



УДК 577.112.5

## КЛОНИРОВАНИЕ И ОПРЕДЕЛЕНИЕ ПЕРВИЧНОЙ СТРУКТУРЫ ФРАГМЕНТОВ ГЕНОВ $\beta$ -СУБЪЕДИНИЦЫ РНК-ПОЛИМЕРАЗ ПРЕДСТАВИТЕЛЕЙ БАЙКАЛЬСКОГО ПИКОПЛАНКТОНА

© 1998 г. О. А. Сутурина, Ю. А. Сутурина, Е. А. Семенова<sup>#</sup>

Лимнологический институт СО РАН, 664033, Иркутск, ул. Улан-Баторская, 3

Поступила в редакцию 06.11.97 г. Принята к печати 25.05.98 г.

С использованием полимеразной цепной реакции (ПЦР) клонированы фрагменты генов  $\beta$ -субъединицы РНК-полимераз представителей пикопланктона озера Байкал. Проведено секвенирование полученных рекомбинантных клонов. Анализ 93 клонов выявил 40 типов последовательностей. Сравнение на нуклеотидном и белковом уровне полученных последовательностей с гомологичными из банка данных подтверждает консерватизм района  $\beta$ -субъединицы РНК-полимеразы, участвующего в формировании активного центра фермента.

*Ключевые слова:*  $\beta$ -субъединица РНК-полимеразы; активный центр; секвенирование; пикопланктон.

РНК-полимераза (КФ 2.7.7.6) – один из ключевых ферментов клетки. Было показано, что  $\beta$ -субъединица бактериальной РНК-полимеразы вовлечена в каталитическую активность фермента [1, 2]. В настоящее время известны полные или частичные аминокислотные последовательности  $\beta$ -субъединицы РНК-полимераз некоторых эубактерий, хлоропластов и гомологов этой субъединицы у дрожжей, дрозофилы, человека, археобактерий и вирусов. Выявлены консервативные районы, участвующие в формировании активного центра [3–10].

В данной работе проанализирована структура фрагментов генов *groV*, кодирующих  $\beta$ -субъединицу РНК-полимераз микроорганизмов, входящих в состав байкальского пикопланктона. К этой группе относятся планктонные организмы размером 0.2–2 мкм [11]. Несмотря на важную роль этого компонента в биосистеме озера Байкал [12], пикопланктонные организмы остаются малоизученными. Одним из направлений исследования этого сообщества является анализ первичной структуры генов  $\beta$ -субъединицы РНК-полимераз у организмов, функционирующих в уникальных условиях озера Байкал.

В результате ПЦР-амплификации на суммарной ДНК пикопланктона с праймерами, выбранными на основе последовательности *groV*-гена *Escherichia coli*, были получены фрагменты ДНК длиной около 600 п. о.

Амплифицированные фрагменты ДНК были клонированы, и определены их частичные нуклео-

тидные последовательности. При анализе 93 клонов выявлено 40 различных типов последовательностей.

Полученные нуклеотидные последовательности и последовательности, комплементарные им, переводили в аминокислотные по трем рамкам считывания. При этом для каждой из последовательностей открытой была только одна из шести возможных рамок считывания.

Было проведено сравнение полученных последовательностей со структурой *groV*-генов, депонированных в EMBL-банк. Анализ на нуклеотидном и белковом уровне показал, что все клонированные последовательности гомологичны фрагментам *groV*-генов эубактерий.

Использование для ПЦР праймеров из консервативных районов, соответствие длины полученных продуктов ПЦР длине данного фрагмента *groV*-гена *E. coli* и других эубактерий, наличие единственной открытой рамки считывания и гомологичность клонируемых последовательностей фрагментам генов *groV* эубактерий позволяет считать полученные последовательности ДНК фрагментами *groV*-генов представителей пикопланктона озера Байкал.

