
КРАТКИЕ
СООБЩЕНИЯ

УДК 577.16

Данные были представлены на Юбилейной научной конференции
“Николай Константинович Кольцов и биология XXI века”,
3–8 октября 2022 г., Москва, Институт биологии развития
им. Н.К. Кольцова РАН

РАЗВИТИЕ ИДЕИ Н.К. КОЛЬЦОВА О ГЕНЕТИЧЕСКОЙ ОРГАНИЗАЦИИ МЕЖДИСКОВ ПОЛИТЕННЫХ ХРОМОСОМ *DROSOPHILA MELANOGASTER*

© 2023 г. И. Ф. Жимулев^{а,*}, Т. Ю. Ватолина^а, В. Г. Левицкий^б,
Т. Д. Колесникова^а, А. В. Цуканов^б

^аИнститут молекулярной и клеточной биологии СО РАН, просп. ак. Лаврентьева, 8/2, Новосибирск, 630090 Россия

^бИнститут цитологии и генетики СО РАН, просп. ак. Лаврентьева, 10, Новосибирск, 630090 Россия

*e-mail: zhimulev@mcb.nsc.ru

Поступила в редакцию 27.12.2022 г.

После доработки 24.03.2023 г.

Принята к публикации 04.04.2023 г.

В работе впервые исследовали организацию промоторов генов развития и генов, необходимых для обще клеточных функций – “домашнего хозяйства” клетки в полном геноме *Drosophila melanogaster*. С помощью биоинформатических методов показано, что гены, промоторы которых расположены в междисках политенных хромосом, обогащены функциями, связанными с обще клеточными процессами, тогда как остальная часть генов (примерно половина генома *Drosophila*) связана с узкоспециализированными процессами, происходящими в ходе развития. В промоторной зоне генов “домашнего хозяйства” обнаружено четыре специфических мотива, которые могут присутствовать у разных генов индивидуально или в различных комбинациях. Существенная часть междисковых промоторов не содержит выявленных мотивов. Анализ, проведенный с помощью Gene Ontology, показал, что для отдельных групп междисковых генов, содержащих в промоторах один мотив или их комбинации, характерно выполнение определенных функций.

Ключевые слова: гены развития, гены “домашнего хозяйства”, промоторы генов, мотивы нуклеотидов в промоторах, междиски, диски, политенные хромосомы, *Drosophila melanogaster*

DOI: 10.31857/S0475145023020064, **EDN:** XFKXAP

ВВЕДЕНИЕ

Вскоре после открытия политенных хромосом в 1934 г. началась бурное обсуждение их структуры и генетической организации, и Н.К. Кольцов (Koltzoff, 1934) сделал ряд предположений, которые привлекли внимание исследователей. Он первым высказал гипотезу о том, что хромосомы этого типа состоят из многих нитей, т.е. они многонитчатые (или политенные), кроме того, он предположил, что гены находятся в междисках политенных хромосом. После этого проблемы генетической организации интерфазных хромосом оказались в центре внимания многих генетиков на долгие годы. Высказывали многочисленные гипотезы, из которых хотелось бы упомянуть три: (Crick, 1971; Paul, 1972; Zhimulev, Belyaeva, 1975). В первых двух авторы предположили, что каждый ген занимает диск и соседний с ним междиск, авторы третьей предположили, что гены домашнего

хозяйства и гены развития занимают соответственно междиски и разные типы дисков.

