



УДК 578.832.1А:578.224'212.4:577.215

**СИНТЕЗ, КЛОНИРОВАНИЕ И ОПРЕДЕЛЕНИЕ ПЕРВИЧНОЙ СТРУКТУРЫ ПОЛНОРАЗМЕРНОЙ ДНК-КОПИИ ГЕНА БЕЛКА NP ВИРУСА ГРИППА А**

*Беклэмиев А. Б., Блинз В. М., Василенко С. К.,  
Головин С. Я., Каргинов В. А., Мамаев Л. В.,  
Микрюков Н. Н., Нетесов С. В., Петренко В. А.,  
Петров Н. А., Фролов И. В.*

*Всесоюзный научно-исследовательский институт молекулярной биологии  
Главмикробиопрома, пос. Кольцово Новосибирской обл.*

Проведены синтез, клонирование и определение первичной структуры полноразмерной ДНК-копии фрагмента 5 генома вируса гриппа типа А. Оценена степень различий в нуклеотидной последовательности между двумя лабораторными вариантами штамма А/PR/8/34. Обсуждается вероятная вторичная структура фрагмента 5.

Ранее нами была определена полная нуклеотидная последовательность генов гемагглютинина (НА) и нейраминидазы (НА) вакцинного штамма вируса гриппа Х/Ленинград/54/1 подтипа Н1N1 [1, 2], который является рекомбинантом штаммов А/PR/8/34 и А/Хабаровск/34457/77, отобранным по антигенным свойствам штамма А/Хабаровск/34457/77. Мы поставили задачу определить принадлежность всех генов рекомбинантного генома к одному из двух родительских штаммов. С этой целью, а также для оценки степени варибельности клонов кДНК вируса гриппа А в настоящей работе определена полная нуклеотидная последовательность фрагмента 5 генома штамма Х/Ленинград/54/1, кодирующего структурный белок нуклеопротеида NP.

Подробные методики синтеза и клонирования кДНК приведены в работе [1]. Методом гибридизации с РНК были отобраны клоны, содержащие последовательности генов гемагглютинина и нейраминидазы, оставшиеся рекомбинантные плазмидные ДНК были подвергнуты картированию с помощью рестриктаз *PstI*, *AvaII*, *MspI*, *BamHI*, *HindIII*, *BglII*. На основании этого анализа и предварительного секвенирования было отобрано пять независимых клонов, содержащих последовательность гена белка NP и обозначенных 64-31А, 04-20, 64-31, 64-26 и 22-19.

Первичную структуру вставок кДНК в этих клонах определяли модифицированным методом Максама — Гилберта [1, 3]. Только рекомбинантная ДНК из клона 64-31А содержала полную копию гена белка NP, что было доказано секвенированием концевых участков вставки.

По многочисленным литературным данным, при синтезе и клонировании кДНК иногда наблюдаются минорные клоны, ДНК из которых различаются не только длиной, но и заменами или пропусками одного или нескольких нуклеотидов, а также перегруппировками различных участков кДНК. Для выявления таких нарушений, а также для выяснения их причины нами было проведено определение первичной структуры для ДНК из нескольких независимо полученных клонов. Так, 50% последовательности гена белка NP было определено в клоне 64-31А, 55% — в 04-20, 20% — в 64-26 и 75% — в клоне 64-31, а всего более 80% последовательности было установлено в двух и более независимых клонах. Схематически стратегия секвенирования изображена на рис. 1. Более 80% последовательности определено по двум комплементарным цепям.

В результате расшифрована полная нуклеотидная последовательность гена белка NP штамма Х/Ленинград/54/1 длиной 1565 нуклеотидов

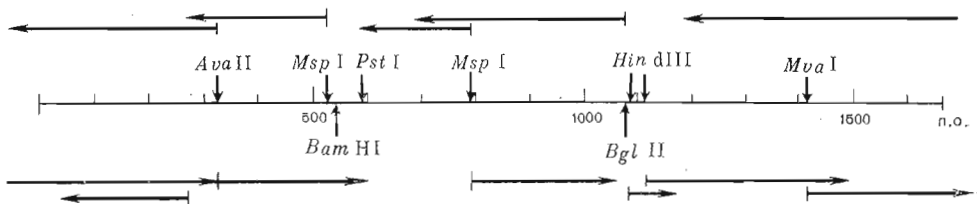


Рис. 1. Схема стратегии секвенирования. Отмечены только сайты узнавания рестриктаз, использованные при определении структуры. Стрелки показывают направление и размеры секвенированной области меченых по 3'-концу фрагментов

(рис. 2). Она имеет единственную протяженную рамку трансляции для белка размером 498 аминокислотных остатков.