На рисунке показаны выведенные из нуклеотидных аминокислотные последовательности участка  $\beta$ -субъединицы РНК-полимераз представителей байкальского пикопланктона.

При сравнении аминокислотных последовательностей  $\beta$ -субъединицы РНК-полимеразы *E. coli* и гомологичных субъединиц РНК-полимераз других организмов (про-, эукариотических и археобактериальных) выявлено 9 консервативных

<sup>#</sup> Автор для переписки (тел.: (3952) 46-06-10).

КЛОНЫ:	1210	1220	1230	1240	1250
4-11	FEDDPSRQQMILCDGRTGDQFERPVTVGVMHVLKLNHLVDDKMHARSTGSPYSLVT				
4-12	KAGIPKNGLVYLNGLSGDRFDQPTTVG I SYMLKLGHMVDDKMHARSTGSPYSLIT				
4-15	KVGVPRSGKTYPYDGGTGDSDQATVGV I HMLKLSHMVDDKMHARSTGSPYSLTT				
4-17	LSNISRTGQTVLYDGRTGEQFDGVPVTVG YMYMLKLNHLVDDKMHARSTGSPYSLVT				
4-19	NQLIGRSGKARLFDGRSGEPYPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
4-26	NQLIGRSGKAKLFDGRSGEPYPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
4-28	NQLIGRSGKARLFDGRSGEPYPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
4-31	KAGIPKNGLVYLNGLSGDRFDQPTTVG I SYMLKLGHMVDDKMHARSTGSPYSLIT				
4-34	NQLIGRSGKARLFDGRSGEPYPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
4-35	TAGIPKPLVYLNGLSGDRFDQPTTVG I SYMLKLGHMVDDKMHARSTGSPYSLIT				
4-38	GQLIGRSGKAKLFDGRSGEPYPTPI SVGNMYILKLNHLVDDKIHARSTGSPNSMIT				
4-39	NQLTGSSSGKARLPDGRPGEPYPTPI RVGYMYILKLNHLVDDKTNERSIGQYLMIT				
4-50	KAGIPKNGLVYLNGLSGDRFDQPTTVG I SYMLKLGHMVDDKMDARSTGSPYSLIT				
4-42	LANISRTGQTVLYDGRTGEQFDGVPVTVG YMYMLKLNHLVDDKMHARSTGSPYSLVT				
4-5	NQLIGRSGKARLFDGRSGEPSPPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
4-52	HAGIPKNGLEYLNGLSGDRFDQPTTVG I SYMLKLGHMVDDKMHARSTGSPYSLIT				
4-53	LPNISRTGQTVLYDGRTGEQFDGVPVTVG YMYMLKLNHLVDDKMHARSTGSPYSLVT				
4-58	NQLIGRSGKARLFDGRSGEPYPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
4-6	MRLTGADGKAKLPDGRTEGEPYQMI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
4-3	LTNISRTGQTVLYDGRTGEQFDGVPVTVG YMYMLKLNHLVDDKMHARSTGSPYSLVT				
2-82	????????????????????VYDNPITVG YMYILKLNHLVDDKIHARSTGSPYSMIT				
3-70	????????????????????HRRCYDNPITVG YMYILKLNHLVDDKIHARSTGSPYSMIT				
3-58	RQSNDRRRRKGASQRSHRRCYDNPITVG YMYILKLNHLVDDKIHARSTGSPYSMIT				
