



ПИСЬМА РЕДАКТОРУ

УДК 577.21 : 578.81
© 1991 г.

Р. Г. Нивинскас, А. А. Раудоникене, Р. И. Вайшкунайте

НУКЛЕОТИДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ ГЕНА 26 БАЗАЛЬНОЙ ПЛАСТИНКИ БАКТЕРИОФАГА Т4

Институт биохимии Литовской АН, Вильнюс

Определение первичной структуры гена 26 вызывает интерес в связи с неординарностью его экспрессии — речь идет как о направлении транскрипции, так и о наличии двух продуктов. Этот ген является поздним геном фага, локализованным в группе генов базальной пластинки в последовательности 25, 26, 51, 27, 28 и 29. При индуцировании транскрипции с P_L -промотора рекомбинантных плазмид pRL705 и pRL707, полученных в результате введения *Bgl*II-фрагмента ДНК с упомянутыми генами базальной пластинки фага в состав векторной плазмиды в различной ориентации [1], в опытах по комплементации *in vivo* [2], а также при анализе продуктов экспрессии генов [3], нами установлено, что в отличие от генов 51, 27, 28 и 29 гены 26 и 25 транскрибируются против часовой стрелки на генетической карте фага, т. е. в направлении транскрипции ранних генов фага Т4. Такое направление в отношении гена 25 было подтверждено и при определении его нуклеотидной последовательности [4]. Следовательно, в участке между генами 26 и 51 происходит инициация транскрипции в противоположных направлениях. Продукт гена 26 был идентифицирован в структуре центральной части базальной пластинки как белок с мол. массой 40—41 кДа [5—7]. В то же время при экспрессии гена 26, находящегося в составе рекомбинантной плазмиды, его продукт идентифицирован нами как белок с мол. массой 24 кДа [3]. Более того, используя для экспрессии гена двухплазмидную систему фага Т7, мы показали, что продуктами экспрессии гена 26 являются два полипептида с мол. массами 24 и 8—9 кДа, причем полипептид размером 8—9 кДа кодируется концевой частью гена 26 [8]. Таким образом, для изучения представляет несомнен-

