

ФИЗИКО-ХИМИЧЕСКИЕ СВОЙСТВА ДНК В РЕГУЛЯТОРНЫХ УЧАСТКАХ ГЕНОМАНечипуренко Ю.Д.¹, Ильичева И.А.¹, Попцова М.С.², и Гроховский С.Л.¹¹Институт молекулярной биологии им. В.А. Энгельгардта РАН,
ул. Вавилова, 32, г. Москва, 119991, РФ²МГУул. Ленинские горы, 1, стр. 2, г. Москва, 119991, РФ
e-mail: nech99@mail.ru

Аннотация. Общая структура и способ действия РНК полимераз эукариот и архей сходны, несмотря на различия в последовательности промоторов из разных видов. Целью нашей работы является поиск общих физико-химических характеристик участков ДНК, которые определяют их свойство как промоторов для РНК полимеразы II (Pol II). В работе проведен анализ профилей большого числа физических и структурных характеристик, усредненных по репрезентативным выборкам из разных видов для животных, растений и одноклеточных грибов. В дополнении к известным динуклеотидным характеристикам ДНК мы впервые использовали индексы, характеризующие расщепление ДНК ультразвуком. Эти индексы несут информацию о свойствах отдельных нитей молекулы ДНК. Особые свойства обнаруживаются для ДНК из многоклеточных организмов в области, которая соединяет ТАТА бокс и начало сайта инициации транскрипции (TSS). Мы связываем их с интенсивностью конформационных движений, которая меняется в противофазе вдоль ДНК. Это видно лучше всего для ДНК из клеток млекопитающих. Обнаруженные нами физико-химические свойства молекулы ДНК могут быть использованы в генетической инженерии для искусственной модуляции силы промоторов.

Ключевые слова: Локальная структура и конформация ДНК, Промоторы РНК полимеразы II, Специфичное расщепление ДНК ультразвуком.

PHYSICO-CHEMICAL PROPERTIES OF DNA IN REGULATORY SITES OF THE GENOMENechipurenko Y. D.¹, Il'icheva I. A.¹, Poptsova M. S.², Grokhovsky, S. L.¹¹Engelhardt Institute of Molecular Biology, Russian Academy of Sciences,
Vavilov str., 32, Moscow, 119991, Russia²Moscow State UniversityLeninskie gory str., 1, bld. 2, Moscow, 119991, Russia
e-mail: nech99@mail.ru

Abstract. The general structure and action of all eukaryotic and archaeal RNA polymerases machinery have an astonishing similarity despite the diversity of core promoter sequences in different species. The goal of our work is to find common characteristics of DNA region that define it as a promoter for the RNA polymerase II (Pol II). The profiles of a large number of physical and structural characteristics, averaged over representative sets of the Pol II minimal core promoters of the evolutionary divergent species from animals, plants and unicellular fungi were analysed. In addition to the characteristics defined at the base-pair steps, we, for the first time, use profiles of the ultrasonic cleavage and DNase I cleavage indexes, informative for internal properties of each complementary strand. Special characteristics of conformational behavior are revealed in metazoans at the region, which connects the end of TATA-box and the transcription start site (TSS). The intensities of conformational motions in the complementary strands are periodically changed in opposite phases. They are noticeable, best of all, in mammals. Obtained results may be useful in genetic engineering for artificial modulation of the promoter strength.

Keywords: Local DNA structure, RNA polymerase II promoter sequences, Sequence-specific ultrasonic cleavage.

Геном живого организма можно сравнить с книжной полкой, на которой хранятся книги-хромосомы с последовательностями букв-нуклеотидов. Первоначальные методы расшифровки последовательностей были основаны на точном вырезании определённого куска текста из такой книги и его аккуратном прочтении. Более сорока лет назад был предложен метод секвенирования, когда вся геномная ДНК разрывается на короткие куски (так называемые "риды"), которые все прочитываются, и потом эти риды, как пазлы, собираются в единый текст, используя перекрывающиеся концевые участки. Совершенствование автоматических секвенаторов и увеличение мощности вычислительных машин сделало именно этот метод доминирующим в настоящее время. Современные секвенаторы могут за сутки прочитать сотни миллионов ридов, каждый длиной несколько десятков или сотен букв. Поскольку в автоматических секвенаторах довольно большой процент ошибок считывания, постольку для надежного перекрывания концевых последовательностей при сборке этих "пазлов", требуется, чтобы суммарная длина прочитанных ридов в несколько десятков раз превосходила длину анализируемого генома.

Затруднением для применения этого метода представляют протяженные повторяющиеся последовательности, которые часто встречаются в геномах эукариот и длина которых может превышать длину отдельных ридов. Например, геном человека содержит сотни тысяч повторяющихся элементов. Это, в свою очередь, может приводить к систематическим ошибкам при процедуре компьютерной сборки (особенно в областях генома, содержащих большое количество повторов).

После секвенирования фрагментов ДНК производится компьютерная сборка всего генома с многократным наложением перекрывающихся участков. В основе такого подхода к секвенированию лежит предположение, что разрывы ДНК на фрагменты происходят случайно и не зависят от последовательности нуклеотидов. Однако при

фрагментации ДНК геномов с помощью методов, основанных на разрыве ДНК под действием гидродинамических сил (ультразвук, небулизация и «Каварис»), обнаружено, что фрагменты из одних областей присутствуют в избытке, а из других - в недостатке. Это свидетельствует о неслучайном характере разрывов.

