаются в этих доменах в районе промоторной области генов.

Недавно нами было показано [10], что искусственно привлечённые в участок хромосомы белки-инсуляторы CHRIZ и CTCF формируют новый междиск. Мы изучили [11] распределение инсуляторных белков CHRIZ и BEAF-32, а также РНК-полимеразы II в последовательностях междисков, содержащих один ген (группа I) и разнонаправленные гены (группа IV). Последовательности хроматина aquamagine, соответствующие междискам, были выравнены относительно стартов инициации транскрипции генов для междисков I группы и относительно центра между стартами инициации транскрипции для междисков IV группы. Несмотря на то что средний размер междисков составил 3,4 т.п.н. (рис. 1а, б), белки занимали не всю последовательность, а были расположены локально в районе промоторной области генов (рис. 2). Можно видеть, что в последовательностях междисков, содержащих один ген, инсуляторные белки CHRIZ и BEAF-32 локализовались непосредственно на старте инициации транскрипции генов, а РНК-полимераза II была сдвинута в область 200 п.н. выше старта инициации транскрипции (рис. 2а). В междисках, содержащих два разнонаправленных гена, белки BEAF-32 и CHRIZ располагались между стартами инициации транскрипции генов, а РНК-полиме-

раза II имела два пика в районе промоторной области генов (рис. 2б).

В промоторах многих генов присутствует паузировавшая РНК-полимераза II, которая генерирует короткие некодирующие РНК длиной 25–60 п.н. Наличие такой приостановленной формы полимеразы способствует созданию доступной конформации хроматина в регуляторной области гена и связыванию факторов транскрипции [12].

В работе [13] короткие некодирующие РНК, производные паузировавшей формы РНК-полимеразы II, были картированы в геноме дрозофилы и определены их 5'- и 3'-концы, поэтому известны их точные последовательности. Было показано, что такие РНК характерны для доменов открытого aquamagine-хроматина (междисков) дрозофилы [9]. Мы исследовали топологию таких некодирующих РНК, а также белков репликативного комплекса (ORC2) [14], ДНКазу I гиперчувствительных сайтов [1], инсерций *P*-элементов (FlyBase г.5.57) и “широких” промоторов [15] в изучаемых междисках.

Все перечисленные структуры находились в области старта инициации транскрипции генов междисков, содержащих гены с одним стартом транскрипции, и имели два независимых пика в междисках с разнонаправленными генами (рис. 3).

Таким образом, несмотря на то что размер индивидуальных междисков составляет в среднем около 2 т.п.н. [6, 7], а по нашим оценкам — около 3,4 т.п.н., инсуляторные белки и другие ранее описанные структуры, характерные для междисков [2], имели узкую локализацию в районе регуляторной области генов, расположенных в междисках. Обогащение хроматина aquamagine белками ORC и 5'-некодирующими областями генов преимущественно с повсеместной активностью [2, 9] подтвердил тот факт, что междиски совмещают в себе процессы инициации транскрипции и репликации генов. По всей вероятности, регуляторная область генов является определяющей в формировании домена открытого декомпактизированного хроматина междисков.

Источник финансирования. Работа поддержана грантом Российского научного фонда № 14–14–00934.

СПИСОК ЛИТЕРАТУРЫ

1. Kharchenko P.V., Alekseyenko A.A., Schwartz Y.B., et al. Comprehensive Analysis of the Chromatin Landscape in *Drosophila melanogaster* // Nature. 2011. V. 471. № 7339. P. 480–485.
2. Zhimulev I.F., Zykova T.Y., Goncharov F.P., Khoro-shko V.A., Demakova O.V., Semeshin V.F., Pok-holkova G.V., Boldyreva L.V., Demidova D.S.,