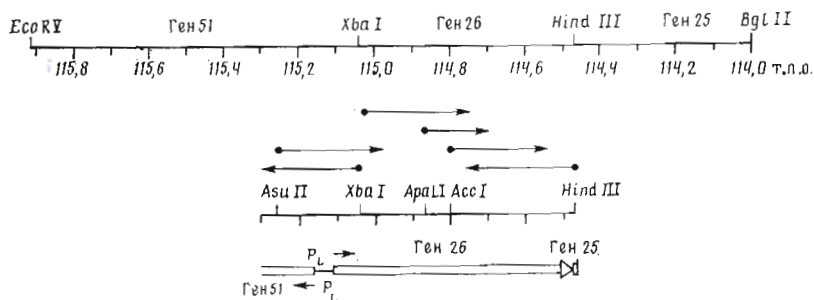


Рис. 1. Локализация гена 26 на генетической карте фага Т4 и стратегия его секвенирования

```

60
*
AsuII
CTTTTGTCTTGTGGTGAC CGATGTTCTATATCGTTCG AACTAACAAAAATCTCGAT
120
*
AATCTTCTACCGTAAATGGT TTAACCGATGAACACCATC TGSTAATTTACAACGAATAA
180
*
P -->
L SD
TGTTCGTCATAGATGCTCCT TTTATTCCTATTATATAATAT GATAAATAAGGAGCTAAAT
228
gp51 SD <--P
gp26 L
ATG TAT GAA TAC AAA TTT GAT GTG AGA GTT GGT TCT AAA ATA ATC AAT
Met Tyr Glu Tyr Lys Phe Asp Val Arg Val Gly Ser Lys Ile Ile Asn
276
*
XbaI
TGT CGC GCA TTC ACG CTT AAA GAA TAT CTA GAA CTT ATT ACT GCC AAA
Cys Arg Ala Phe Thr Leu Lys Glu Tyr Leu Glu Leu Ile Thr Ala Lys
324
*
AAT AAT GGT TCC GTA GAA GTA ATT GTT AAA AAG CTA ATC AAA GAC TGC
Asn Asn Gly Ser Val Glu Val Ile Val Lys Lys Leu Ile Lys Asp Cys
372
*
ACA AAT GCA AAA GAT TTA AAC CGC CAA GAA TCA GAA CTA TTG TTG ATT
Thr Asn Ala Lys Asp Leu Asn Arg Gln Glu Ser Glu Leu Leu Leu Ile
420
amS105 amNG114
CAT TTA TGG GCA CAT TCT CTC GGT GAA GTT AAT CAC GAA AAC TCC TGG
His Leu Trp Ala His Ser Leu Gly Glu Val Asn His Glu Asn Ser Trp
468
*
ApaLI
AAG TGC ACC TGT GGA ACT GAA ATA CCA ACC CAT ATA AAT CTA TTA CAT
Lys Cys Thr Cys Gly Thr Glu Ile Pro Thr His Ile Asn Leu Leu His
516
*
amN131 AccI
ACA CAA ATA GAT GCA CCA GAA GAC CTC TGG TAT ACA CTA GGT GAC ATT
Thr Gln Ile Asp Ala Pro Glu Asp Leu Trp Tyr Thr Leu Gly Asp Ile
564
*
AAA ATT AAA TTC CGA TAC CCT AAA ATT TTT GAT GAT AAA AAT ATA GCC
Lys Ile Lys Phe Arg Tyr Pro Lys Ile Phe Asp Asp Lys Asn Ile Ala
612
*
CAC ATG ATA GTA TCA TGT ATA GAA ACG ATT CAT GCT AAC GGG GAA AGC
His Met Ile Val Ser Cys Ile Glu Thr Ile His Ala Asn Gly Glu Ser
660
*
ATT CCA GTT GAA GAC TTA AAT GAA AAG GAA CTA GAA GAT TTA TAT TCT
Ile Pro Val Glu Asp Leu Asn Glu Lys Glu Leu Glu Asp Leu Tyr Ser
708
*
ATC ATC ACA GAG TCA GAT ATT GTA GCT ATA AAA GAT ATG CTT TTA AAG
Ile Ile Thr Glu Ser Asp Ile Val Ala Ile Lys Asp Met Leu Leu Lys
756
*
CCT ACC GTT TAT TTG GCT GTT CCA ATT AAG TGT CCA GAG TGT GGA AAA
Pro Thr Val Tyr Leu Ala Val Pro Ile Lys Cys Pro Glu Cys Gly Lys
804
*
ACC CAT GCT CAC GTA ATA AGA GGC CTC AAA GAG TTC TTT GAG TTA CTA
Thr His Ala His Val Ile Arg Gly Leu Lys Glu Phe Phe Glu Leu Leu
827
gp25 HindIII
TAATG GCA AAT ATT AAT AAG CTT
***
Met Ala Asn Ile Asn Lys Leu

```

Рис. 2. Нуклеотидная последовательность фрагмента ДНК фага Т4 с геном 26. Подчеркнуты сайты действия рестриктаз *Asu*II, *Xba*I, *Apa*LI, *Acc*I и *Hind*III, потенциальные участки локализации амбер-мутаций, а также последовательность Шайна—Далгарно (SD). Обведены специфические последовательности поздних промоторов (P_L) генов 26 и 51 фага Т4

ный интерес как структура регуляторной области между генами 26 и 51, так и сама последовательность нуклеотидов гена 26.

В настоящей работе для определения нуклеотидной последовательности гена 26 методом Сенгера [9] различные фрагменты ДНК фага Т4 были выщеплены из упомянутых рекомбинантных плазмид и введены в состав

фаговых векторов M13mp10 и M13mp11. Стратегия секвенирования изображена на рис. 1. Нуклеотидная последовательность ДНК длиной в 827 нуклеотидов представлена на рис. 2. Видно, что кодирующая белок последовательность гена 26 состоит из 624 нуклеотидов. Она завершается терминирующим кодоном ТАА, последний нуклеотид которого является первым нуклеотидом иницирующего кодона гена 25. В структуре гена 26 обнаружены три потенциальных амбер-участка (TGG), причем локализация конкретных амбер-мутаций в этих участках основана на последовательности их расположения, установленной нами ранее в опытах по спасению маркера с использованием делеционных производных плазмид [8]. На расстоянии шести нуклеотидов от иницирующего кодона гена 26 расположен регуляторный сигнал трансляции — последовательность Шайна — Далгарно, а перед ней — промотор гена 26, содержащий типичную для поздних промоторов фага Т4 последовательность ТАТАААТА [10, 11]. Поздний промотор гена 26 перекрывается с поздним промотором гена 51, в случае которого синтез мРНК иницируется с комплементарной нити ДНК по направлению, характерному для транскрипции поздних генов фага Т4. Секвенированный фрагмент содержит 130 нуклеотидов белок-кодирующей последовательности гена 51, находящейся на расстоянии всего лишь 50 нуклеотидов от гена 26.

