



ПИСЬМА РЕДАКТОРУ

УДК 577.113.4

МОЛЕКУЛЯРНОЕ КЛОНИРОВАНИЕ ГЕНОВ,
КОДИРУЮЩИХ ФАКТОРЫ НЕКРОЗА ОПУХОЛЕЙ.
ПОЛНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ НУКЛЕОТИДОВ ГЕНОМНОЙ
КОПИИ ФНО- α МЫШИ

Шахов А. П., Недоспасов С. А.

Институт молекулярной биологии Академии наук СССР, Москва

Факторы некроза опухолей (ФНО) являются белковыми иммуномедиаторами с широким спектром биологического действия, в частности с противоопухолевым эффектом *in vivo* и *in vitro* [1]. Недавно установлена [2–4] молекулярная структура белков ФНО человека, а также кодирующих их генов. В отличие от интерферонов ФНО имеют низкую видовую специфичность, так что ФНО человека, мыши и кролика активны и в гетерологичных системах [1–3]. Определена структура комплементарной ДНК (кДНК) гена ФНО мыши [5]. Кроме того, детально изучены хромосомные (геномные) копии генов ФНО человека [4, 6, 7]; при этом установлено, что уникальные гены, кодирующие ФНО и лимфотоксин, тандемно расположены в коротком сегменте генома длиной ~6 тыс. нуклеотидов (тыс. п.о.), нуклеотидная последовательность которого была определена [7]. Наконец, мы осуществили клонирование хромосомных копий генов ФНО и лимфотоксина мыши и установили их тандемное расположение на хромосоме 17 [8].

В настоящем сообщении мы приводим результаты определения последовательности нуклеотидов и экзон-интронной структуры гена ФНО мыши. Ген имеет длину ~2,5 тыс. п.о. и перекрывает два соседствующих *EcoRI*-фрагмента длиной 2,8 и 0,7 тыс. п.о. Эти фрагменты были переклонированы в одноцепочечный вектор, и их структура определена по методу Сэнгера [9] с использованием различных синтетических олигонуклеотидов в качестве затравок для синтеза (отмечены на рис. 1). Структура большинства олигонуклеотидов основывалась на известной последовательности кДНК гена ФНО мыши, другие были выбраны по ходу установления последовательности хромосомной копии гена. Кроме того, тем же методом были субклонированы и расшифрованы субфрагменты, образующиеся из *EcoRI*-фрагмента длиной 2,8 тыс. п.о. после гидролиза рестриктазами *HindIII* и *PvuII*.

Нами определена последовательность нуклеотидов фрагментов ДНК длиной ~3,4 тыс. п.о. Часть этой последовательности, содержащая собственно ген ФНО мыши, приведена на рис. 1. В начале последовательности выделен сигнальный участок инициации транскрипции, в конце — сигнал полиаденилирования мРНК (отмечены на рис. 1). Для того чтобы выявить участки, кодирующие белок ФНО, было проведено сопоставление полученной структуры с последовательностью кДНК [5], которое показало, что ген содержит 4 экзона и 3 интрона (рис. 2). Точные границы экзонов и интронов были установлены с учетом закономерностей организации генов высших организмов [10]. В результате нами выявлены участки, кодирующие предшественник зрелого белка и сам белок ФНО (рис. 1). Установленная организация мышиного гена полностью соответствует