3-67	????????????????RSQSVDRRCYGNPITVG YMYILKLNHLVDDKIHARSTGSPYSMIT				
3-57	NQSIGRSGKAKLFDGRSGEPYPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIK				
3-59	NQLIGRSGKARLFDGRSGEPYPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIK				
2-5	??RYKTGKTPLRNGRTGELYDNPITVG YMYILKLNHLVDDKIHARSTGSPYSMIT				
2-69	????????????????NGLTVDKFDQPTTVG I IYMLKLGHLVDDKMHARSTGSPYSLIT				
2-10	????????????IYMTVKSGDCFHHTKTVG I IYMLKLGHLVDDKMHARSTGSPYSLIT				
2-4	????????????????MVLVSHTRDPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
2-9	HAKSRSPQQPHLFDGRTGEAFERPTTIG YMHFLKLNHLVDDKMHARSTGSPYSLVT				
3-68	NQLIGRSGKARLFDGRSGEPYPTPI SVGYMYILKLNHLVDDKIHARSTGSPYSMIT				
3-65	??TIKTTGKTPLFNGRTGAKYDNDITVG YMYILKLNHLVDDKIHARSTGSPYSMIT				
2-1	????????GKAQLFDWLSRCAYLTLILVGYMY I LKLYHLVDDKIHARSTGSPYSMIT				
2-84	????????????????ATGEVHDNPITVG YMYILKLNHLVDDKIHACSTGSPYLMIM				
2-77	NATCWTKRKSKIIGWRTGQPYKEATVGYMY I LKLNHLVDDKIHARSTGSPYSMIT				
3-71	?????RAATALIRWPHCEAFERPTTIG YMHFLKLNHLVDDKMHARSTGSPHSLK				
2-76	????????RKSITYLKAAPGEAFDRPTTIG YMHFLKLNHLVDDKMHARSTGSPYSLVT				
2-83	???????AATAFIRSSHFEAFERPTTIG YMHFLKLNHLVDDKMHARSTGSPYSLVT				
2-6	????????????????RRCHDNPITVGHTH I SKLNHSVDDKIHARSTGSPYSMIT				
Consensus	...I.R.GK...L.DGR.GE...PITVG YMYILKLNHLVDDKIHARSTGSPYSMIT				
ЭУБАКТЕРИИ:					
<i>E. coli</i>	LGDLPTSGQIRLYDGRTGEQFERPVTVG YMYMLKLNHLVDDKMHARSTGSPYSLVT				
<i>Anabaena</i>	????????????????????AFDRPVTIG VAYMLKLNHLVDDKIHARSTGSPYSLVT				
<i>B. subtilis</i>	EAGMSRDAKTVLYDGRTGEPFDNRVSVG I MYMLKLAHMVDDKIHARSTGSPYSLVT				
<i>M. leprae</i>	DVMVGGDGKAVLFDGRSGEPFPYPTVVG YMYIMKLNHLVDDKIHARSTGSPYSMIT				
ХЛОРОПЛАСТЫ:					
<i>C. reinh.</i>	LLDPNHPGKIRLFDGRNSECFDQTVTVG I AYVCLKVHMVDDKMHARSTGSPYSLVT				
<i>N. tabacum</i>	VFEPEYPGKSRIFDGRTGNPFEPQVPI IGRPYILKLIHQVDDKIHGRSSGHYALVT				
ЭУКАРИОТЫ:					
<i>S. cer.</i>	EHGYQSRGFVEMYNGHTGKKLMAQIFFGPTYYQRLRHMVDDKIHARAGPQVLT				
<i>D. mel.</i>	EYGYHLRGNEVMYNGHTGRKINAQVFLGPTYYQRLKHMVDDKIHARARGPQVILV				
<i>H. sapiens</i>	DYGYHLRGNEVLYNGFTGRKITSQIFIGPTYYQRLKHMVDDKIHARARGPIQILN				
АРХЕБАКТЕРИИ:					
<i>T. celer</i>	ELGFKHSGREVMYDGITGRLEADVFGVI IYYQRLNHMVADKMHARSRGPVQVLT				
<i>S. acid.</i>	LGHLPDSTEV-VYDGRTGQKLSRILFGI VVYQKLNHMVADKMHARARGPQVILT				

