



УДК 577.113.6:578.245

ОПРЕДЕЛЕНИЕ ПЕРВИЧНОЙ СТРУКТУРЫ МУТАНТНЫХ ГЕНОВ ЛЕЙКОЦИТАРНОГО ИНТЕРФЕРОНА $\alpha 2$ ЧЕЛОВЕКА

*Петренко В. А., Татъков С. И., Ерошкин А. М.,
Болдырев А. Н., Поздняков П. И., Сиволобова Г. Ф.*

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

Определена полная первичная структура 10 мутантных генов лейкоцитарного интерферона $\alpha 2$ человека (IFN $\alpha 2$) при использовании четырех олигонуклеотидов-праймеров, содержащих этильные заместители по фосфатной группе. Описан подход к выбору структуры праймеров, основанный на построении профиля локального сходства гена IFN и векторной ДНК фага M13mp7.

Одним из этапов исследования структурно-функциональной организации IFN посредством локализованного мутагенеза является установление нуклеотидной последовательности мутантных генов [1]. Для этой цели обычно используют метод Сэнгера [2], который позволяет с одного праймера в одном геле достоверно установить структуру фрагмента ДНК длиной около 200 нуклеотидов. Исходя из длины интерферона $\alpha 2$ (166 аминокислот), для прочтения на одном геле полной последовательности его мутантных генов требуется несколько праймеров.

Олигонуклеотиды-праймеры для секвенирования должны удовлетворять следующим требованиям: 1) не связываться с участками как самого гена, так и векторной молекулы вне заданного района, т. е. обладать специфичностью связывания с матрицей; 2) содержать минимальное количество оснований, комплементарных тем, которые подвергаются моди-

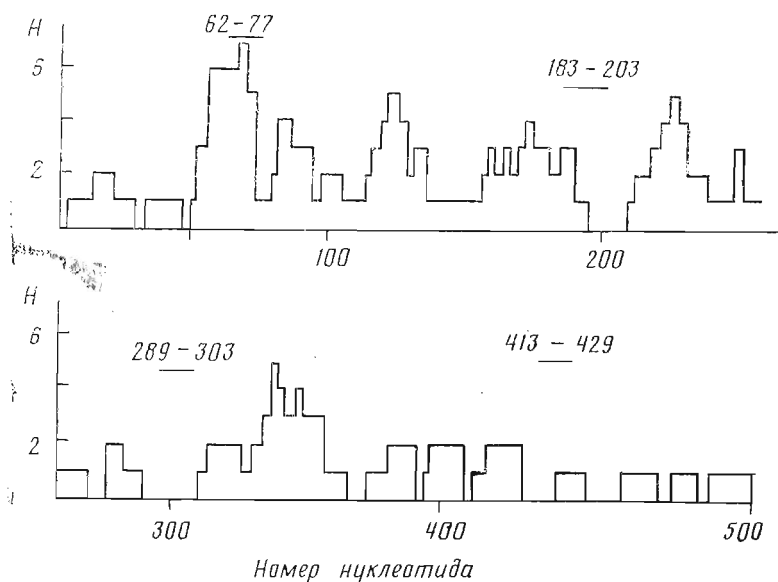


Рис. 1. Анализ гомологии гена IFN и ДНК M13mp7. H — частота встречаемости нуклеотидов гена IFN в совпадающих фрагментах гена и ДНК M13mp7. По оси абсцисс приведена нумерация нуклеотидов с 5'-конца не кодирующей цепи гена IFN; вынесены участки, выбранные для связывания с олигонуклеотидами-праймерами

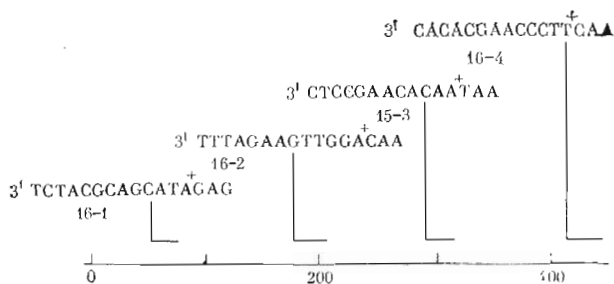


Рис. 2. Схема определения нуклеотидной последовательности мутантных генов IFN по методу Сэнгера. Горизонтальными линиями указано положение праймеров на гене IFN (нумерация нуклеотидов с 5'-конца некодирующей цепи). Крестиками обозначены фосфатные группы, содержащие этильные заместители