1 CCGGCCCTCT TCCSAGGGCT ATAAAGGGGC CGTCTGCACA GCSAGCCAGC ACAAAGCTCCC TCAGCGGAGGA
 71 CAGCAAGGGA CTAGCCAGGA GGGAGAACAG AAACCTCCAGA ACATCTTGGG AATAGCTCCC AGAAAAGCAA
 141 GCAGCAACC AGCAGGTTT TGTCCSTTC ACTCACTGGC CCAAGGGGCC .ACATCTCCST CCAAGAAAAGA
 211 met ser thr glu ser met ile ary asp val glu leu ala glu glu ala leu pro
 CACCATG ACC ACA GAA AGC ATG ATC CGC GAC GTG GAA CTG GCA GAA GAG GCA CTC CCC
 269 gln lys met gly gly phe gln asn ser arg arg cis leu cis leu ser leu phe ser
 CAA AAG ATG GCG GGC TTC CAG AAC TCC AGG CGG TGC CTA TGT CTC AGC CTC TTC TCA
 phe leu leu val ala gly ala thr thr leu phe cis leu leu asn phe gly val ile
 TTC CTG CTT GTG GCA GGG GCC ACC ACG CTC TTC TGT CTA CTG AAC TTC GGG GTG ATC
 326 gly pro gln arg asp glu
 GGT CCC CAA AGG GAT GAG GTAGTGTC TCGGCACCCT TATCTCTGGCT CACAAGCAAG ACGGGTTAGG
 383 AGGGCAAGAA GACAGTGTGA GGGAAAGAA TGGGCTAATG GGCAGGGCAA GGTGGAGGAG AGTGTGGAGG
 449 GGACACAGTC AGGACCTCGG ACCCATGCGT CCAGCTGACT AAACATCCTT CTCTCGGATGC ACAGAGAGAT
 589 GAATGAACGA ACAAATGTGT TCACACGTGG AGAGATCTGG AAAGATGTGG CCAGGGGAAG AGGGGATAAG
 659 CAAGAGATTA AACTCAGAGA CAGAATGAG AGAGGCATGA GAGATAAGGA GGAAGATGAA GGGGAGATAA
 729 CGGGAGATCA AGCACAGAGG GCACCGGCAGA AAGAAGCCGT GGGTTGGACA GATGAATGAA TGAAGAAGAA
 799 AACACAAAGT GGGGGTGGG TGGGGCAAAG AGGAACTGTA AGCGGGGCAA TCAGCCGGGA GCTTCTCCTT
 869 TGGGGTGACT CTGTCTTAAC TAACCTCCTT TTCCTACACA GAAG TTC CCA AAT GGC CTC CCT CTC
 934 file ser ser met ala gln thr leu thr Leu A
 ATC AGT TCT ATG GCC CAG ACC CTC ACA CTC AGT AAGTGTTCCTT ACACCTCTCT CTTAATTTAA
 997 GATGGAGAAG GGCAGTTAGG CATGGGATGA GATGGGGTGG GGGGAAAACCT TAAAGCTTTG GTTTGGGAGG
 1067 AAAGGGTCT AAGTGCATAG ATGCTTGTCTG GAAAGCCATA AAGGCTCATC CTGSCCTTG TCTCTTCCCC
 rg Ser Ser Ser Gln Ash Ser Ser Asp Lys Pro Val Ala His Val Val A
 1137 TCCAGGA TCA TCT TCT CAA AAT TCG AGT GAC AAG CCT GTA GCC CAC CTC GTA GGTAAGAT
 1197 TTCTTTACAT GTGCTTGTAG AATGAAGGGG CATGATTTTG GGGGGGGGGT TGAGGGGTGT CGAGCCAGGC
 1267 TGAGAAAAGA CAGAGCTCTT AGAGACAGCA CGTGAGAGTC AGAGCAGTGA CTCAAAAGCA AGGCATCAGG
 1337 GGGCCACCCG GGACCTCATA GCCAATGGGA TGTGGAAGA CAGAGGGTGC AGGAAACCCGA AGTGAAGTGT

