



УДК 547.963.32.02

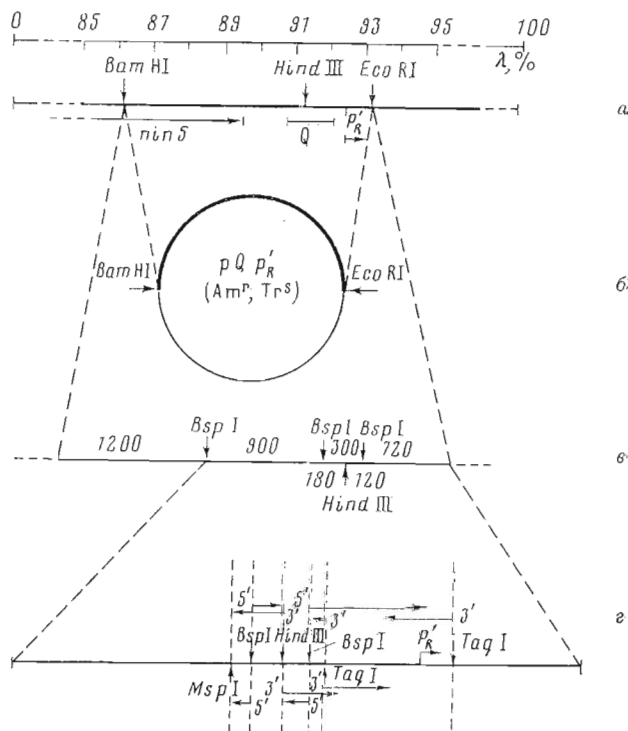
## НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ОБЛАСТИ ДНК ФАГА $\lambda$ , КОДИРУЮЩЕЙ ГЕН Q

*Петров Н. А., Сергунский О. И., Микрюков Н. И.,  
Каргинов В. А., Еравченко В. В.*

*Всесоюзный научно-исследовательский институт молекулярной биологии,  
Новосибирск*

На поздней стадии развития флага  $\lambda$  иницируется транскрипция с промотора  $p'_R$  [1], обуславливающая экспрессию генов, ответственных за сборку фоговой частицы и лизис бактериальной клетки. Для эффективной транскрипции с этого промотора необходим продукт фогового гена Q [2]. Белок Q, по-видимому, не стимулирует синтез РНК с промотора  $p_R$ , а обеспечивает продолжение транскрипции короткой 6S РНК, иницированной с промотора  $p'_R$  [3], т. е. является антитерминатором.

Получение плазмиды  $pQp'_R$ , содержащей ген Q и промотор  $p'_R$ , и стратегия, использованная при установлении первичной структуры гена Q



*a* – физико-генетическая карта ДНК  $\lambda$  в области гена Q и промотора  $p'_R$  по [9];  
*б* – кольцевая карта плазмиды  $pQp'_R$ , полученной *in vitro* на основе ДНК  $pBR322$  (жирной линией выделен  $EcoRI$ - $BamHI$ -фрагмент ДНК флага  $\lambda$ ); *в* – расположение  $BspI$ -сайтов на  $EcoRI$ - $BamHI$ -фрагменте, содержащем ген Q и промотор  $p'_R$ ; *г* – 5'- и 3'-меченные  $^{32}P$  субфрагменты, использованные при анализе первичной структуры. Плазмидную ДНК выделяли как описано в работе [10]. Гидролиз ДНК эндонуклеазами рестрикции и выделение фрагментов из геля проводили как описано в работе [11]

MetArgLeuGluSerValAlaLysPheHisSerProLysSer

5' TGCATTACAGTAAACAGGGAGAAAGGCCATCGAGACTCGAAAGCGTAGCTAAATTCATTCCGCCAAAAGCC  
 3' AACTAAGTCCATTGTCCTTTCCCGGTACTCTGAGCTTTCCGCATCGATTAAACCTAAGCGCTTTTTCG  
 10 20 30 40 50 60 70

ProMetMetSerAspSerProArgAlaThrAlaSerAspSerLeuSerGlyThrAspValMetAlaAla  
 CCGATGATGAGCGACTCACCCAGGGCCACGGCTTCTGACTCTCTTCCGGTACTGATGTGATGGCTGCT  
 GGCTACTACTGGCTGAGTGGCCCGGTGCCGAAGACTGACAGAAAGCGCATGACTACCCATCCGACCGA  
 80 90 100 110 120 130

MetGlyMetAlaGlnSerGlnAlaGlyPheGlyMetAlaAlaPheCysGlyLysHisGluLeuSerGln  
 ATGGGGATGGCGCAATCAACAAGCGGGATTTCGGTATGGCTGCATTTCCGGTAAAGCACCAACTCAGCCAG  
 TAGCCCTACCGGTTAGTCTGCCCTAAGCCATACCGCATAGACCGCATTCCTGCTTTCGATCCGGCTC  
 140 150 160 170 180 190 200