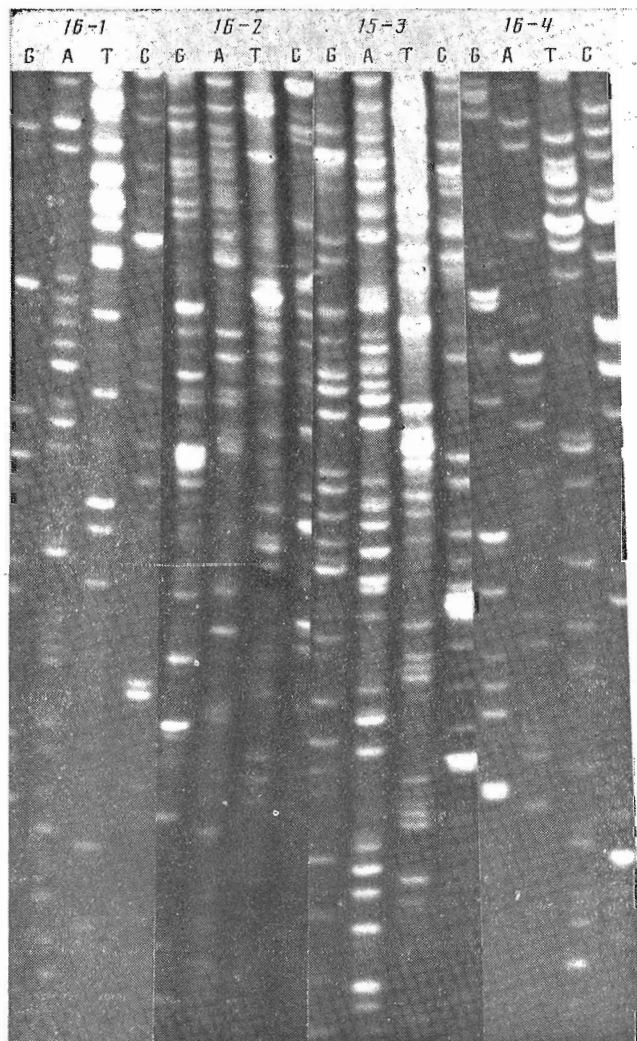


Рис. 3. Определение нуклеотидной последовательности мутантного гена IFN с помощью четырех праймеров по методу Сэнгера. Прочитаны районы гена: 158 (праймер 16-1), 75-174 (16-2), 162-273 (15-3), 328-407 (16-4)

фикации в составе ДНК (в данном случае G, так как в качестве мутагена использован бисульфит натрия).

Для удовлетворения этих требований с помощью компьютерного анализа были определены районы IFN, имеющие мало сходных участков

	10				20				30				40				50				60				70			
IFN- α 2	ATG	TGT	GAT	CTG	CCG	CAG	ACT	CAC	TCT	CTG	GGT	TCT	CGT	CGT	ACT	CTG	ATG	CTG	CTG	GCT	CAG	ATG	CGT	CGT				
MIF-228				3					8			5				6			5									
MIF-229																							7					
MIF-231				3		0			8		5				4			5										
MIF-232				3		0			6	8					4													
MIF-233				3		0					5																	
MIF-234				34					8						4													
MIF-236				34		0			8		5								9									
MIF-237									6	8		5			4													
MIF-238									6		5				4				9									
MIF-239											5				4								3					

	80				90				100				110				120				130				140			
IFN- α 2	ATC	TCT	CTT	TTC	TGC	CTG	AAA	GAC	CGT	CAT	GAC	TTC	GGT	TTC	CCG	CAG	GAA	GAG	TTC	GGT	AAC	CAG	TTC	AAA	GAC	TTC		
MIF-228					6																							
MIF-229		7	9		6				9						7	8								8				
MIF-231	5	7	9												7	8	1									4		
MIF-232	5	7	9												8	1								8				
MIF-233	5				6																							
MIF-234		7							9	0								1										
MIF-236					6	0	1																			4		
MIF-237					6				9						7	8												
MIF-238					4						3	8														8		
MIF-239											0				7													

	150				160				170				180				190				200				210			
IFN- α 2	CAG	AAA	GCT	GAA	ACT	ATC	CCT	GTT	CTG	CAT	GAA	ATG	ATC	CAG	CAA	ATC	TTC	AAC	CTG	TTC	TCT	ACT	AAA	GAC	TTC	AAA	GAC	
MIF-228						4									3													
MIF-229					8		3											8			4		9				6	
MIF-231					8		2											8					9				5	
MIF-232					8		3											8					9				5	
MIF-233					8		2	3						3			2	5					6				6	
MIF-234	5			2																			6				9	
MIF-236						2	4				9								9		4	6					5	
MIF-237						8	2	3			9			3					9		4	6					5	
MIF-238							4				2																6	
MIF-239						8												2	5								6	