Нуклеотидная последовательность гена 26 кодирует полипептид, состоящий из 208 аминокислотных остатков с мол. массой 23,9 кДа. Полипептид содержит 32 кислые и 33 основные аминокислоты, вычисленная изоэлектрическая точка pI 6,06. Таким образом, мол. масса продукта гена 26, установленная нами ранее при его экспрессии в составе рекомбинантных плазмид, хорошо согласуется с мол. массой, вычисленной исходя из нуклеотидной последовательности этого гена. Видимо, в структуре базальной пластинки белок 40—41 кДа, обнаруживаемый как продукт гена 26 [5—7], фактически представляет собой комплекс полипептида *M* 23,9 кДа с другими полипептидами, возможно и с полипептидом *M* 8—9 кДа, кодируемым концевой частью гена 26. В установленной нуклеотидной последовательности гена 26 не обнаруживается отдельной открытой рамки считывания для кодирования полипептида *M* 8—9 кДа; следовательно, этот полипептид, как и предполагалось [8], кодируется концевой частью той же самой открытой рамки считывания гена 26. Использование АТГ в положении 568 как иницирующего кодона для синтеза этого небольшого полипептида сомнительно, так как в этом случае последовательность Шайна — Далгарно была бы представлена всего лишь двумя нуклеотидами (AG). В качестве иницирующего кодона могут служить и некоторые другие кодоны [12], поэтому окончательный ответ на этот вопрос даст только выделение этого небольшого второго полипептида гена 26 и определение его N-концевой аминокислотной последовательности.

СПИСОК ЛИТЕРАТУРЫ

1. Нивинскас Р. Г. // Генетика. 1988. Т. 24. № 1. С. 34—41.
2. Клауса В. И., Нивинскас Р. Г. // Генетика. 1988. Т. 24. № 1. С. 42—51.
3. Клауса В. И., Нивинскас Р. Г. // Генетика. 1990. Т. 26. № 2. С. 359—362.
4. Gruidl M. E., Canan N. C., Mosig G. // Nucl. Acids Res. 1988. V. 16. № 20. P. 9862.
5. Kozloff L. M. // Bacteriophage assembly / Ed. Dubow M. N. Y.: Alan R. Liss, Inc. 1981. P. 327—342.
6. Kozloff L. M., Zorzopulos J. // J. Virol. 1981. V. 40. № 3. P. 635—644.
7. Kozloff L. M., Lute M. // J. Virol. 1984. V. 52. № 2. P. 344—349.
8. Вайнштейн Р. И., Нивинскас Р. Г. // Молекулярн. биология. 1990. Т. 24. Вып. 2. С. 379—390.
9. Sagner F., Nicklen S., Coulson A. R. // Proc. Nat. Acad. Sci. USA. 1977. V. 74. № 12. P. 5463—5467.
10. Christensen A. C., Young E. T. // Bacteriophage T4 / Eds Mathews C. K., Kutner E. M., Mosig G., Berget P. B. Washington, D. C.: Amer. Soc. Microbiol. 1983. P. 184—188.

11. Kassavetis G. A., Zentner P. G., Geiduschek E. P. // J. Biol. Chem. 1986. V. 261. № 30. P. 14256—14265.
12. Stormo G. D. // Maximizing gene expression // Eds Reznikoff W., Gold L. Stoneham: Butterworth Publ., 1986. P. 195—224.

Поступило в редакцию
20.XI.1989

После доработки
13.VI.1990

R. H. NIVINSKAS, A. A. RAUDONIKIENE, R. J. VAISKUNAITE
SEQUENCE OF THE BACTERIOPHAGE T4 BASEPLATE GENE 26

Institute of Biochemistry, Lithuanian Academy of Sciences, Vilnius

We have determined the sequence of a 827 bp region comprising gene 26 of bacteriophage T4. The coding region consists of 624 bp directing the synthesis of a polypeptide 208 amino acid residues long, with a calculated molecular mass of 23.9 kDa. The upstream sequence contains the consensus sequence for T4 late promoter.