Рис. 1. Полная последовательность нуклеотидов хромосомной копии гена ФНО- α мыши. Показано 2709 нуклеотидов по цепи, кодирующей мРНК. Сигнальные последовательности для инициации транскрипции и полиденитрирования мРНК подчеркнуты сплошной чертой. Последовательность аминокислот показана над экзонами (для зрелой части белка, начало и конец которого отмечены звездочками, аминокислоты написаны с заглавной буквы). Последовательностям олигонуклеотидам, которые использовались для секвенирования, подчеркнуты штриховой линией (при совпадении с кодирующей цепью — над последовательностью, при совпадении с комплементарной цепью — под последовательностью). Стрелки на штриховых линиях указывают на направлении секвенирования. Кроме того, штриховой линией со стрелками в обе стороны подчеркнуты рестрикционные сайты, по которым проводилось субклонирование фрагментов ДНК. В этих случаях определение последовательности осуществили от универсальной затравки [11], начиная с выделенного рестриктного сайта. Строчные буквы в последовательности олигонуклеотидов (4-й экзон) указывают на точечные замены, соответствующие различиям в структуре генов мыши и человека [7, 8]

1407 CCGTAGCTGC TGAGGCTCAG GATGTGGAGT GTCAACTAAG AGGGTGACAC TGACTCAATC CTCCCCCCCC
1477 CCCTCAG CA AAC CAC CAA GTG GAG GAC CAG CTG GAG TGG CTG AGC CAG CGC GCC AAC
1534 Ala Leu Leu Ala Asn Gly Met Asp Leu Lys Asp Asn Gln Leu Val Val Pro Ala Asp
GCC CTC CTG GCC AAC GGC ATG GAT CTC AAA GAC AAC CAA CTA GTG GTG CCA GCC GAT
1591 Gly Leu Tyr Leu Val Tyr Ser Gln Val Leu Phe Lys Gly Gln Gly Cis Pro Asp Tyr
GGG TTG TAC CTT GTC TAC TCC CAG GTT CTC TTC AAG GGA CAA GGC TGC CCC GAC TAC
1648 Val Leu Leu Thr His Thr Val Ser Arg Phe Ala Ile Ser Tyr Gln Glu Lys Val Asn
GTG CTC CTC ACC CAC ACC GTC AGC CGA TTT GCT ATC TCA TAC CAG GAG AAA GTC AAC
1705 Leu Leu Ser Ala Val Lys Ser Pro Cis Pro Lys Asp Thr Pro Glu Gly Ala Glu Leu
CTC CTC TCT GCC GTC AAG AGC CCC TGC CCC AAG GAC ACC CCT GAG GGG GCT GAG CTC
1762 Lys Pro Trp Tyr Glu Pro Ile Tyr Leu Gly Val Phe Gln Leu Glu Lys Gly Asp
AAA CCC TGG TAT GAG CCC ATA TAC CTG GGA GTC TTC CAG CTG GAG AAG GGG GAC
1819 Gln Leu Ser Ala Glu Val Asn Leu Pro Lys Tyr Leu Asp Phe Ala Glu Ser Gly Gln
CAA CTC ACC GCT GAG GTC AAT CTG CCC AAG TAC TTA GAC TTT GCG GAG TCC CGG CAG
1876 Val Tyr Phe Gly Val Ile Ala Leu ***
GTC TAC TTT GGA GTC ATT GCT CTG TGA AGGAAAT GGGTGTTCAT CCATTCTCTA CCCAGCCCCC
1940 ACTCTGACCC CTTTACTCTG ACCCCCTTTAT TGTCTACTCC TCAGAGCCCC CAGTCTGTAT CCTTCTAACT
2010 TAGAAAGGGG ATTATGGCTC AGGTCCAAAC TCTGTGCTCA GAGCTTTCAA CACTACTCA GAAACACAAG
2080 ATGCTGGGAC AGTGACCTGG ACTGTGGGCC TCTCATGCAC CACCATCAAG GACTCAAAG GCGTTTCCGA
2150 ATTCACTGGA GTCTCGAATG TCCATTCCCTG AGTCTGCAA AGGCAGAGTG GTCAGGTTGC CTCTGTCTCA
2220 GAATGAGGCT GGATAAGATC TCAGGCCCTC CTACCTTCAG ACCTTTCAG ATTCTTCCCT GAGGTGCAAT
2290 GCCAAGCCTT CCTCACAGAG CCAGCCCCCC TCAATTTATA TTTGCACCTA TTTATTTATTA TTTATTTATT
2360 ATTTATTTAT TTGCTTATGA ATGTATTTAC TTGGAAGGCC GGGGTGTCTT GGAGGACCCA GTGTGGGAAG
2430 CTGTCTTCAG ACAGACATGT TTTCTGTGAA AACGGACTG AGCTGTCCCC ACCTGGCCCTC TCTACCTTGT
2500 TGCCTCCTCT TTTGCTTAAT GTTTAAAACA AATAATTTAT CTAACCCAAAT TGTCTTAATA ACCCTGATTT
2570 GGTGACCAGG CTGTGCTAC ATCACTGACC CTCTTGCCCC ACGGGAGCCG TGACTGTAAAT CGCCTACGGT
2640 CAATTGAGAA AAATAAAGAT CGCTTGAAA AGAAATGTGA TTTCTGTCTT GGGATGAAGT CTGCATCCAT