	220				230				240				250				260				270				280			
IFN- α 2	TCT	TCT	GCT	GCT	TGG	GAC	GAA	ACC	CTG	CTG	GAC	AAA	TTC	TAC	ACT	GAA	CTG	TAC	CAG	CAA	CTC	AAC	GAC	CTC	TTC	AAA	GAC	
MIF-228	8	1													8								2					
MIF-229	8					4									8												5	
MIF-231				4		4		9	1																			
MIF-232				4	7				1						5												5	
MIF-233								9	1						8													
MIF-234				1				9			9																	
MIF-236										4							5				0	1		4		2	8	
MIF-237									0					5		0							7				8	
MIF-238									9						0						0	1						
MIF-239									9	1					8	0					0							

	290				300				310				320				330				340				350				360			
IFN- α 2	GAG	GCT	TGT	GTT	ATT	CAG	GGC	GTT	GGT	GTT	ACC	GAA	ACT	CCG	CTG	ATG	AAA	GAG	GAT	AGC	ATC	CTG	GCT	GTT	TTC	AAA	GAC					
MIF-228																																
MIF-229																	9															
MIF-231																																
MIF-232																																
MIF-233																																
MIF-234																																
MIF-236																6																
MIF-237																																
MIF-238																																
MIF-239																																

	370				380				390				400				410				420				430						
IFN- α 2	CGC	AAG	TAT	TTC	CAG	CGT	ATC	ACT	CTT	TAC	CTG	AAG	GAA	AAG	AAG	TAT	AGC	CCG	TGT	GCT	TCG	GAA	GTT	GTT	TTC	AAA	GAC				
MIF-228	3					6																									
MIF-229						2					0																				
MIF-231						6																									
MIF-232	3																														
MIF-233	3					6	1				1																				
MIF-234								1																							
MIF-236						2					5																				
MIF-237						2																									
MIF-238																															
MIF-239						2					3																				

	440				450				460				470				480				490				500						
IFN- α 2	CGT	GCT	GAG	ATC	ATG	CGT	TCC	TTC	TCC	CTG	TCT	ACT	AAc	CTG	CAG	GAA	TCT	CTG	CGT	TCT	AAG	GAA	TAG	TTC	AAA	GAC					
MIF-228								89				7																			
MIF-229																	1														
MIF-231		7										7																			
MIF-232																															
MIF-233																															
MIF-234								8			6																				
MIF-236								2			6	9	0		4																
MIF-237	3							23			8																				
MIF-238								2			89																				
MIF-239	3	7									6		0																		

Рис. 4. Нуклеотидная последовательность мутантных генов интерферона. Приведена последовательность исходного гена IFN-2; для мутантных генов MIF показаны лишь положения замен С на Т (например, цифра 4 между 160 и 170 для MIF-228 означает, что модифицирован С-164)

в векторной молекуле ДНК М13mp7. Профиль локального сходства (H) гена IFN и ДНК М13mp7 (рис. 1) показывает, сколько сходных участков содержат разные районы гена IFN и ДНК М13mp7. Участки с низкими значениями профиля локального сходства были использованы при выборе структуры праймеров для секвенирования.

Первичную структуру 3'-концевой области гена определяли, используя универсальный 17-звенный праймер [3], а структуру средней части — с помощью трех синтетических олигонуклеотидов, комплементарных районам 187—203, 289—303, 413—429. Для секвенирования начала гена IFN была дополнительно выбрана последовательность 62—77, включающая участок с минимальным значением профиля локального сходства ($H = 1$) (рис. 1).

Определяющее значение в методе Сэнгера имеет качество фрагмента: Кленова ДНК-полимеразы I *E. coli*. Эндо- и 5'-экзонуклеазные примеси в препарате искажают картину секвенирования. Для повышения воспроизводимости метода мы использовали фосфотриэфирные аналоги олигонуклеотидов (рис. 2): ранее было показано, что такие производные устойчивы к действию экзонуклеаз и могут применяться в качестве затравок для ферментативного синтеза ДНК *in vitro* [4].

Как видно из рис. 3, все праймеры специфично связываются с расчетными районами, что определило возможность анализа нуклеотидной последовательности 10 мутантных генов IFN (рис. 4).

Авторы благодарят А. Е. Никулина за предоставление программы поиска вырожденных повторов.