Секвенированные клоны ДНК обладают рядом особенностей, по-видимому, характерных для метода клонирования с использованием обратной транскриптазы. Так, клон 04-20 содержит последовательность гена белка NP с 1-го по 873-й нуклеотид. Обрыв приходится на центр последовательности 866–881, представляющей собой четырехкратный повтор олигонуклеотида ССТГ. Остальные клоны содержали последовательность примерно от района 300–500 до 1442-го нуклеотида, вслед за которым был расположен повтор инвертированного участка последовательности с 287-го по 1-й нуклеотид. Такое строение клонов, очевидно, обусловлено частичной взаимной комплементарностью удаленных участков последовательности и может происходить либо в результате рекомбинации при клонировании кДНК, либо в результате ошибки ревертазы.

Вторая причина кажется более вероятной, поскольку подобные перегруппировки были замечены лишь при клонировании NP-гена, тем более что аналогичные факты наблюдались при клонировании этого гена и другими авторами. Так, Винтер с сотр. [4] описывает обнаруженные при клонировании кДНК гена белка NP вируса гриппа А/PR/8/34 инвертированные повторы участков кДНК размером 150–200 нуклеотидов. Авторы объясняют это наличием в исходной РНК коротких самокомплементарных участков, из-за которых и происходит неправильное копирование РНК ревертазой, и подробно описывают возможные механизмы этого явления [4].

На основе этих данных, а также собственных аналогичных результатов мы предположили, что РНК гена белка NP имеет весьма прочную вторичную структуру, причем концы двуспиральных участков в ней являются теми местами, в которых и «ошибается» ревертаза.

На рис. 3 в форме (+)-цепи изображена вторичная структура протяженного участка гена белка NP, построенная на основе экспериментальных данных о строении минорных клонов кДНК и состоящая из трех областей. Наличие области I в районе концов РНК предполагается еще в работе [5]. Данные, указывающие на существование области II, получены в настоящей работе, а аналогичные данные, касающиеся области III, приведены в работе [4].

Петли в изображенной на рис. 3 структуре, вероятно, также структурированы и взаимодействуют между собой. Так, образование комплементарной связи между нуклеотидами 816–819 из петли «580» (см. рис. 3) и петлей «4» (968–971) установлено в работе [4].

Как видно из рис. 3, имеющихся экспериментальных данных о структуре минорных клонов кДНК гена белка NP вируса гриппа явно недостаточно для точного построения его вторичной структуры. Однако ценность уже имеющихся результатов сомнению не подлежит, и, по-видимому, изучение искажений нативной последовательности РНК в минорных клонах ее кДНК может явиться не менее действенным методом в исследовании вторичной структуры РНК, чем традиционные методы определения двуспиральных участков РНК.

Что касается гетерогенности исходной РНК, то нами обнаружены два различия в нуклеотидной последовательности гена белка NP из различных клонов:

MetAlaSerGlnGly 5

AGCAAAAGCAGGGTAGATAATCACTCACTGAGTGACATCAAATCATGGCGTCTCAAGGC 60  
C

ThrLysArgSerTyrGluGlnMetGluThrAspGlyGluArgGlnAsnAlaThrGluIle 25  
ACCAAACGATCTTACGAACAGATGGAGACTGATGGAGAACCAGCAATGCCACTGAAATC 120  
G

ArgAlaSerValGlyLysMetIleGlyGlyIleGlyArgPheTyrIleGlnMetCys\*Thr 45  
AGAGCATCCGTCGGAAAAATGATGGTGGAAATGGACGATCTACATCCAATGTGCACC 180  
A

GluLeuLysLeuSerAspTyrGluGlyArgLeuIleGlnAsnSerLeuThrIleGluArg 65  
GAACTCAAACCTCAGTGATTATGAGGGACGTTGATCCAAAACAGCTTAACAATAGAGAGA 240  
T

MetAlaLeuSerAlaPheAspGluArgArgAsnLysTyrLeuGluGluHisProSerAla 85  
ATGGTGCTCTCTGCTTTTGACGAAAGGAGAAATAAATACCTTGAAGAACATCCCAGTGCG 300  
G

GlyLysAspProLysLysThrGlyGlyProIleTyrArgArgValAsnGlyLysTrpMet 105  
GGAAAGATCCTAAGAAAACCTGGAGGACCTATATACAGGACAGTAAACGGAAAGTGGATG 360  
G A