AsnAspLysGlnLysAlaIleAsnTyrLeuMetGlnPheAlaHisLysValSerGlyLysTyrArgGly  
 AACGACAAAACAAAAGGCTATCAACTATCTGATGCAATTTCCACACAAAGGATATCCGGGAAATACCCCGGT  
 TTGCTGTTTTCTTTCCGATAGTGTGATAGACTACGTTAAACGTGTCTTCCATACCCCTTTATGCGCCCA  
 210 220 230 240 250 260 270

ValAlaLysLeuGluGlyAsnThrLysAlaLysValLeuGlnValLeuAlaThrPheAlaTyrAlaAsp  
 CTGGCGAAGCTTGAAGGAAATCTAAGGCAAAAGGCTACTGCAAGTCTCCGCAACATTCCTTATCCGGAT  
 CACCGCTTCGAACCTCCCTTTATGATTCCTTCCATGACGTTCCAGCGCTTCCATCCGCAACCCCTA  
 280 290 300 310 320 330 340

TyrCysArgSerAlaAlaThrProGlyAlaArgCysArgAspCysHisGlyThrGlyArgAlaValAsp  
 TATTGCCGTAGTGGCGCCAGCCGGGGCAAGATGCAAGATTTGCCATGCTACAGGCGCTGCGGTTGAT  
 ATAACGGCATCACCCCGCTGCGGCCCTTCTACGCTCTTAACGGTATCCATCCCGCCAGCCCAACTA  
 350 360 370 380 390 400 410

IleAlaLysThrGluLeuTrpGlyArgValValGluLysGluCysGlyArgCysLysGlyValGlyTyr  
 ATTGCCAAAACAGAGCTGTGGGGAGACTTCTCCAGAAAGACTGCGGAACATGCCAAAGGCTCCGGCTAT  
 TAACGTTTTGTCTCGACACCCCTCTCAACAGCTCTTTCTCACGCGCTTCTTCCGCAACCCGATA  
 420 430 440 450 460 470 480

SerArgMetProAlaSerAlaAlaTyrArgAlaValThrMetLeuIleProAsnLeuThrGlnProThr  
 TCAAGGATGCCAGCAAGCGCAGCATATCGCGCTGTGACGATGCTAATCCCAAACCTACCCAAACCCACC  
 ACTTCCATCGGTCTGCTGCGTGTATAGCCCGACACTGCTACGATTAGGCTTTGGAATGGTGGGCTG  
 490 500 510 520 530 540 550

TrpSerArgThrValLysProLeuTyrAspAlaLeuValValGlnCysHisLysGluGluSerIleAla  
 TGGTCAAGCACTGTAAGCCGCTGTATGACCGCTCTGGTGTGCAATGCCACAAAAGAGTCAATCCGA  
 ACCAGTCCGTCGACAAATTCGGCGACATACTGCGAGACCCACCACTGCGTTCGTTCTTCTCAGTTAGGCT  
 560 570 580 590 600 610 620

AspAsnIleLeuAsnAlaIleThrArgEnd  
 GACAAATTTTGAATCCGATCAGCTTACGAGCATGATCCCCACGGATGGCAGATATTAACGGCATGA  
 CTGTTGTAACAATCAGTCTGTCGCAATCGTCTACTAGCGTCCCTACCGTGTATAATTCGCTACT  
 630 640 650 660 670 680 690

TATTGACTTATTGAATAAAAATTGGCTAAATTTGACTCAACCGATGGGTTAATTCGCTCGGT 3'  
 ATAACSTAATAACTAATTTTAAACCCATTTAAACTGAGTTGCTACCCAAATTAAGCGAGCCA 5'

700 710 720 730 740 750

Нуклеотидная последовательность области ДНК фага  $\lambda$ , содержащей ген Q и начало транскриптора 6S РНК, и возможная аминокислотная последовательность белка Q. Точками помечены нуклеотиды, комплементарные 3'-концу 16S РНК, так называемая SD-последовательность [12]. Подчеркнуты последовательности промотора  $p_R'$  в районе -10-го и -35-го нуклеотидов от начала 6S РНК

Настоящая работа посвящена установлению полной первичной структуры гена Q. Дополнительно мы приводим уточненную структуру промотора  $p_R'$ , определенную нами ранее [4], которая совпадает с опубликованной позднее в работе [5]. Данная структура была уточнена при определении нуклеотидной последовательности начиная от более близко расположенных к промоторной области сайтов рестрикционных эндонуклеаз *TaqI*, *HinfI* и *TaqXI*.

*EcoRI*+*BamHI*-фрагмент ДНК фага  $\lambda$ , содержащий ген Q и транскриптор 6S РНК, клонировали в составе pBR322 (схема a, б) и его <sup>32</sup>P-меченные субфрагменты (схема в, г) использовали для установления первичной структуры гена Q и промотора  $p_R'$ . Получение меченых фрагментов ДНК и определение последовательности нуклеотидов этих фрагментов проводили как описано ранее [4].