Экспериментальная часть

Выбор нуклеотидной последовательности праймеров. С помощью программы поиска вырожденных повторов в исходном гене и векторной ДНК М13mp7 находили совпадающие фрагменты длиной 9 и более нуклеотидов, не менее 13 нуклеотидов с двумя несовпадениями и 17 и более нуклеотидов с четырьмя несовпадениями (например, фрагменты (9,0), (13,2) и (17,4)). Из списка более коротких фрагментов исключали входившие в список более длинных фрагментов. Для каждого нуклеотида из гена IFN подсчитывали, сколько раз он входил в состав совпадающих фрагментов (9,0), (13,2), (17,4) в гене IFN и векторной ДНК. Отложив по оси ординат число этих случаев, а по оси абсцисс — положение нуклеотида в гене, построили диаграмму локального совпадения (H) для всего гена (рис. 1). Районы гена, которые не имеют гомологичных участков в молекуле вектора (на диаграмме им соответствуют нулевые значения профиля H), перспективны для выбора участков связывания с олигомерами-праймерами. Окончательный выбор участков связывания проводили с учетом требования о минимальном содержании цитозинов в ДНК.

Синтез олигодезоксирибонуклеотидов-праймеров. Олигонуклеотиды, содержащие этильные заместители, 5'GAG⁺ATACGACGCATCT, AAC⁺AGGTTGAAGATTT, AAC⁺TTCCSAAGCASC, синтезировали по схеме $[2 + (2^* + 2)] + \{(2 + 2) + [2 + (2 + 2)]\}$; 5'AAT⁺AACASAAAGCCTC — по схеме $[2 + (2^* + 2)] + \{(2 + 2) + [2 + (1 + 2)]\}$ (положение фосфотриэфирного звена в олигонуклеотиде обозначено крестиком). Моно- и динуклеотидные защищенные блоки получали по методам [4, 5], динуклеотидные блоки с этильными заместителями по межнуклеотидной фосфатной группе (на схемах помечены звездочкой) — по методам [4, 6]. Сборку целевых олигонуклеотидов, полное деблокирование и выделение проводили согласно [4]. Структуру синтезированных олигонуклеотидов подтверждали методом Максама — Гилберта [7], а положение фосфотриэфирных узлов — анализом продуктов щелочного расщепления [4].

Секвенирование ДНК. Одноцепочечную ДНК выделяли по методу [8] и секвенировали по методу Сэнгера [2] с использованием фосфотриэфирных праймеров. Дезокси- и дезоксирибонуклеозидтрифосфаты синтезированы в лаборатории, фрагмент Кленова ДНК-полимеразы I *E. coli* — производства НИКТИ БАВ (Бердск).

ЛИТЕРАТУРА

1. Петренко В. А., Татиков С. И., Сиволобова Г. Ф., Болдырев А. Н., Колокольцов А. А., Ершкин А. М., Куличков В. А. // Биоорганич. химия. 1987. Т. 13. № 2. С. 259—262.
2. Sanger F., Coulson A. R., Barrel B. G., Smith A. G. H., Roe B. A. // J. Mol. Biol. 1980. V. 143. № 2. P. 161—178.
3. Duckworth M. L., Gail M., Goelet P., Hong G. F., Singh M., Titmas R. C // Nucl. Acids Res. 1981. V. 9. № 7. P. 1691—1706.

4. Петренко В. А., Поздняков П. И., Киприянов С. М., Болдырев А. Н., Семенова Л. Н., Сиволобова Г. Ф. // Биоорг. химия. 1986. Т. 12. № 8. С. 1088—1100.
5. Stawinski J., Hozumi T., Narang S. A. // Can. J. Chem. 1976. V. 54. № 3. P. 670—672.
6. Петренко В. А., Поздняков П. И., Сиволобова Г. Ф., Шубина Т. Н. // Биоорг. химия. 1980. Т. 6. № 3. С. 431—435.
7. Маннатис Т., Фрич А., Сэмбрук Дж. Методы генетической инженерии. Молекулярное клонирование. М.: Мир, 1984. С. 416—420.
8. Петренко В. А., Сиволобова Г. Ф., Семенова Л. Н., Болдырев А. Н., Каргинов В. А., Гуторов В. В. // Молекул. генетика, микробиол. и вирусол. 1985. № 8. С. 38—44.

Поступила в редакцию
21.VIII.1987
После доработки
23.XI.1987

SEQUENCE OF MUTANT GENES OF HUMAN LEUKOCYTE INTERFERON

PETRENKO V. A., TATKOV S. I., EROSHKIN A. M.,
BOLDYREV A. N., POZDNYAKOV P. I., SIVOLOBOVA G. F.

*All-Union Institute of Molecular Biology,
Kol'tsovo, Novosibirsk Region*

Nucleotide sequences of 10 mutant genes of human leukocyte interferon $\alpha 2$ (IFN), with the use of 4 oligonucleotide primers containing ethyl substituents at phosphate groups were determined. To design primer sequences, an approach based on the local similarity profile of the IFN gene and M13mp7 vector DNA is described.