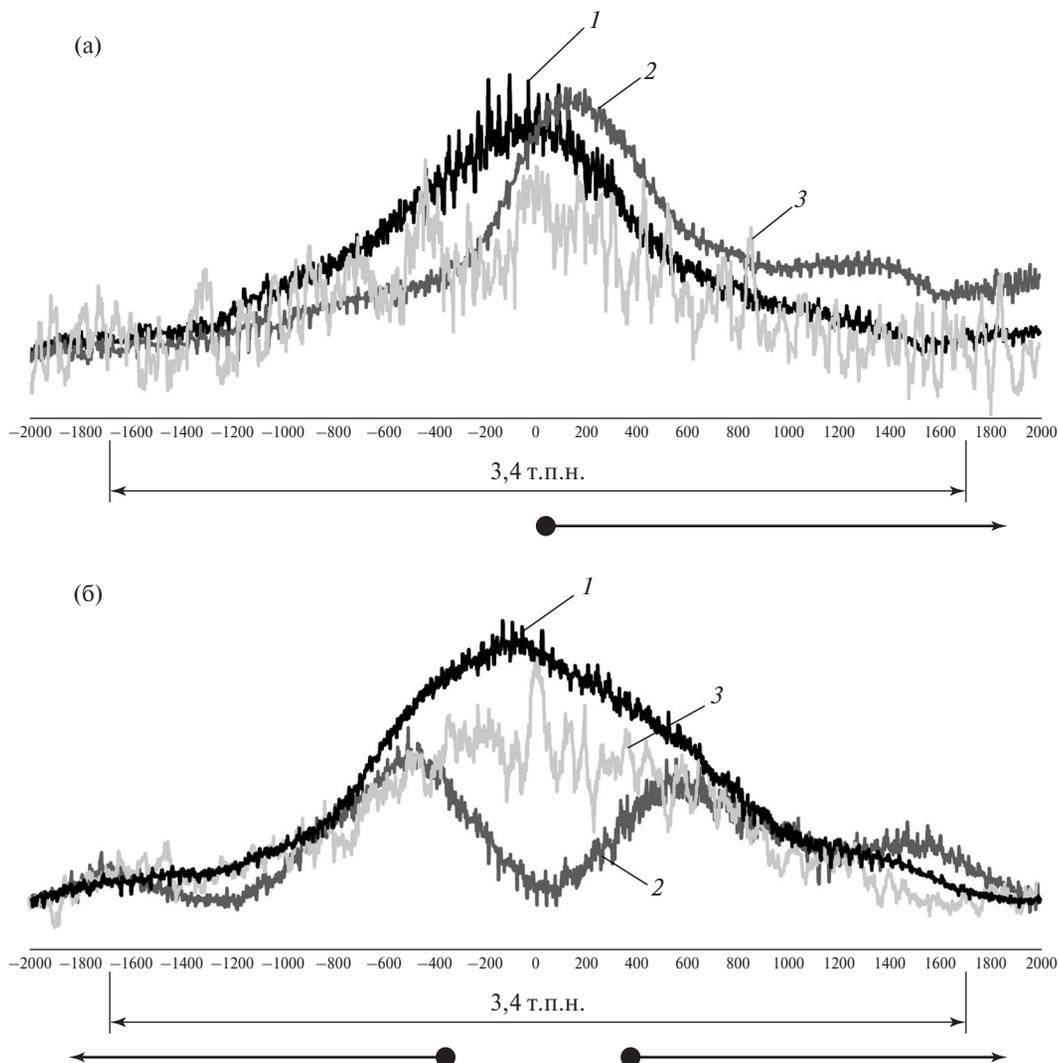


Рис. 2. Локализация белков в последовательностях ДНК, соответствующих междискам политенных хромосом. Здесь и на рис. 3 точка “0” по оси абсцисс соответствует старту инициации транскрипции генов. Цифрами обозначены средняя интенсивность связывания белков в культуре клеток S2 (данные modENCODE): 1 — CHRIZ (CHROMATOR), 2 — РНК-полимераза II, 3 — BEAF-32.