Сравнение аминокислотных последовательностей β-субъединицы РНК-полимеразы в районе активного центра у представителей байкальского пикопланктона. Звездочкой отмечен сайт мечения β-субъединицы РНК-полимеразы *E. coli* [13]. Нумерация приведена по первичной структуре *E. coli*. В консенсусной последовательности приведены аминокислотные остатки, встречающиеся в большинстве анализированных последовательностей, инвариантные позиции подчеркнуты; точки – переменные аминокислотные остатки. *Anabaena* sp., *B. subtilis* – *Bacillus subtilis*, *M. leprae* – *Mycobacterium leprae*, *C. reinh.* – *Chlamidomonas reinhardtii*, *N. tabacum* – *Nicotiana tabacum*, *S. cer.* – *Saccharomyces cerevisiae*, *D. mel.* – *Drosophila melanogaster*, *H. sapiens* – *Homo sapiens*, *T. celer* – *Thermococcus celer*, *S. acid.* – *Sulfolobus acidocaldarius*.

в эволюции районов [9, 14]. С помощью метода высокоселективного аффинного мечения было показано, что аминокислотные остатки VIII и IX районов участвуют в связывании иницирующего

субстрата [8, 13]. Так, в β-субъединице РНК-полимеразы *E. coli* Lys-1065 и His-1237 избирательно метятся аналогами иницирующего субстрата [13, 15]. Участие этих аминокислотных остатков в

формировании каталитического центра РНК-полимеразы подтверждено методом сайт-направленных мутаций [15, 16].

Клонированные нами фрагменты *groV*-гена кодируют консервативный район IX, расположенный в С-концевом участке  $\beta$ -субъединицы РНК-полимеразы. Анализ полученных последовательностей показал, что аминокислотный остаток His, соответствующий сайту мечения в структуре  $\beta$ -субъединицы РНК-полимеразы *E. coli* (His-1237), инвариантен во всех клонах. На основе полученных структур выведен консенсус аминокислотных последовательностей (рисунок). Кроме указанного выше инвариантного остатка гистидина во всех клонах неизменными остаются аминокислотные остатки Gly-1228, Lys-1234, Leu-1235, Val-1239, Asp-1240, Asp-1241, Lys-1242, Ser-1247, Gly-1249 (нумерация по структуре *E. coli*). Высокий консерватизм данных аминокислотных остатков позволяет предположить их функциональную значимость.

Для исследуемого образца пикопланктона был проведен филогенетический анализ на основе последовательностей рибосомной РНК малой субъединицы [17], согласно которому представители байкальского пикопланктона входят в состав восьми групп: Cyanobacteria, Actinomycetes, Flavobacteria, Planctomyces,  $\alpha$ -,  $\beta$ -Proteobacteria, Chlorophyta (принадлежность к этой группе организмов определяли по последовательностям генов 16S рРНК хлоропластов) и одну таксономически неопределенную группу. Неполнота банка по нуклеотидным последовательностям *groV*-гена для разных групп бактерий (отсутствуют последовательности этого гена представителей Cyanobacteria,  $\alpha$ -,  $\beta$ -Proteobacteria и др.) не позволила нам четко определить таксономическую принадлежность исследованных клонов. В дальнейшем эта задача может быть решена при определении последовательностей *groV*-гена у чистых культур байкальских микроорганизмов, а также при расширении банка данных по последовательностям этого гена.

Высокий консерватизм полученных структур позволяет использовать их в качестве молекулярных зондов для идентификации малоизученных байкальских бактерий.

### ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

В работе использовали акриламид, бисакриламид, трис(гидроксиметил)аминометан, агарозу (Sigma, США), ДНК-полимеразу из *Thermus aquaticus* (Сибэнзим, Новосибирск), олигодезоксирибонуклеотидные праймеры (Pharmacia, Швеция), ТА-вектор рCR<sup>(TM)</sup>II (In vitro Gene, США).