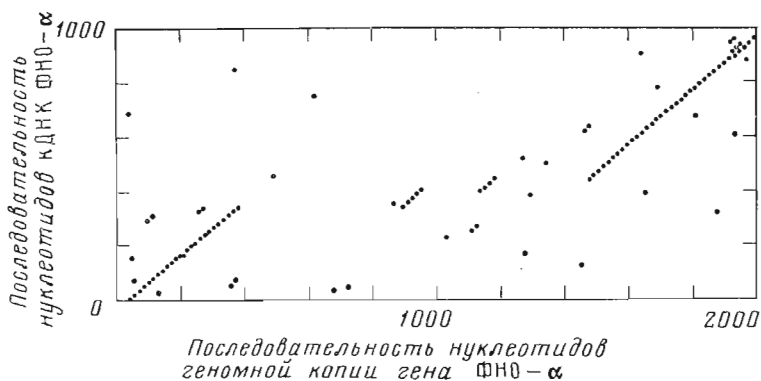


Рис. 2. Экзон-интронная структура гена ФНО- α мыши по результатам сопоставления последовательности нуклеотидов кДНК [5] и хромосомной копии гена (см. рис. 1)

структуре гена ФНО- α человека [4, 6, 7], при этом первые, третьи и четвертые экзоны генов обоих видов гомологичны приблизительно на 80%. Наоборот, последовательности интронов, в особенности первых, наиболее длинных, сильно дивергированы и содержат большое количество несовершенных повторов. Этим же свойством обладает и первый интрон гена лимфоциткиназы (ФНО- β) человека [6, 7]. Наконец, нами выявлены гомологичные последовательности в некодирующих частях генов, которые могут играть роль в регуляции их экспрессии.

Матрица точечных гомологий между кДНК (нуклеотиды 1–1000 [5]) и хромосомной копией гена ФНО- α (нуклеотиды 1–2000, рис. 1) построена на ПЭКВМ «Искра-222» с помощью программы Revelon, разработанной Е. И. Головановым (Институт молекулярной генетики АН СССР).

Впервые установленная нами полная структура гена фактора некроза опухолей мыши хорошо согласуется с определенной ранее последовательностью кДНК [5], хотя нами обнаружен ряд отличий (см. рис. 1). В опубликованной ранее [5] последовательности нуклеотидов кДНК имеются следующие отличия от структуры экзонов гена ФНО- α , приведенной на рисунке: 116 — Т, 2151 — Т, 2271 — Т, 2292 — С, 2293 — А, 2389 — С (номера нуклеотидов соответствуют приведенным на рисунке). Следует иметь в виду, что кДНК была получена в результате синтеза с помощью ревертазы на матрицах мРНК *in vitro*, а этот процесс нередко сопровождается ошибками копирования. Поэтому в случае разночтений предпочтение следует отдавать последовательности нуклеотидов хромосомной копии гена.

