

Сравнительные характеристики первичной и пространственной структуры генов тРНК архей, прокариот и эукариот

Шумилина Т.Г.¹, 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 имени И.М. Сеченова Министерства здравоохранения Российской Федерации (Сеченовский Университет), Москва, Россия

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

³ Институт цитологии и генетики СО РАН, Новосибирск, Россия

* tshumilina@mail.ru

Ключевые слова: промоторы генов тРНК, ТАТА-бокс, боксы А и В

Мотивация и цель: Структурное сходство, присущее РНК-полимеразам (Pols) бактерий, архей и эукариот, отражает общее происхождение систем транскрипции у всех организмов на Земле [1]. ТАТА-связывающий белок (ТВР) является компонентом общих факторов транскрипции (GTFs) у всех трех ДНК-зависимых РНК-полимераз эукариот, археи используют его ортолог. Ранее, при анализе базальных промоторов генов, транскрибируемых Pol II, нами была обнаружена характерная сингулярная область, отвечающая за непосредственное взаимодействие с белком ТВР [2–4]. Целью данной работы было обнаружение аналогичных сингулярных областей, отвечающих за регуляцию транскрипции генов тРНК у архей, прокариот и эукариот.

Методы и алгоритмы: Были проанализированы текстовые и структурные характеристики фрагментов геномов, содержащих гены тРНК и их upstream-областей для наборов генов тРНК представителей архей: *M. barkeri str. Fusaro* (63 гена тРНК), *H. volcanii DS2* (53 гена тРНК), прокариот: *E. coli str. K-12 substr. MG1655* (89 генов тРНК), *B. subtilis subsp. subtilis str. 168* (86 генов тРНК) и низших одноклеточных эукариот *S. pombe 972h-* (171 ген тРНК). Анализировали последовательности участков, расположенных в позициях от –50 до +30 п.н. относительно старта транскрипции, и в позициях от N-60 до N, где N – координата последнего нуклеотида гена тРНК. Получены профили частот встречаемости нуклеотидов, их лого-представления, а также профили структурных и физико-химических свойств.

Результаты: У всех исследованных организмов были обнаружены две сингулярные области, расположенные внутри генов. Их координаты отвечают расположению боксов А и В. Анализ первичной структуры выявил определенные отличия частот встречаемости G:C и A:T пар, характер которых позволяет определить домен, к которому относится конкретный организм. На профилях структурных параметров одноклеточных эукариот *S. pombe 972h-* заметно некоторое снижение жесткости двойной спирали к изменениям угла Roll на участке от –28 до –36 н.п., что может свидетельствовать о более низких энергетических затратах на структурные перестройки ДНК при взаимодействии с

белком ТВР именно в этой позиции. Такого эффекта на профилях представителей архей и прокариот мы не обнаружили.

Выводы: Текстовые и структурные параметры двойной спирали ДНК фрагментов нуклеотидных последовательностей генов тРНК и примыкающих к ним областей выше старта транскрипции (до –50 н.п.) из геномов архей (*M. barkeri* str. *Fusaro* и *H. volcanii* DS2), прокариот (*E. coli* str. *K-12 substr. MG1655* и *B. subtilis* subsp. *subtilis* str. 168) и низших эукариот (*S. pombe* 972h-) свидетельствует о присутствии двух сингулярных областей внутри генов тРНК. Их нуклеотидные последовательности у представителей трех доменов отличаются незначительно. Особенности пространственной структуры двойной спирали ДНК выше старта транскрипции в геномах представителя эукариот *S. pombe* 972h-, свидетельствует о возможности избирательного связывания белка ТВР на участке от –28 до –36 н.п., который можно назвать ТАТА-боксом.