- Babenko V.N., Demakov S.A., Belyaeva E.S.* Genetic Organization of Interphase Chromosome Bands and Interbands in *Drosophila melanogaster* // PLoS One. 2014. V. 9. № 7. e101631.
- Zhimulev I., Belyaeva E., Makunin I., Pirrotta V., Volkova E., Alekseyenko A.* Influence of the *SuUR* Gene on Intercalary Heterochromatin in *Drosophila melanogaster* Polytene Chromosomes // Chromosoma. 2003. V. 111. № 6. P. 377–398.
 - Vatolina T.Yu., Boldyreva L.V., Demakova O.V., Demakov S.A., Kokoza E.B., Semeshin V.F., Babenko V.N., Goncharov F.P., Belyaeva E.S., Zhimulev I.F.* Identical Functional Organization of Cell Line Interphase and Polytene Chromosomes in *Drosophila melanogaster* // PLoS One. 2011. V. 6. № 10. e25960.
 - Khoroshko V.A., Levitsky V.G., Zykova T.Y., Antonenko O.V., Belyaeva E.S., Zhimulev I.F.* Chromatin Heterogeneity and Distribution of Regulatory Elements in the Late-Replicating Intercalary Heterochromatin Domains of *Drosophila melanogaster* Chromosomes // PLoS One. 2016. V. 11. № 6. e0157147.
 - Zhimulev I.F., Belyaeva E.S., Semeshin V.F., Koryakov D.E., Demakov S.A., Demakova O.V., Pokholkova G.V., Andreyeva E.N.* Polytene Chromosomes: 70 Years of Genetic Research // Int. Rev. Cytol. 2004. V. 241. P. 203–275.
 - Beermann W.* Chromomeres and Genes // Results and Problems in Cell Differentiation. 1972. V. 4. P. 1–33.
 - Зыкова Т., Попова О., Хорошко В., Левицкий В., Лавров С., Жимулев И.* Генетическая организация доменов открытого типа, расположенных в междисках политенных хромосом дрозофилы // ДАН. 2018. Т. 483. № 1. С. 98–102.
 - Zykova T., Levitsky V., Belyaeva E., Zhimulev I.* Polytene Chromosomes — A Portrait of Functional Orga-