Ошибки, возможные при определении первичной структуры ДНК этим методом, объясняются главным образом феноменом «компрессии» [6], вызываемым неполной денатурацией фрагментов ДНК при электрофорезе в полиакриламидном геле, а также присутствием незначительного количества минорного основания — 5-метилцитозина [7] в последовательностях CCAGG и CCTGG ДНК плазмиды, выращенной на штамме *E. coli*

С600. Поэтому более надежным является определение структуры ДНК по обоим комплементарным цепям.

В настоящей работе для выявления последовательностей ССАГГ и ССТГГ применяли рестрикционную эндонуклеазу *Taq*XI [8], гидролизующую ДНК именно в этих участках, что исключало возможность ошибки из-за наличия мизорного основания. Отдельные участки нуклеотидной последовательности гена Q были определены по одной цепи фрагмента ДНК. В этом случае для повышения надежности определения структуры проводили дважды с противоположных концов каждой цепи ДНК, а эксперименты дублировались. Для устранения «компрессии» экспериментально подбирали оптимальные условия электрофоретического разделения продуктов частичного расщепления ДНК, включая напряжение и процентное содержание акриламида и метиленбисакриламида в разделяющем полиакриламидном геле. Наличие непрерывной «рамки трансляции» в области ДНК, которая должна кодировать последовательность белка Q, косвенно свидетельствует о надежности определения структуры.

На рисунке приведена первичная структура участка ДНК длиной 750 пар оснований. В начале этой последовательности имеется участок, комплементарный 3'-концу 16S рРНК, расположенный на расстоянии пяти нуклеотидов от иницирующего AUG-кодона. При трансляции мРНК с этого иницирующего кодона может образоваться полипептид длиной в 207 аминокислотных остатков (рисунок), имеющий молекулярную массу 23 114. По данным электрофореза в полиакриламидном геле [9], подвижность белка Q соответствует полипептиду с молекулярной массой 23 000, что согласуется с величиной, выведенной нами из первичной структуры ДНК.

В соответствии с генетической картой ДНК ффага  $\lambda$  за геном Q следует промотор  $p'_R$  [9]. На рисунке видно, что в нескольких десятках (около 40) пар оснований от терминирующего кодона TAG расположена последовательность промотора  $p'_R$  и начало 6S РНК. Этот факт указывает также на то, что мы определили первичную структуру гена Q.

Итак, совокупность приведенных выше данных свидетельствует о том, что установлена полная нуклеотидная последовательность гена Q — анти-терминатора транскрипции с промотора  $p'_R$ .

#### ЛИТЕРАТУРА

1. Roberts J. Proc. Nat. Acad. Sci. USA, 1975, v 72, № 9, p. 3300–3304.
2. Herskowitz I. Ann. Rev. Genet., 1974, v. 7, p. 289–324.
3. Кравченко В. В., Василенко С. К., Гилева И. П., Зайцев Б. Н. Биоорг. химия, 1980, т. 6, № 2, с. 217–222.
4. Петров Н. А., Каргинов В. А., Микрюков Н. Н., Серпинский О. И., Кравченко В. В., Василенко С. К. Биоорг. химия, 1979, т. 5, № 12, с. 1873–1875.
5. Siebenlist U., Simpson R. B., Gilbert W. Cell, 1980, v. 20, № 2, p. 269–281.
6. Kramer F. R., Mills D. R. Proc. Nat. Acad. Sci. USA, 1978, v. 75, № 11, p. 5334–5338.
7. Maxam A. M., Gilbert W. G. Proc. Nat. Acad. Sci. USA, 1977, v. 74, № 2, p. 560–564.
8. Грачев С. А., Мамаев С. В., Гуревич А. И., Игошин А. В., Колосов М. Н., Слюсаренко А. Г. Биоорг. химия, 1981, т. 7, № 4, с. 628–630.
9. Szybalski E. H., Szybalski W. A. Gene, 1979, v. 7, № 3–4, p. 217–270.
10. Кравченко В. В., Василенко С. К., Гилева И. П. Докл. АН СССР, 1980, т. 252, № 2, с. 1255–1257.
11. Kravchenko V. V., Vasilenko S. K., Grachev M. A. Gene, 1979, v. 7, № 3–4, p. 181–195.
12. Shine J., Dalgarno I. Proc. Nat. Acad. Sci. USA, 1974, v. 71, № 4, p. 1342–1346.

Поступило в редакцию  
31.VIII.1981

#### NUCLEOTIDE SEQUENCE OF THE BACTERIOPHAGE $\lambda$ DNA REGION CONTAINING GENE Q

PETROV N. A., SERPINSKI O. I., MIKRYUKOV N. N.,  
KARGINOV V. A., KRAVCHENKO V. V.

All-Union Research Institute of Molecular Biology, Novosibirsk

The primary structure of  $\lambda$  DNA region between 90,8% and 92,6% of the  $\lambda$  genome length has been determined. This region contains gene Q and promoter  $p'_R$ . The sequence at the right side of the promoter  $p'_R$  can be translated up to the terminator, and polypeptide (product of gene Q) with the length of 207 amino acids-residues can be formed.