**Отбор проб воды** осуществлялся 25.07.95 г. на станции мыс Толстый (51°45'N, 104°38'E) с глубины 10 м. Пробу объемом около 10 л фильтровали

через 2-мкм фильтр для отделения от крупных форм. Фильтрат затем концентрировали до конечного объема 50 мл с помощью установки для тангенциальной фильтрации с 0.22-мкм фильтрами и использовали для выделения ДНК.

**Суммарную ДНК** пикопланктона выделяли по стандартной методике с использованием лизоцима, додецилсульфата натрия и протеиназы К с последующей фенольной экстракцией [18].

**Аmplификация фрагмента гена  $\beta$ -субъединицы РНК-полимеразы.** Для амплификации использовали праймеры следующей структуры, выбранные на основе последовательности *E. coli* (депонированной в EMBL-банк под номером V00340):

E3: 5'GGTGACAAGATGGCAGGTCGTCACGG (3187-3212),

E4: 5'TCCATCTCCCGAAACGCTGACCACC (3821-3796).

Аmplификацию осуществляли в 20 мкл буфера, содержащего 10 мМ Трис-НСl (рН 8.9), 40 мМ КСl, 0.1 мг/мл БСА, 2 мМ MgCl<sub>2</sub>, dATP, dCTP, dGTP и dTTP (0.2 мМ каждый), праймеры E3 и E4 (0.5–1 мкМ каждый), 1–10 нг суммарной ДНК пикопланктона, 1–2 ед. акт. *Taq*-ДНК-полимеразы, в течение 35 циклов в следующем режиме: денатурация (94°C, 60 с), отжиг (52°C, 70 с), полимеризация (72°C, 150 с). Использовался амплификатор фирмы "Лина" (Ангарск). Продукты ПЦР анализировали электрофорезом в 1% агарозном геле и выделяли с помощью электроэлюции.

**Клонирование** амплифицированных фрагментов осуществляли в ТА-векторе рCR<sup>(TM)</sup>II. Трансформацию проводили согласно [19]. Отбирали колонии, имеющие lac<sup>-</sup>-фенотип. Клоны, содержащие вставки ДНК нужного размера, выявляли с помощью амплификационного анализа (3 мМ MgCl<sub>2</sub>, температура отжига 55°C, время полимеризации 120 с, остальные условия описаны выше) с использованием универсальных праймеров к плазмиде: (R: 5'TCGTTGTAACGACGGCCAGT, L: 5'CAGCTATGACCATGATTACGCC).

**Первичную структуру** полученных амплификационных фрагментов рекомбинантных клонов определяли с помощью дидезоксинуклеотидных терминаторов в ходе линейной ПЦР согласно Муррей [20].

Полученные нуклеотидные последовательности сравнивали с EMBL-банком для выявления гомологичных последовательностей. Для анализа использовался пакет программ MicroGenie Sequence Analysis Program (version 5.0).

Авторы благодарят директора Лимнологического института СО РАН М.А. Грачева за поддержку работы и большую заинтересованность в проведении исследований пикопланктона озера Байкал.

Мы также благодарим заведующего лабораторией молекулярной энзимологии Е.Ф. Зайчикова за помощь при выполнении данной работы.