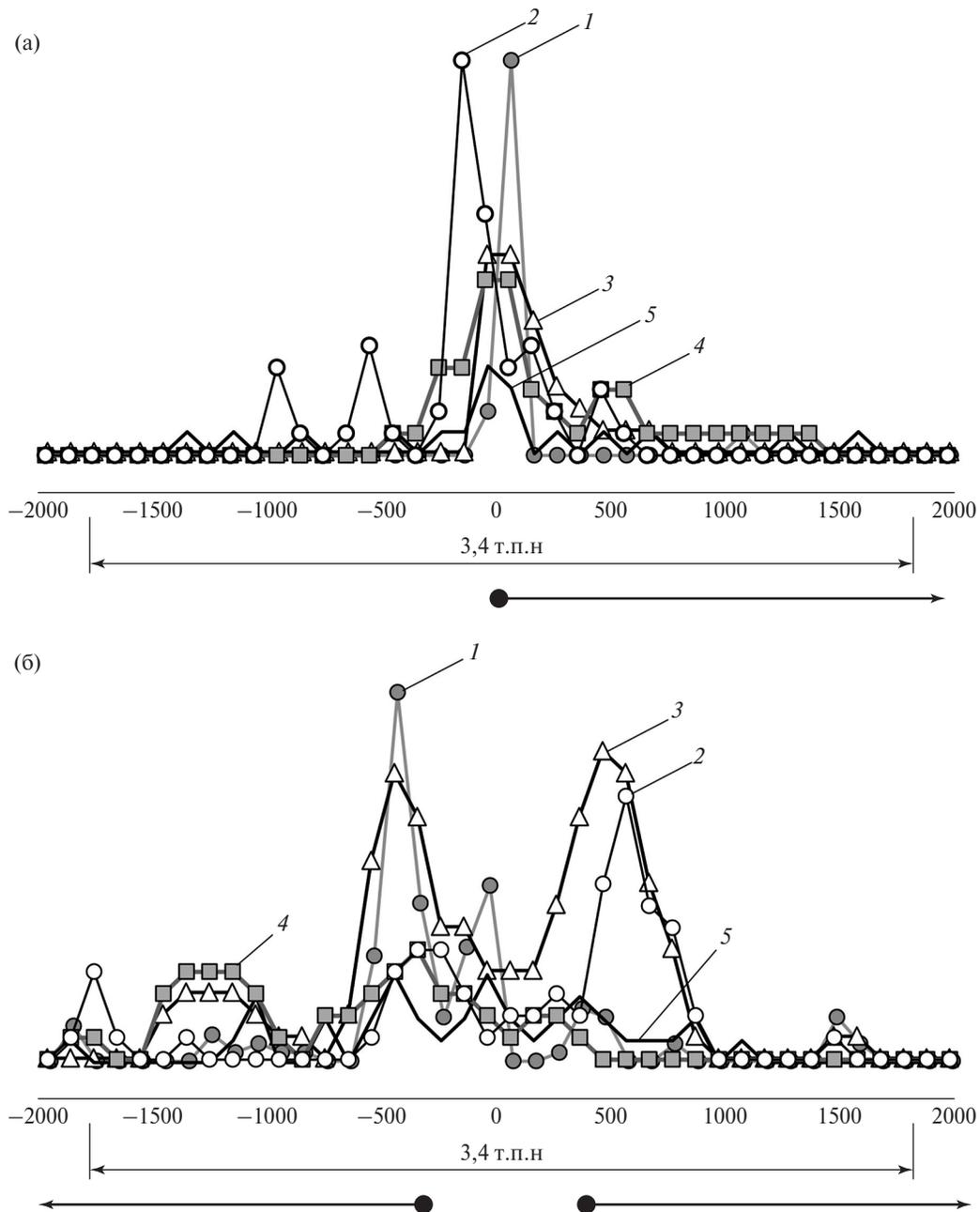


Рис. 3. Локализация элементов генома в последовательностях ДНК, соответствующих междискам политенных хромосом. Цифрами обозначено среднее число аннотаций характеристик в исследуемых междисках: 1 — короткие РНК, производные паузированной формы РНК-полимеразы II в эмбрионах *D. melanogaster* [13]; 2 — инсерции *P*-элементов (FlyBase г.5.57); 3 — широкие (Broad) промоторы [15]; 4 — белки репликативного комплекса (ORC2) в культуре клеток S2 [14]; 5 — гиперчувствительные к ДНКазе I сайты (DHS) в культуре клеток S2 [1].

- nization of the *Drosophila* Genome // Curr. Genomics. 2018. V. 19. № 3. P. 179–191.
10. Pokholkova G., Demakov S., Andreenkov O., Andreenkova N., Volkova E., Belyaeva E., Zhimulev I. Tethering of CHROMATOR and dCTCF Proteins Results in Decompaction of Condensed Bands in the *Drosophila melanogaster* Polytene Chromosomes but Does Not Affect Their Transcription and Replication Timing // PLoS One. 2018. V. 13. № 4. e0192634.
 11. Vogelmann J., Le Gall A., Dejardin S., Allemand F., Gamot A., Labesse G., Cuvier O., Nègre N., Cohen-Gonsaud M., Margeat E., Nöllmann M. Chromatin Insulator Factors Involved in Long-Range DNA Interactions and Their Role in the Folding of the *Drosophila* Genome // PLoS Genet. 2014. V. 10. № 8. e1004544.
 12. Henriques T., Gilchrist D., Nechaev S., Bern M., Muse G., Burkholder A., Fargo D., Adelman K. Stable

- Pausing by RNA Polymerase II Provides an Opportunity to Target and Integrate Regulatory Signals // *Mol. Cell*. 2013. V. 52. № 4. P. 517–528.
13. *Nechaev S., Fargo D., dos Santos G., Liu L., Gao Y., Adelman K.* Global Analysis of Short RNAs Reveals Widespread Promoter-Proximal Stalling and Arrest of Pol II in *Drosophila* // *Science*. 2010. V. 327. № 5963. P. 335–338.
 14. *Eaton M.L., Prinz J.A., MacAlpine H.K., Tretyakov G., Kharchenko P.V., MacAlpine D.M.* Chromatin Signatures of the *Drosophila* Replication Program // *Genome Res*. 2011. V. 21. № 2. P. 164–174.
 15. *Hoskins R.A., Landolin J.M., Brown J.B., Sandler J.E., Takahashi H., et al.* Genome-Wide Analysis of Promoter Architecture in *Drosophila melanogaster* // *Genome Res*. 2011. V. 21. № 2. P. 182–192.

ARCHITECTURE OF PROMOTERS OF HOUSE-KEEPING GENES IN POLYTENE CHROMOSOME INTERBANDS OF *Drosophila melanogaster*

T. Yu. Zykova, V. G. Levitsky, Academician of the RAS I. F. Zhimulev

Received November 23, 2018

This is the first study to investigate the molecular-genetic organization of polytene chromosome interbands located on both molecular and cytological maps of *Drosophila* genome. The majority of the studied interbands contained one gene with a single transcription initiation site; the remaining interbands contained one gene with several alternative promoters, two or more unidirectional genes, and “head-to-head” arranged genes. In addition, intricately arranged interbands containing three or more genes in both unidirectional and bidirectional orientation were found. Insulator proteins, ORC, *P*-insertions, DNase I hypersensitive sites, and other open chromatin structures were situated in the promoter region of the genes located in the interbands. This area is critical for the formation of the interband, an open chromatin region in which gene transcription and replication are combined.

Keywords: *Drosophila*, organization of genome, polytene chromosomes, interbands, promoters.