

Локальная структура ДНК базальных промоторов генов тРНК эукариот

Савина Е.А.^{1, 2*}, Орлов Ю.Л.^{1, 3}, Лебедев Г.С.¹, Анашкина А.А.^{1, 2}, Ильичёва И.А.²

¹ ФГАОУ ВО Первый Московский государственный медицинский университет им. И.М. Сеченова
Министерства здравоохранения Российской Федерации (Сеченовский Университет),
Москва, Россия

² Институт молекулярной биологии им. В.А. Энгельгардта РАН, Москва, Россия

³ Институт цитологии и генетики СО РАН, Новосибирск, Россия

* savina_e_a_1@staff.sechenov.ru

Ключевые слова: транскрипция тРНК; ТАТА-связывающий белок; ультразвуковое расщепление ДНК; расщепление ДНКазой I; локальная структура ДНК

Мотивация и цель: Три ДНК-зависимые РНК-полимеразы осуществляют транскрипцию всех типов молекул РНК в клетках эукариот. Это РНК-полимераза I (Pol I), которая транскрибирует предшественник рРНК, РНК-полимераза II (Pol II), которая транскрибирует предшественники информационных РНК и различных регуляторных некодирующих РНК, и РНК-полимераза III (Pol III), которая транскрибирует предшественники малых РНК, включая все тРНК, 5SрРНК, 7SLРНК, малых ядерных РНК U6 [1], а также мобильные генетические элементы, называемые SINE [2]. В растениях дополнительно присутствуют еще две РНК-полимеразы [3]. Ключевым фактором транскрипции Pol II является белок (ТВР), распознающий ТАТА-бокс в базальных промоторах генов матричной РНК. Ранее мы показали, что позиция связывания его в окрестностях –28 п. н. от старта транскрипции определяется локальными структурными свойствами ДНК на этом участке [4–6]. Этот белок присутствует и в транскрипционных комплексах Pol I и Pol III. Целью данной работы было выявление аналогичного сайта связывания ТВР в области, предшествующей генам тРНК в геномах эукариот.

Методы и алгоритмы: Мы использовали коллекцию тРНК 7 эукариот: *S. pombe* 972h- (171 ген), *S. cerevisiae* S288c (273 гена), *A. thaliana* TAIR 10 (642 гена), *D. melanogaster* (BDGP Rel. 6/dm6) (290 генов), *C. elegans* W BceI235/ce 11 (580 генов), *Mus musculus* GRCm39/mm39 (400 генов) и *H. sapiens* GRCh38/hg38 (429 генов), находящуюся в базе данных геномных тРНК (GtRNAdb). Эта база содержит гены тРНК, выявленные на полных геномах с помощью компьютерного поиска программой tRNAscan-SE. Мы проанализировали нуклеотидный текст и изменения локальной пространственной структуры фрагментов длиной 80 н. п. в окрестностях начала генов тРНК (–50/+30). Были определены профили частот встречаемости мононуклеотидов, их лого-представления, а также профили изменений физических и структурных параметров интенсивностей ультразвукового расщепления, и расщепления ДНКазой I. Для этой цели нами была написана программа на Python 3.10.

Результаты: Оказалось, что профили, характеризующие тексты участков, предшествующих генам (от –50 до –1), существенно различаются у разных представителей эукариот, а начальные участки самих генов тРНК (от +1 до +30) у всех эукариот имеют сходные последовательности и одинаковый консенсус бокса A.

В отличие от этого профили структурных и физических параметров, включая характеристики интенсивностей расщеплений ДНК ультразвуком и с помощью фермента ДНКазы I, выявляют сходство локальной структуры участков, предшествующих генам (-50/-1). Эти структурные особенности, которые наиболее выражены у *S. pombe* 972h-, показаны на рис. 1. Сравнение структурных профилей всех семи изученных организмов указывают на расширение малой бороздки ДНК и низкую энергию изгиба в сторону несакхарного желоба, немного варьирующую на всем протяжении, вплоть до начала гена. Такие особенности структуры и динамики ДНК вблизи генов тРНК существенно отличают их от базальных промоторов матричной РНК.