При анализе продуктов ультразвукового расщепления рестриктных фрагментов ДНК с помощью электрофореза в полиакриламидных гелях нами было обнаружено явление специфичности такого расщепления: частота (вероятность) двунитевых разрывов зависит от последовательности нуклеотидов [1]. Было показано, что частота разрывов зависит также от температуры, pH и ионной силы раствора, но не зависит от используемой частоты ультразвука. Наиболее часто двойная спираль разрывается по динуклеотиду d(CpG). Было обнаружено, что фосфатная группа остается на 5'-конце образующихся фрагментов, а характер процесса имеет черты механохимической реакции. Позже мы показали, что при разрыве двухспиральной ДНК при помощи ультразвука частоты разрыва цепи зависят от последовательности нуклеотидов [2-7]. В результате проведенного анализа 5'-концевых участков фрагментов, выровненных на соответствующие референтные геномы, мы обнаружили, что частоты разрывов разных ди- и тетрануклеотидов на концах фрагментов находятся в хорошем соответствии с данными ультразвукового расщепления ДНК. Частоты ультразвукового расщепления можно использовать для физического картирования генома, то есть определения участков, имеющих аномалии по ряду физических характеристик [8]. Эти аномалии связаны, как правило, с регуляторными участками генов.

В работе [9] мы показали, что промоторные участки ряда генов из разных организмов имеют такого рода аномалии, и таким образом, использование наших данных о частотах разрывов ди- и тетрануклеотидов позволяет обнаруживать промоторные участки. Мы рассмотрели разные физико-химические характеристики участков ДНК, которые могут определять их свойства как промоторов для РНК полимеразы II (Pol II). В работе проведен анализ профилей большого числа физических и структурных характеристик, усредненных по репрезентативным выборкам из разных видов для животных, растений и одноклеточных грибов. Индексы, характеризующие расщепление ДНК ультразвуком, несут информацию о свойствах отдельных нитей молекулы ДНК (в отличие от других индексов). Особые свойства обнаруживаются для ДНК из многоклеточных организмов в области, которая соединяет ТАТА бокс и начало сайта инициации транскрипции (TSS). Таким образом, можно утверждать, что наряду с такими известными локальными характеристиками молекулы ДНК, как твист, ширина бороздок и т.п. наша характеристика – индекс расщепления ДНК ультразвуком, которая, как мы показали ранее [9] связана с конформацией и подвижностью сахара в молекуле ДНК, помогает выявить регуляторные участки ДНК и обнаружить сходные свойства промоторов из разных видов.

Список литературы / References:

1. Гроховский С.Л., Ильичева И.А., Нечипуренко Д.Ю., Панченко Л.А., Полозов Р.В., Нечипуренко Ю.Д. Биофизика, 2008, т.53, с. 417-425. [Grokhovsky S.L. et al Biophysics, 2008, vol. 53, pp. 417-425 (In Russ.)]
2. Нечипуренко Ю.Д., Головкин М.В., Нечипуренко Д.Ю., Ильичева И.А., Панченко Л.А., Полозов Р.В., Гроховский С.Л. Журнал структурной химии, 2009, т. 50, с. 1040-1047. [Nechipurenko Yu.D. et al J. Strukt. Khimii 2009, vol. 50, pp. 1040-1047 (In Russ.)]
3. Grokhovsky S.L., Il'icheva I.A., Nechipurenko D.Yu., Golovkin M.V., Panchenko L.A., Polozov R.V., Nechipurenko Yu.D. Biophysical Journal, 2011, vol.100, no.1, pp. 117-125.
4. Grokhovsky S., Il'icheva I., Nechipurenko D., Golovkin M., Taranov G., Panchenko L., Polozov R. and Nechipurenko Yu. Gel Electrophoresis - Principles and Basics, Dr. Sameh Magdeldin Ed., 2012, ISBN: 978-953-51-0458-2, InTech.
5. Poptsova M.S., Il'icheva I.A., Nechipurenko D.Y., Panchenko L.A., Khodikov M.V., Oparina N.Y., Polozov R.V., Nechipurenko Y.D., Grokhovsky S.L. Sci Rep., 2014, vol. 4, p. 4532.
6. Grokhovsky S.L., Il'icheva I.A., Nechipurenko D.Yu., Golovkin M.V., Panchenko L.A., Polozov R.V. and Nechipurenko Yu.D. Advances in Engineering Research, 2014, vol. 8, pp. 213-236, ISSN: 2163-3932, ISBN: 978-1-63321-282-4.
7. Нечипуренко Д.Ю., Ильичева И.А., Ходыков М.В., Попцова М.С., Нечипуренко Ю.Д., Гроховский С.Л. Биофизика, 2014, т. 59, с. 1061-1070. [Nechipurenko D.Yu. et al Biophysics, 2014, vol. 59, pp. 1061-1070 (In Russ.)]
8. Нечипуренко Ю.Д., Нечипуренко Д.Ю., Ильичева И.А., Головкин М.В., Панченко Л.А., Полозов Р.В., Гроховский С.Л. Журнал «Компьютерные исследования и моделирование», 2010, т. 2, с. 419-428. [Nechipurenko Yu.D. et al J. Computer Res. and Modelling, 2010, vol. 2, pp. 419-428 (In Russ.)]
9. Il'icheva I.A., Khodikov M.V., Poptsova M.S., Nechipurenko D.Yu., Nechipurenko Yu.D., Grokhovsky S.L. Structural features of DNA that determine RNA polymerase II core promoter. *BMC genomics*, 2016, vol. 17, p. 973.