Ранее нами был предложен метод одновременной локализации дисков и междисков на физической карте ДНК и на цитологической карте, что позволило совместить расположение генов разных типов на хромосомах и карте ДНК. На следующем этапе, используя различия в спектрах белков и модификаций гистонов в участках локализации генов разных типов, разработали математическую модель организации хроматина (4НММ) (Zhimulev et al., 2014). С помощью этой модели весь хроматин в интерфазных ядрах клеток дрозофилы был разделен на четыре состояния (названы, как принято по названию оптического цвета – аквамарин, лазурит малахит, рубин). Каждый цвет соответствует степени деконденсации хроматина и открытости его для транскрипции. Самый деконденсированный – аквамарин, самый конденса-

рованный — рубин, соответствуют междискам и черным дискам соответственно, лазурит — это серые разрыхленные диски (Zykova et al., 2018). Благодаря этой модели все гены генома дрозофилы были распределены в две группы — 6562 гена домашнего хозяйства (в литературе также используется термин “гены с повсеместной активностью”), промоторы которых расположены в участках хроматина типа аквамарин (междиски политенных интерфазных хромосом), а тело гена (экзоны и интроны) — в хроматине лазурит (в серых дисках политенных хромосом). Оставшиеся 5664 гена развития расположены в остальной части генома.

В области исследований организации промоторов у генов высших Metazoa достигнуты огромные успехи. Сейчас хорошо известно, что в окрестностях точки инициации транскрипции — Inr, ниже и выше на 100–200 пар нуклеотидов, располагаются специфические мотивы (короткие специфические для различных функции последовательности нуклеотидов), совершенно необходимые для инициации и продолжения транскрипции (Lenhard et al., 2012; Danino et al., 2015; Anderson, Sandelin, 2020; Ramalingam et al., 2021). Тем не менее в понимании структуры промоторов все-таки довольно много неизученного. Во-первых, до сих пор нет исследований полногеномной организации промоторов у генов развития и “домашнего хозяйства”, поскольку до сих пор не было метода, позволяющего разделить эти группы генов, во-вторых, нет четкого полногеномного разделения на гены “домашнего хозяйства” и гены развития, несмотря на то, что сведения об этом уже накапливаются и показано, что промоторы этих групп генов могут различаться. В данной работе мы попытались найти наиболее важные закономерности в организации мотивов в промоторах генов “домашнего хозяйства” и генах развития в полном геноме дрозофилы.

МАТЕРИАЛЫ И МЕТОДЫ

Мы проанализировали 13 574 белок-кодирующих гена *Drosophila melanogaster* из базы данных FlyBase (R 5.57). В анализ были взяты 5'-области генов от –300 до +200 п.н. относительно сайтов начала транскрипции и полностью картированные в доменах состояний хроматина аквамарин, лазурит, малахит и рубин, как описано ранее (Levitsky et al., 2020). Из анализа были удалены все околомоторные участки, которые имели кодирующие последовательностям в пределах анализируемой области. Таким образом, набор околомоторных участков соответствовал 6562, 874, 1628 и 3162 генам аквамарин, лазурит, малахит и рубин соответственно. Мы применили инструмент STREME для *de novo* поиска мотивов в околомоторных участках для каждого состояния хроматина в трех тестах, относительно либо других состояний хроматина, либо набора последовательностей,

содержащего случайные геномные локусы. Мы применили поправку Бонферрони к обогащенным мотивам: значение $p/N < 0.05$, где p означает статистическую значимость мотива, а N обозначает количество мотивов, найденных STREME (<https://meme-suite.org/meme/doc/streme.html>). Использовали инструмент ТОМТОМ (<https://meme-suite.org/meme/tools/tomtomp>) для сравнения обогащенных мотивов с мотивами для известных факторов транскрипции из баз данных CISBP (<http://cisbp.ccb.utoronto.ca/>) и JASPAR (<https://jaspar.uio.no/>).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Воспользовавшись данными модели 4НММ, мы отобрали гены “домашнего хозяйства” и гены развития в полном геноме дрозофилы и затем обработали последовательности нуклеотидов в промоторах обеих групп генов с помощью инструмента STREME (Sensitive, Thorough, Rapid, Enriched Motif Elicitation). Эта программа позволяет выявлять *de novo* обогащение мотивами в определенных районах генома. Изучены околомоторные участки (от –300 до +200 п.н. вокруг точки инициации транскрипции) у 6562 генов из аквамарин хроматина при сравнении их с последовательностями генов контрастного состояния — рубин (3162 генов), а также против генов полного генома дрозофилы и против суммы генов в трех состояниях, исключая аквамарин (5664 гена). Наши результаты показали, что можно выделить достаточно много мотивов, по которым различаются промоторы генов, локализованных в разных состояниях хроматина. Поэтому для ужесточения степени достоверности оценок из этих значений отобрали только те, которые удовлетворяют следующим дополнительным критериям: максимально низкое значение *E-value*, максимальное обогащение по сравнению с фоном, наличие сведений в других базах данных, а также воспроизводимость в разных вариантах сравнений. Оказалось, что этим критериям удовлетворяют только четыре мотива: M1BP (fd19B), BEAF-32 (pnr, DREF), cpr (sna, sage, grh), kni (CG7928, slbo) (рис. 1a). Локализация этих мотивов в остальных хроматинах, помимо аквамарина, показана на уровне оси абсцисс (рис. 1a).