ArgGluLeuIleLeuTyrAspLysGluGluIleArgArgIleTrpArgGlnAlaAsnAsn 125  
AGAGAACTCATCCTTTATGACAAAGAAGAAATAAGGCGAATCTGGCGCCAAGCTAATAAT 420

GlyAspAspAlaThrAlaGlyLeuThrHisMetMetIleTrpHisSerAsnLeuAsnAsp 145  
GGTGACGATCAACGGCTGGTCTGACTCACATGATGATCTGGCATTCCAATTTGAATGAT 480

AlaThrTyrGlnArgThrArgAlaLeuValArgThrGlyMetAspProArgMetCys\*Ser 165  
GCAACTTATCAGAGGACAAGAGCTCTTGTTCGCACCGGAATPGATCCAGGATGTGCTCT 540  
G

LeuMetGlnGlySerThrLeuProArgArgSerGlyAlaAlaGlyAlaAlaValLysGly 185  
CTGATGCAAGGTTCAACTCTCCCTAGGAGGCTGGAGCCCGAGGTGTCAGTCAAAGGA 600

ValGlyThrMetValMetGluLeuValArgMetIleLysArgGlyIleAsnAspArgAsn 205  
GTTGGAAACAATGGTGATGGAATTGGTCAGAATGATCAAACCTGGGATCAATGATCGGAAC 660  
G

PheTrpArgGlyGluAsnGlyArgLysThrArgIleAlaTyrGluArgMetCys\*AsnIle 225  
TTCTGGAGGGGTGAGAATGGACGAAAAACAAGAAATGCTTATGAAAGAAATGTGCAACATT 720

LeuLysGlyLysPheGlnThrAlaAlaGlnLysAlaMetMetAspGlnValArgGluSer 245  
CTCAAAGGGAAATTTCAAACCTGCTGCACAAAAAGCAATGATGGATCAAGTGAGAGAGAGC 780

Asp  
ArgAsnProGlyAsnAlaGluPheGluAspLeuThrPheLeuAlaArgSerAlaLeuIle 265  
CGGAACCCAGGGAATGCTGAGTTCGAAGATCTCACTTTTCTAGCACGGTCTGCACCTATA 840  
G

LeuArgGlySerValAlaHisLysSerCys\*LeuProAlaCys\*ValTyrGlyProAlaVal 285  
TTGAGAGGGTTCGGTTGCTCACAAGTCTGCCTGCCTGCCTGTGTGTATGGACCTGCCGTA 900

AlaSerGlyTyrAspPheGluArgGluGlyTyrSerLeuValGlyIleAspProPheArg 305  
GCCAGTGGGTACGACTTGGAAAGGGAGGGATACCTCTAGTCGGGAATAGACCCTTTCAGA 960  
A

LeuLeuGlnAsnSerGlnValTyrSerLeuIleArgProAsnGluAsnProAlaHisLys 325  
CTGCTTCAAACAGCCAAGTGTACAGCCCTAATCAGACCAAATGAGAATCCAGCACACAAG 1020

SerGlnLeuValTrpMetAlaCys\*HisSerAlaAlaPheGluAspLeuArgValLeuSer 345  
AGTCAACTGGTGTGGATGCATGCCATTCTGCCGATTTGAAGATCTAAGAGTATTAAGC 1080  
G

Val  
PheIleLysGlyThrLysValLeuProArgGlyLysLeuSerThrArgGlyValGlnIle 365  
TTCATCAAACGGGACGAAGGTGCTCCCAAGAGGGAAGCTTCCACTAGAGGAGTTCAAATT 1140  
G

AlaSerAsnGluAsnMetGluThrMetGluSerSerThrLeuGluLeuArgSerArgTyr 385  
GCTTCCAATGAAAATATGGAGACTATGGAATCAAGTACACTTGAAGTGAAGAAGCAGGTAC 1200

TrpAlaIleArgThrArgSerGlyGlyAsnThrAsnGlnGlnArgAlaSerAlaGlyGln 405  
TGGGCCATAAGGACCAGAACTGGAGGAAACACCAATCAACAGACGGCATCTGCGGGCCAA 1260

1) в ДНК клона 04-20 в районе нуклеотида 750 имеется последовательность из четырех остатков А, в то время как в ДНК клона 64-31 и нуклеотидной последовательности гена белка NP штамма А/PR/8/34 [6] — пять. По-видимому, «потеря» одного основания обусловлена ошибкой обратной