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

Comparative characteristics of the primary and spatial structure of tRNA genes in archaea, prokaryotes and eukaryotes

Shumilina T.G.^{1*}, Orlov Yu.L.^{1, 3}, Lebedev G.S.¹, Anashkina A.A.^{1, 2}, Il'icheva I.A.²

¹ Sechenov Moscow State Medical University (Sechenov University), Moscow, Russia

² Engelhardt Institute of Molecular Biology, Moscow, Russia

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

* tshumilina@mail.ru

Key words: tRNA gene promoters; TATA box; boxes A and B

Motivation and Aim: The structural similarity shared by RNA polymerases (Pols) of bacteria, archaea, and eukaryotes reflects the common origin of transcription systems in all organisms on Earth [1]. TATA binding protein (TBP) is a component of common transcription factors (GTFs) in all three DNA-dependent RNA polymerases of eukaryotes, and archaea use its ortholog. Previously, when analysing the basal promoters of genes transcribed by Pol II, we discovered a characteristic singular region responsible for direct interaction with the TBP protein [2–4]. The goal of this work was to discover similar singular regions responsible for the regulation of tRNA gene transcription in archaea, prokaryotes, and eukaryotes.

Methods and Algorithms: The textual and structural characteristics of genome fragments containing tRNA genes and their upstream regions were analyzed for tRNA gene sets of archaeal representatives: *M. barkeri* str. *Fusaro* (63 tRNA genes), *H. volcanii* DS2 (53 tRNA genes), prokaryotes: *E. coli* str. *K-12 substr. MG1655* (89 tRNA genes), *B. subtilis* subsp. *subtilis* str. 168 (86 tRNA genes) and lower unicellular eukaryotes *S. pombe* 972h- (171 tRNA genes). The sequences of regions located at positions from –50 bp were analyzed up to +30 bp relative to the start of transcription, and in positions from N-60 to N, where N is the coordinate of the last nucleotide of the tRNA gene.

Frequency profiles of mononucleotides, their logo representations, as well as profiles of structural and physicochemical properties were obtained.

Results: In all organisms studied, two singular regions located inside genes were discovered. Their coordinates correspond to the location of boxes A and B. Analysis of the primary structure revealed certain differences in the frequencies of occurrence of G:C and A:T pairs, the nature of which makes it possible to determine the domain to which a particular organism belongs. Profiles of structural parameters of unicellular eukaryotes *S. pombe* 972h-, show a slight decrease in the rigidity of the double helix in response to changes of the Roll angle in the region from –28 to –36 bp, which may indicate lower energy costs for structural rearrangements of DNA when interacting with TBP protein in this particular position. We did not find such an effect on the profiles of representatives of archaea and prokaryotes.

Conclusion: Textual and structural parameters of the DNA double helix of fragments of nucleotide sequences of tRNA genes and their upstream regions from the genomes of archaea (*M. barkeri* str. Fusaro and *H. volcanii* DS2), prokaryotes (*E. coli* str. K-12 substr. MG1655 and *B. subtilis* subsp. *subtilis* str. 168) and lower eukaryotes (*S. pombe* 972h-) indicates the presence of two singular regions within tRNA genes. Their nucleotide sequences differ slightly among representatives of the three domains. Features of the spatial structure of the double helix above the start of transcription in eukaryotes (*S. pombe* 972h-) indicate the possibility of selective binding of the TBP protein in the region from –28 to –36 bp, which can be called the TATA box.

Funding: The research is supported by a grant from the Russian Science Foundation No. 24-24-00563 on the topic “Development of digital educational programs in biomedicine”.

Список литературы/References:

1. Hamada M., Huang Y., Lowe T.M., Maraia R.J. Widespread use of TATA elements in the core promoters for RNA polymerases III, II, and I in fission yeast. *Mol Cell Biol.* 2001;21(20):6870-6881
2. Il'icheva I.A., Khodikov M.V., Poptsova M.S. et al. Structural features of DNA that determine RNA polymerase II core promoter. *BMC Genomics.* 2016;17(1):973
3. 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(18):10873
4. 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