Аналогичная работа была проведена для сравнений генов в максимально контрастном состоянии рубин (данные не приводятся). Выявлен один наиболее устойчивый мотив — ТАТА-бокс.

Оказалось, что среди промоторов генов домашнего хозяйства у дрозофилы возможны разнообразные комбинации мотивов. Кроме того, 33% генов этого типа вообще не имеют ни одного из этих мотивов (рис. 1б). Аналогичная работа была проведена для анализа промоторов генов развития, имеющих только очень специфический

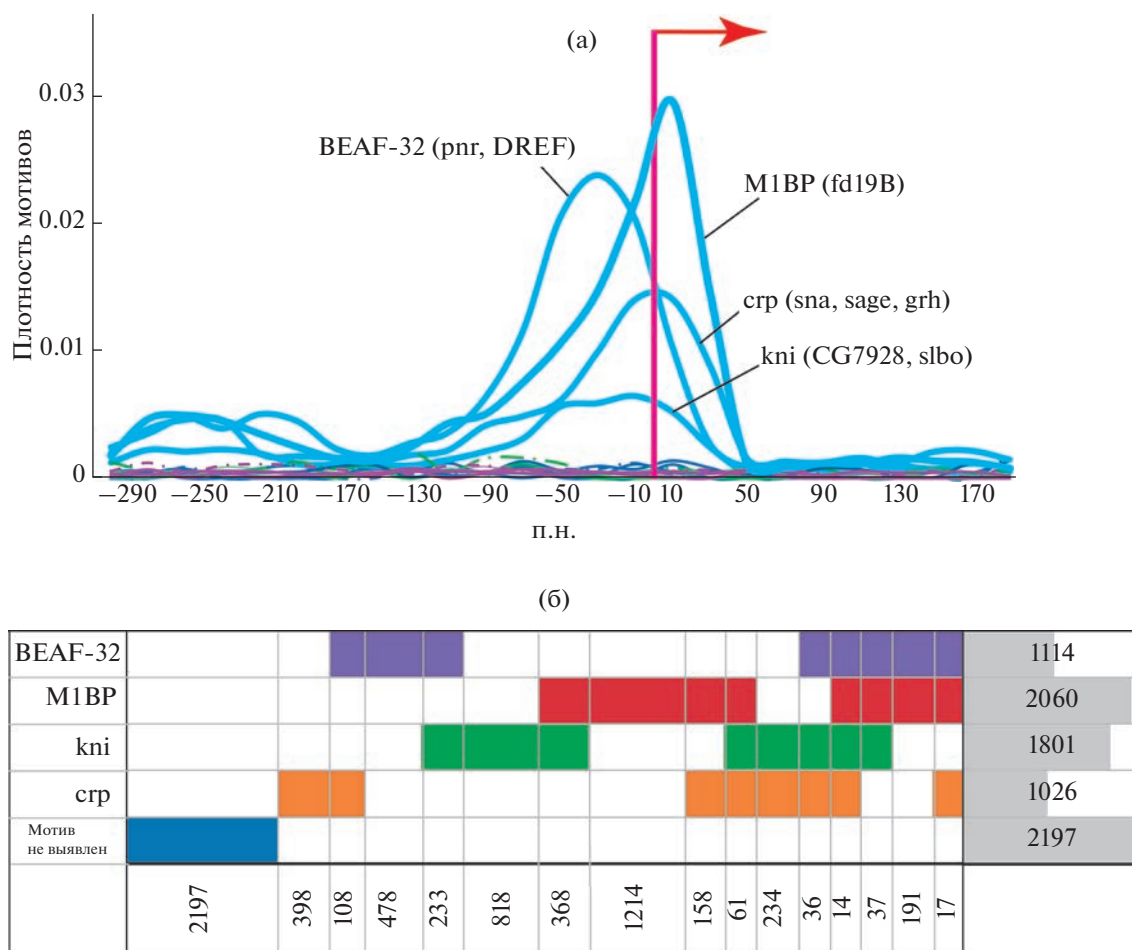


Рис. 1. Характеристики промоторных участков 6562 генов домашнего хозяйства дрозофилы. (а) Суммарная локализация мотивов относительно точки инициации транскрипции (красная стрелка). По оси абсцисс показано расстояние до точки инициации транскрипции, по оси ординат частота локализации мотива в данной координате околопромоторного участка. Приведены наиболее употребительные названия мотива, а в скобках указаны названия, которые встречаются в других статьях/базах данных. Изученные мотивы присутствуют только в промоторных участках хроматина аквамарины (голубая линия), все остальные промоторные области не содержат этих мотивов (зеленая, синяя, фиолетовая линии на уровне нулевого значения по оси абсцисс). (б) Числа генов домашнего хозяйства, имеющих только один мотив (вертикальная колонка цифр справа) или комбинации мотивов (горизонтальная последовательность цифр внизу). Прямоугольниками разных цветов обозначено наличие изучаемого мотива (отсутствие цвета в клетке обозначает отсутствие мотива в данном наборе генов).