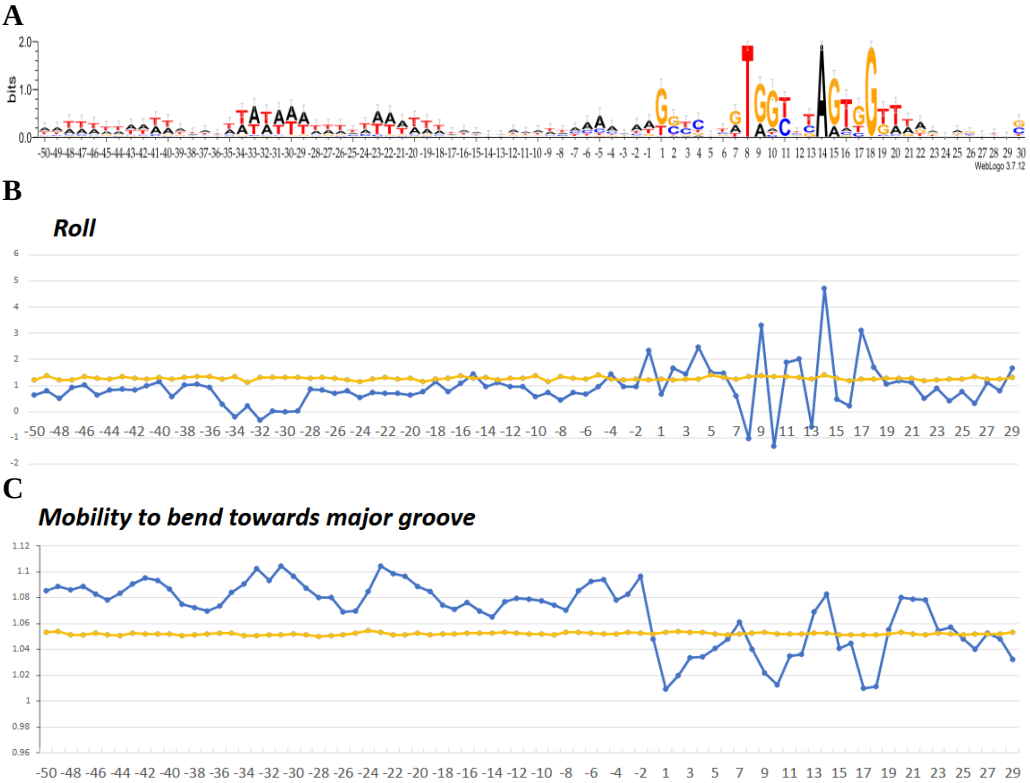


Рис. 1. А. Лого-представление областей от -50 до +30 для 171 гена тРНК *S. pombe* 972h-. В. Структурные параметры Roll. С. Mobility to bend towards major groove для областей от -50 до +30 генов тРНК *S. pombe* 972h- (синий цвет) и случайных последовательностей (желтый)

Выводы: Особенности структуры ДНК на участках базальных промоторов генов тРНК, которые заметно отличают их от промоторных участков генов матричной РНК, транскрибируемых РНК-полимеразой II, позволяют предположить, что при транскрипции генов тРНК ТАТА-связывающий белок не играет ключевой роли в формировании транскрипционного комплекса РНК-полимеразы III.

Финансирование: Исследование поддержано грантом Российского научного фонда № 24-24-00563 на тему «Разработка цифровых образовательных программ в биомедицине».

Local structure of basal DNA promoters in eukaryotic tRNA genes

Savina E.^{1, 2*}, Orlov U.^{1, 3}, Lebedev G.¹, Anashkina A.^{1, 2}, Il'icheva I.²

¹ Sechenov First Moscow State Medical University of the Russian Ministry of Health (Sechenov University), Moscow, Russia

² V.A. Engelhardt Institute of Molecular Biology, RAS, Moscow, Russia

³ Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

* savina_e_a_1@staff.sechenov.ru

Key words: tRNA transcription; TBP; ultrasonic cleavage; DNase I cleavage; local structure of DNA

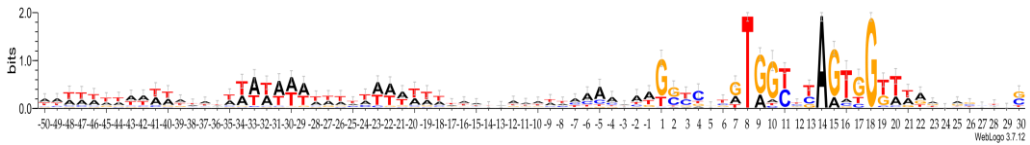
Motivation and Aim: Transcription of all types RNA molecules is carried out by three DNA-dependent RNA polymerases in eukaryotic cells. RNA polymerase I (Pol I) synthesizes the rRNA precursor, RNA polymerase II (Pol II) synthesizes messenger RNAs and various regulatory non-coding RNAs, and RNA polymerase III (Pol III) synthesizes small RNAs, including all tRNAs, 5SrRNAs, 7SLRNAs, U6 small nuclear RNAs [1], as well as mobile genetic elements referred to as SINEs [2]. In plants also function two additional RNA polymerases [3]. A key general transcription factor in Pol II machinery is TATA-binding protein (TBP) that recognizes the TATA box in the basal promoters of RNA genes. We have previously shown that its binding position in the vicinity of –28 bp from the transcription start (TSS) is determined by local structural properties of DNA at this site [4–6]. TBP is the component of Pol I and Pol II transcription complexes also. The aim of this work was to find analogical binding site of TBP in the region upstream tRNA genes in eukaryotic genomes.

Methods and Algorithms: We utilized the tRNA collection of 7 eukaryotic organisms: *S. pombe* 972h- (171 genes), *S. cerevisiae* S288c (273 genes), *A. thaliana* TAIR 10 (642 genes), *D. melanogaster* (BDGP Rel. 6/dm6) (290 genes). *C. elegans* WbceI235/ce 11 (580 genes), *Mus musculus* GRCm39/mm39 (400 genes) and *H. sapiens* GRCh38/hg38 (429 genes), from genomic tRNA database (GtRNAdb). It contains tRNA genes identified on full genomes by tRNAscan-SE – computer search program. We analyzed the nucleotide text and changes in the local spatial structure of fragments 80 np in length in the vicinity of tRNA gene starts (–50/+30). Profiles of mononucleotide occurrence frequencies, their logo-presentation, and profiles of changes in physical and structural parameters, ultrasonic cleavage and DNAase I cleavage intensities were determined. For this purpose, we wrote a program in Python 3.10.