Авторы благодарны Э. Кавашиме, Б. К. Чернову, В. Н. Добрынину, С. А. Филишову, Р. Л. Турецкой, В. А. Метт, М. М. Азизову за содействие в работе, а также Г. П. Георгиеву за поддержку этого исследования.

ЛИТЕРАТУРА

1. Old L. J. // *Science*. 1985. V. 230. № 4726. P. 630–632.
2. Aggarwal B. B., Kohr W. J., Moffat B., Hass P. E., Spencer S. A., Henzel W. J., Bringman T. S., Nedwin G. E., Goeddel D. V., Harkins R. N. // *J. Biol. Chem.* 1985. V. 260. № 4. P. 2345–2354.
3. Pennica D., Nedwin G. E., Hayflick J. S., Seeburg P. H., Derynck R., Palladino M. A., Kohr W. J., Aggarwal B. B., Goeddel D. V. // *Nature*. 1984. V. 312. № 5996. P. 724–729.
4. Shirai T., Yamaguchi H., Ito H., Todd C. V., Wallace R. B. // *Nature*. 1985. V. 313. № 6005. P. 803–806.
5. Fransen L., Muller P., Marmenout A., Tavernier J., Van der Heyden J., Kawashima E., Cholett A., Tizard R., Van Heuverswin H., Van Vliet A., Ruyschaert M. R., Fiers W. // *Nucl. Acids Res.* 1985. V. 13. № 12. P. 4417–4429.
6. Nedwin G. E., Naylor S. L., Sakaguchi A. J., Smith D., Jarrett-Nedwin J., Pennica D., Goeddel D. V., Gray P. W. // *Nucl. Acids Res.* 1985. V. 13. № 17. P. 6361–6373.
7. Nedospasov S. A., Shakhov A. N., Turetskaya R. L., Mett V. A., Azizov M. M., Georgiev G. P., Korobko V. G., Dobrynin V. N., Filippov S. A., Bystrov N. S., Boldyre-

- va E. F., Chuvpilo S. A., Chumakov A. M., Shingarova L. N., Ovchinnikov Yu. A. // Cold Spring Harbor Symp. Quant. Biol. 1987. V. 51.
8. Nedospasov S. A., Hirt B., Shakhov A. N., Dobrynin V. N., Kawashima E., Accolla R. S., Jongeneel C. V. // Nucl. Acids Res. 1986. V. 14. № 19. P. 7713-7728.
 9. Sanger F., Nicklen S., Coulson A. R. // Proc. Nat. Acad. Sci. USA. 1977. V. 74. № 12. P. 5463-5467.
 10. Breathnach R., Chambon P. // Ann. Rev. Biochem. 1981. V. 50. P. 349-383.
 11. Messing J., Vieira J. // Gene. 1982. V. 19. № 3. P. 269-276.

Поступило в редакцию
18.X.1986

**MOLECULAR CLONING OF THE GENES CODING FOR TUMOR NECROSIS
FACTORS: COMPLETE NUCLEOTIDE SEQUENCE OF THE GENOMIC COPY
OF TNF-ALPHA IN MICE**

SHAKHOV A. N., NEDOSPASOV S. A.

Institute of Molecular Biology, Academy of Sciences of the USSR, Moscow

Using Sanger's technique and synthetic oligonucleotides derived from the known structure of cDNA, complete nucleotide sequence of the genomic copy of the mouse TNF-alpha gene has been determined. The length of the gene from transcription initiation site down to polyadenylation signal is about 2620 bp. Comparison with the previously described cDNA sequence revealed existence of 4 exons and 3 introns in the positions homologous to those of human TNF gene. The fourth exon codes for 88% amino acids of mature TNF-alpha.