IleSerIleGlnProThrPheSerValGlnArgAsnLeuProPheAspArgThrThrVal 425  
ATCAGCATACAAGCTACGTTCTCAGTACAGAGAAATCTCCCTTTTGACAGAACAACCGTT 1320

Thr  
MetAlaAlaPheSerGlyAsnThrGluGlyArgThrSerAspMetArgThrGluIleIle 445  
ATGGCAGCATTCAGTGGGAATACAGAGGGGAGAACATCTGCATCAGGACCGAAATCATA 1380  
C

ArgMetMetGluSerAlaArgProGluAspValSerPheGlnGlyArgGlyValPheGlu 465  
AGGATGATGGAAAGTCCAAGACCAGAAGATCTGTCTTTCCAGGGCGGGGAGTCTTCGAG 1440

LeuSerAspGluLysAlaAlaSerProIleValProSerPheAspMetSerAsnGluGly 485  
CTCTCGGACGAAAAGGCAGCGAGCCCGATCGTGCCTTCTTTGACATGAGTAATGAAGGA 1500

SerTyrPhePheGlyAspAsnAlaGluGluTyrAspAsn  
TCTTATTTCTTCGGAGACAAATGCCAGGAAATACGATAATTAAGAAAAATACCCCTTGTTF 1560  
C

СТАСТ

Рис. 2. Нуклеотидная последовательность гена белка NP из штамма X/Ленинград/54/1 в форме (+)-цепи и выведенная на ее основании аминокислотная последовательность. Внизу приведены нуклеотидные замены по сравнению с соответствующим геном штамма A/PR/8/34 варианта Cambridge, а сверху — отвечающие им аминокислотные замены

транскриптазы, поскольку это приводит к преждевременной терминации рамки трансляции. Аналогичное явление отмечалось в работе [7];

2) в последовательности гена из клона 64-31A в положении 1341 был определен остаток С, а в клоне 64-31, так же как и в штамме A/PR/8/34 — основание Т. Это говорит, вероятно, о наличии сосуществующих вариантов нуклеотидной последовательности некоторых генов в популяции вируса, поскольку замена основания не меняет смысла кодона. В отношении гена белка NP это отмечено также в работе [6].

Сравнение первичных структур NP-генов штаммов A/PR/8/34 и X/Ленинград/54/1 показало, что имеется 15 нуклеотидных замен (см. рис. 2), из которых только три приводят к консервативным заменам аминокислот Asp→Asn, Val→Leu, Thr→Ser, а остальные 12 являются заменами в 3-м положении кодона. При сравнении же двух штаммов вируса гриппа А — A/PR/8/34 и A/NT/60/68 [8] обнаруживается 125 нуклеотидных различий, приводящих к 30 аминокислотным заменам в их белках NP. Следует сделать вывод, что фрагмент 5 рекомбинантного генома вируса X/Ленинград/54/1 происходит от штамма A/PR/8/34. Точнее будет называть исходный штамм A/PR/8/34, взятый для получения рекомбинанта, штаммом A/PR/8/34, вариант СССР.

Сравнивая эти два лабораторных варианта (Cambridge и СССР) штамма A/PR/8/34, можно оценить степень полиморфности гена белка NP по нуклеотидной последовательности. Она составляет 4,8 различий на 1000 нуклеотидов. Аналогичная величина в случае вариантов Cambridge и Mount Sinai этого штамма для фрагмента 8, кодирующего неструктурные белки NS, составляет 5,6 (10 различий [9, 10]), для участка гена гемагглютинина, кодирующего HA<sub>1</sub>-субъединицу, — 6,0 (12 различий [11, 12]), а для гена М (фрагмент 7 генома) — 3 (6 различий [13, 14]).

К сожалению, этими примерами ограничивается набор секвенированных последовательностей для двух вариантов этого штамма. Неизвестно, как возник этот полиморфизм и каков вклад возможных мутаций при лабораторном пассировании вируса. Однако степень полиморфности очень близка для разных генов. Этот результат говорит об отсутствии специального механизма мутаций в генах поверхностных белков и подтверждает данные работ [15, 16].

### Экспериментальная часть

В работе использовался вирус гриппа типа А X/Ленинград/54/1 (H1N1-подтипа), являющийся рекомбинантом природного штамма А/Хабаровск/34457/77 и высокопродуктивного штамма A/PR/8/34, отобраным по антигенным свойствам природного штамма. Вакцинный штамм X/Ленинград/54/1 — отечественного производства.