Results: It turned out that the profiles characterizing the texts of the sites preceding genes (from –50 to –1) differ significantly among the representatives of eukaryotes, while the sequences of tRNA genes themselves (from +1 to +30) in all eukaryotes are similar, as well as their consensus of box A. In contrast, all profiles of structural and physical parameters, including characterization of DNA cleavage intensities by ultrasound and by the DNAase I enzyme, reveal similarities in the local structure of the regions preceding the genes (–50/–1). These structural features, which are most pronounced in *S. pombe* 972h-, are shown in Figure 1. Comparison of the structural profiles of all seven organisms studied indicate a widening of the minor groove of DNA and a low bending energy towards the major groove, slightly varying up to the gene start. These features of

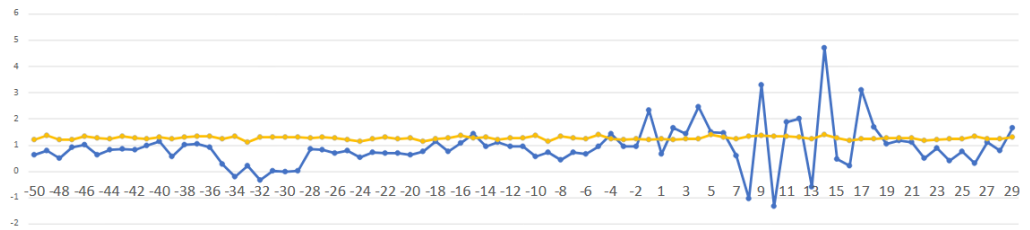
DNA structure and dynamics near tRNA genes essentially distinguish them from the basal promoters of matrix RNA.

A



B

Roll



C

Mobility to bend towards major groove

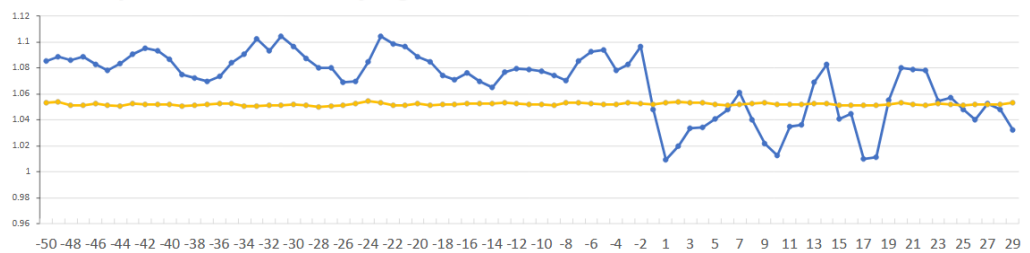


Fig. 1. A. Logo representation of the –50 to +30 regions for 171 *S. pombe* 972h- tRNA genes. B. Structural parameters Roll. C. Mobility to bend toward major groove for the –50 to +30 regions of *S. pombe* 972h- tRNA genes (blue) and random sequences (yellow)

Conclusion: These features of local structure of DNA in the vicinity of tRNA genes, which markedly distinguish them from the promoter sites of mRNA genes transcribed by RNA polymerase II, suggest that TATA-binding protein doesn't play a key role in the transcriptional complex formation of RNA polymerase III.

Funding: The study is supported by Russian Science Foundation (No. 24-24-00563) on the topic “Development of digital educational programs in biomedicine”.

Список литературы/References

1. Girbig M., Misiaszek A.D., Müller C.W. Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases. *Nat Rev Mol Cell Biol.* 2022;23(9):603-622
2. Vassetzky N.S., Kramerov D.A. SINEBase: a database and tool for SINE analysis. *Nucleic Acids Res.* 2013;41(D1):D83-D89
3. Zhou M., Law J.A. RNA Pol IV and V in gene silencing: rebel polymerases evolving away from Pol II's rules. *Curr Opin Plant Biol.* 2015;27:154-164
4. Il'icheva I.A., Khodikov M.V., Poptsova M.S., Nechipurenko D.Y., Nechipurenko Y.D., Grokhovsky S.L. Structural features of DNA that determine RNA polymerase II core promoter. *BMC Genomics.* 2016;17:1-21
5. Melikhova A.V., Anashkina A.A., Il'icheva I.A. Evolutionary invariant of the structure of DNA double helix in RNAP II Core Promoters. *Int J Mol Sci.* 2022;23:10873. doi 10.3390/ijms231810873
6. Savina E.A., Shumilina T.G., Tumanyan V.G., Anashkina A.A., Il'icheva I.A. Core promoter regions of antisense and long intergenic non-coding RNAs. *Int J Mol Sci.* 2023;24(9):8199