мотив – ТАТА-бокс (данные не приводятся). Мотивы Mes2, GAGA, ttk обладают меньшей специфичностью. Также обнаружено, что около 50% генов развития не имеют ни ТАТА, ни каких-либо других из найденных мотивов. В аналогичном сравнении локализации мотивов у генов развития максимальная специфика обнаружена только для ТАТА-бокса в положении от –75 до –25 п.н. от точки инициации транскрипции. Анализ, проведенный с помощью Gene Ontology, показал, что для отдельных групп междискковых генов, содержащих в промоторах один мотив или их комбинации, характерно выполнение определенных функций.

Таким образом впервые описаны специфические мотивы в промоторах генов двух разных типов на уровне всего генома и показано, что наборы мо-

тивов у генов “домашнего хозяйства” и генов развития совершенно разные, а наиболее известные мотивы не являются универсальными.

БЛАГОДАРНОСТИ

Авторы выражают благодарность Центру коллективного пользования “Молекулярная и клеточная биология” при ИМКБ СО РАН.

ФИНАНСИРОВАНИЕ

Данное исследование выполнено при финансовой поддержке раздела Государственного задания ИМКБ РАН № FWGZ-2021-0014 (полногеномный анализ, Gene Ontology анализ) и грантов РФФИ № 19-14-00051-П

(обработка данных по локализации мотивов) и РФФ № 21-14-00240 (поиск мотивов).

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

При выполнении данного исследования манипуляции с экспериментальными животными не проводились.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют, что какой-либо конфликт интересов отсутствует.

ИНФОРМАЦИЯ О ВКЛАДЕ АВТОРОВ

И.Ф. Жимулев выполнял основную работу, инициировал проведение исследований, участвовал в написании статьи, В.Г. Левицкий, А.В. Цуканов планировал и выполняли компьютерные операции, анализировали результаты и участвовали в написании текста статьи, Т.Ю. Ватолина обрабатывала результаты компьютерных операций, участвовала в планировании исследования, обсуждении результатов и написании статьи, Т.Д. Колесникова участвовала в написании текста статьи и обсуждении результатов.