Рис. 3. Предполагаемая вторичная структура участка РНК фрагмента 5 вируса гриппа изображена в форме (+)-цепи кДНК. Обозначено количество нуклеотидов в петлях и положение петель в последовательности фрагмента 5: «267» — 21—287, «59» — 309—367, «580» — 381—960, «4» — 968—971, «192» — 984—1175, «179» — 1197—1375, «26» — 1388—1413, «107» — 1443—1549. Стрелками отмечены места «перескока» полимеразы при копировании. Пунктирной стрелкой показано место вероятной комплементарной связи между петлями «580» и «4»

Подробное описание всех использованных в работе методов и реактивов приведено в нашей работе [4].

Авторы выражают признательность В. В. Старикову и Б. Г. Гриненкову за предоставление препаративных количеств вируса гриппа, Ю. С. Скоблову за синтез меченых нуклеотидов с высокой удельной активностью, А. П. Донченко за обработку результатов на ЭВМ, С. Х. Дегтяреву за предоставление высокоочищенных рестриктаз.

#### ЛИТЕРАТУРА

1. Беклемисhev А. Б., Блинов В. М., Василенко С. К., Головин С. Я., Гугоров В. В., Каргинов В. А., Мамаев Л. В., Микрюков Н. Н., Нетесов С. В., Петренко В. А., Петров Н. А., Сапдагчиев Л. С. *Биоорганическая химия*, 1984, т. 10, № 11, с. 1535—1543.
2. Беклемисhev А. Б., Блинов В. М., Василенко С. К., Головин С. Я., Каргинов В. А., Мамаев Л. В., Микрюков Н. Н., Нетесов С. В., Петренко В. А., Петров Н. А., Фролов И. В. *Биоорганическая химия*, 1985, т. 11, № 5, с. 628—635.
3. Maxam A. M., Gilbert W. *Proc. Nat. Acad. Sci. USA*, 1977, v. 74, № 2, p. 560—564.
4. Fields S., Winter G. *Gene*, 1981, v. 15, № 2/3, p. 207—214.
5. Robertson J. S. *Nucl. Acids Res.*, 1979, v. 6, № 12, p. 3745—3757.
6. Winter G., Fields S. *Virology*, 1981, v. 114, № 2, p. 423—428.
7. Hiti A. L., Nayak D. P. *J. Virol.*, 1982, v. 4, № 2, p. 730—734.
8. Huddleston J. A., Brownlee G. G. *Nucl. Acids Res.*, 1982, v. 10, № 3, p. 1029—1037.
9. Baez M., Taussing R., Zazra J. J., Young J. F., Palese P., Reisfield A., Skalka A. *Nucl. Acids Res.*, 1980, v. 8, № 23, p. 5845—5858.
10. Krystal M., Buonaquario D., Young J. F., Palese P. *J. Virol.*, 1983, v. 45, № 2, p. 547—554.
11. Caton A. J., Brownlee G. G., Yewdell J. W., Gerhard W. *Cell*, 1982, v. 31, № 2, p. 417—427.
12. Winter G., Fields S., Brownlee G. G. *Nature*, 1981, v. 297, № 5818, p. 72—75.
13. Allen H., McCauley J., Waterfield M., Gething M. J. *Virology*, 1980, v. 107, № 2, p. 548—551.
14. Winter G., Fields S. *Nucl. Acids Res.*, 1980, v. 8, № 9, p. 1965—1975.
15. Air G. M. *Proc. Nat. Acad. Sci. USA*, 1981, v. 78, № 12, p. 7639—7643.
16. Portner A., Webster R. G., Bean W. J. *Virology*, 1980, v. 104, № 1, p. 235—238.

Поступила в редакцию  
31.VII.1984  
После доработки  
28.XI.1984

#### SYNTHESIS, CLONING AND SEQUENCING OF A FULL-LENGTH DNACOPY OF NP GENE OF THE INFLUENZA VIRUS A

BEKLEMISHEV A. B., BLINOV V. M., VASSILENKO S. K., GOLOVIN S. Ya.,  
KARGINOV V. A., MAMAIEV L. V., MIKRIUKOV N. N., NETESOV S. V.,  
PETRENKO V. A., PETROV N. A., FROLOV I. V.

All-Union Research Institute of Molecular Biology,  
Koltzovo, Novosibirsk region

Complete nucleotide sequence of the cloned full-length DNA copy of the influenza virus A (H1N1) RNA segment 5 has been determined. The degree of nucleotide difference between two variants of A/PR/8/34 strain is estimated. The possible secondary structure of the segment 5 is deduced from the nucleotide sequence of some clones.