## СПИСОК ЛИТЕРАТУРЫ

1. Grachev M.A., Kolocheva T.I., Lukhtanov E.A., Mustaev A.A. // *Eur. J. Biochem.* 1987. V. 163. P. 113–121.
2. Panka D., Dennis P.D. // *J. Biol. Chem.* 1985. V. 260. P. 1427–1431.
3. Ovchinnikov Yu.A., Monastyrskaya G.S., Gubanov V.V., Guryev S.O., Chertov O.Yu., Modyanov N.N., Grinkevich V.A., Makarova I.A., Marchenko T.V., Polovnikova I.N., Lipkin V.M., Sverdlov E.D. // *Eur. J. Biochem.* 1981. V. 116. P. 621–629.
4. Sweetser D., Nonet M., Young R.A. // *Proc. Natl. Acad. Sci. USA.* 1987. V. 84. P. 1192–1196.
5. Klenk H.-P., Zillig W. // *J. Mol. Evol.* 1994. V. 38. P. 420–432.
6. Ohme M., Tanaka M., Chunwongse J., Shinozaki K., Sugiura M. // *FEBS Lett.* 1986. V. 200. P. 87–90.
7. Falkenburg D., Dworniczak B., Faust D.M., Bautz E.K.F. // *J. Mol. Biol.* 1987. V. 195. P. 929–937.
8. Riva M., Carles C., Sentenac A., Grachev M.A., Mustaev A.A., Zaychikov E.F. // *J. Biol. Chem.* 1990. V. 265. P. 16498–16503.
9. Kontermann R., Sitzler S., Seifarth W., Petersen G., Bautz E.K.F. // *Mol. Gen. Genet.* 1989. V. 219. P. 373–380.
10. Скамров А.В., Розовская Т.А., Гольдман М.А., Феофанова Е.С., Бибилашвили Р.Ш. // *Молекулярн. биология.* 1996. Т. 30. С. 61–75.
11. Li W.K.W., Platt T. // *Sci. Prog., Oxf.* 1987. V. 71. P. 117–132.
12. Nagata T., Takai K., Kawanohe K., Kim D.-S., Nakazato R., Guselnikova N., Bondarenko N., Mologawaya O., Kostrova T., Drucker V., Satoh Y., Watanabe Y. // *J. Plankton Res.* 1994. V. 16. P. 945–959.
13. Grachev M.A., Lukhtanov E.A., Mustaev A.A., Zaychikov E.F., Abdukayumov M.N., Rabinov I.V., Richter V.I., Skoblov Y.S., Chistyakov P.G. // *Eur. J. Biochem.* 1989. V. 180. P. 577–585.
14. Sagitov V., Nikiforov V., Goldfarb A. // *J. Biol. Chem.* 1993. V. 268. P. 2195–2202.
15. Mustaev A., Kashlev M., Lee J., Polyakov A., Lebedev A., Zalenskaya K., Grachev M., Goldfarb A., Nikiforov V. // *J. Biol. Chem.* 1991. V. 266. P. 23927–23931.
16. Kashlev M., Lee J., Zalenskaya K., Nikiforov V., Goldfarb A. // *Science.* 1990. V. 248. P. 1006–1009.
17. Семенова Е.А., Кузнецов К.Д. // *Молекулярн. биология.* 1998. Т. 32. С. 583–589.
18. Schmidt T.M., DeLong E.F., Pace N.R. // *J. Bacteriol.* 1991. V. 173. P. 4371–4378.
19. Клонирование ДНК. Методы: Пер. с англ. / Ред. Д. Гловер. М.: Мир, 1988. С. 140–173.
20. Murray V. // *Nucl. Acids Res.* 1989. V. 21. P. 8889.

## Cloning and Sequencing of Fragments of RNA Polymerase $\beta$ -Subunit Genes from Representatives of Baikal Picoplankton

O. A. Suturina, Yu. A. Suturina, and E. A. Semenova<sup>#</sup>

*Lymnological Institute, Siberian Division, Russian Academy of Sciences, ul. Ulan-Batorskaya 3, Irkutsk, 664033 Russia*

Fragments of the RNA polymerase  $\beta$ -subunit genes from representatives of picoplankton of Lake Baikal were cloned using polymerase chain reaction (PCR), and the resulting recombinant clones were sequenced. An analysis of 93 clones revealed 40 different sequences. A comparison of these nucleotide sequences and the corresponding amino acid sequences deduced with their homologues from databanks confirmed the high conservation of the RNA polymerase  $\beta$ -subunit region involved in the formation of the enzyme active site.

*Key words: RNA polymerase  $\beta$ -subunit, active site, sequencing, picoplankton*

<sup>#</sup> To whom correspondence should be addressed; phone: +7 (3952) 46-0610.