СПИСОК ЛИТЕРАТУРЫ

Andersson R., Sandelin A. Determinants of enhancer and promoter activities of regulatory elements // *Nat. Rev. Genet.* 2020. V. 21. P. 71–87.
Crick F. General model for the chromosomes of higher organisms // *Nature.* 1971. V. 234. P. 25–27.

Danino Y.M., Even D., Ideses D. et al. The core promoter: At the heart of gene expression // *Biochim. Biophys. Acta.* 2015. V. 1849(8). P. 1116–1131.

Koltzoff N.K. The structure of the chromosomes in the salivary glands of *Drosophila* // *Science.* 1934. V. 60. P. 312–313.

Lenhard B., Sandelin A., Carninci P. Metazoan promoters: emerging characteristics and insights into transcriptional regulation // *Nat. Rev. Genet.* 2012. V. 13(4). P. 233–245.

Levitsky V.G., Zykova T.Yu., Moshkin Y.M., Zhimulev I.F. Nucleosome positioning around transcription start site correlates with gene expression only for active chromatin state in *Drosophila* interphase chromosomes // *International J. Molecular Sciences.* 2020. V. 21(23). P. 9282.

Paul J. General theory of chromosomes structures and gene activation in eukaryotes // *Nature.* 1972. V. 238. P. 444–446.

Ramalingam V., Natarajan M., Johnston J. et al. TATA and paused promoters active in differentiated tissues have distinct expression characteristics // *Mol. Syst. Biol.* 2021. V. 17. P. e9866.

Zhimulev I.F., Belyaeva E.S. Proposals to the problem of structural and functional organization of polytene chromosomes // *Theor. Appl. Genet.* 1975. V. 45(8). P. 335–340.

Zhimulev I.F., Zykova T.Yu., Goncharov F.P. et al. Genetic organization of polytene chromosome bands and interbands in *Drosophila melanogaster* // *PLoS One.* 2014. V. 9(7). P. e101631.

Zykova T.Yu., Levitsky V.G., Belyaeva E.S. et al. Polytene chromosomes – a portrait of functional organization of the *Drosophila* genome // *Curr. Genomics.* 2018. V. 19(3). P. 179–191.

Development of N.K. Koltsov Idea about Genetic Organization of Interbands in *Drosophila melanogaster* Polytene Chromosomes

I. F. Zhimulev^{1, *}, T. Yu. Vatolina¹, V. G. Levitsky², T. D. Kolesnikova¹, and A. V. Tsukanov²

¹*Institute of Molecular and Cellular Biology SB RAS, prosp. ak. Lavrentieva 8/2, Novosibirsk, 630090 Russia*

²*Institute of Cytology and Genetics SB RAS, prosp. ak. Lavrentieva 10, Novosibirsk, 630090 Russia*

*e-mail: zhimulev@mcb.nsc.ru

Here for the first time, the organization of promoters of developmental gene promoters and promoter of genes necessary for general cellular functions—the “housekeeping” of the cell in the complete genome of *Drosophila melanogaster* were studied. Using bioinformatic methods, it has been shown that the genes whose promoters are located in the interbands of polytene chromosomes are enriched in functions associated with general cellular processes, while the rest of the genes (about half of the *Drosophila* genome) are associated with highly specialized processes occurring during development. In the promoter zone of the housekeeping genes, four specific motifs were found that can be present in different genes individually or in various combinations. A significant part of interband promoters do not contain identified motifs. The analysis carried out using Gene Ontology showed that certain groups of interband genes containing one motif in promoters or their combinations are characterized by the performance of certain functions.

Keywords: developmental genes, housekeeping genes, gene promoters, nucleotide motifs in promoters, interbands, bands, polytene chromosomes, *Drosophila melanogaster*