

УДК 577.2:631:581.115:542

А. Ю. Губский¹, асп., **О. В. Жук**², док. биол. наук, проф.¹ Одесский национальный университет им. И. И. Мечникова,
кафедра физиологии человека и животных,
ул. Дворянская, 2, Одесса, 65026, Украина² Опольский университет,
кафедра молекулярной и экспериментальной биологии,
ул. Олеска, 48, Ополье, 45052, Польша

ПОЛИМОРФИЗМ СТРУКТУРНЫХ ЭЛЕМЕНТОВ СИГНАЛЬНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ РЕКОМБИНАЦИИ ГЕНОВ ИММУНОГЛОБУЛИНОВ И Т-КЛЕТОЧНЫХ РЕЦЕПТОРОВ ЧЕЛОВЕКА

Установлена структура 547 сигнальных последовательностей рекомбинации трех генов иммуноглобулинов и четырех генов Т-клеточных рецепторов человека. Нуклеотидный состав 50 % гептамеров и 89 % наномеров не соответствует каноническим последовательностям САСАGТG и АСАААААСС. Только 84 % RSS содержат 12- или 23- нуклеотидные спейсеры. Структура последовательностей рекомбинации иммуноглобулинов менее полиморфна, чем Т-клеточных рецепторов.

Ключевые слова: гептамеры, наномеры, RSS.

В генах иммуноглобулинов (IG) и Т-клеточных рецепторов (TCR) сигнальные последовательности рекомбинации (RSS) располагаются на границах интронов, примыкающих к V-, D-, J- сегментам. Они являются субстратом белков RAG1 и RAG2 в реакции эндонуклеазного расщепления ДНК. В результате этого образуются функциональные IG и TCR гены [1]. RSS состоят из трех типов элементов: гептамеров, спейсеров и наномеров. Структура гептамеров, в целом, соответствует последовательности САСАGТG, а наномеров АСАААААСС. Первые три позиции гептамера (САС), пятая, шестая и седьмая позиции наномера (ААА) являются наиболее консервативными [2, 3]. Размер спейсерного участка, как правило, постоянен и соответствует 12- или 23- нуклеотидам (12RSS и 23RSS). Однако, у некоторых RSS он может изменяться на один нуклеотид, образуя 11-, 13-, 22- и 24- нуклеотидные производные (11RSS, 13RSS, 22RSS, 24RSS). Спейсер отделяет гептамер от наномера и проявляет слабую консервативность нуклеотидного состава [4]. В настоящее время, как у человека, так и мыши, в IG и TCR генах установлена структура многих RSS. Информацию о них можно найти в литературе и электронных биологических базах данных Entrez, IMGT и т. д.

Нами была поставлена задача: провести полный сравнительный анализ нуклеотидного состава гептамеров и наномеров всех сигналь-

ных последовательностей рекомбинации IG, TCR генов человека, дать количественную оценку степени полиморфизма их структур.

Материалы и методы

В работе исследована первичная структура ДНК генов тяжелой, κ -, λ - цепей иммуноглобулинов (IGH, IGK, IGL) и α -, δ -, γ -, β - цепей Т-клеточных рецепторов (TCRA, TCRD, TCRG, TCRB) человека. Последовательность ДНК указанных генов с координатами их V-, D-, J- сегментов взята из электронной биологической базы данных Entrez в виде 7 файлов текстового формата (<http://www.ncbi.nlm.nih.gov/Entrez/>). В указанном источнике последовательности RSS представлены для V-сегментов генов TCRA и TCRD, а так же для V-, D-, J- сегментов гена TCRB.

Альтернативный поиск сигнальных последовательностей рекомбинации провели в 583 интронных участках, размером 50 нуклеотидов, примыкающих к 551 V-, D-, J- сегментам. Если в анализируемой области гена, на расстоянии друг от друга равном 12 или 23 парам оснований, обнаруживались семи- и девятинуклеотидные последовательности, которые соответствовали структурам CACAGTG и ACAAAAACC как минимум по трем или четырем основаниям, то найденные элементы мы рассматривали как гептамеры и наномеры RSS, разделенные 12- или 23- нуклеотидными спейсерами. Делая поправку на изменчивость длины спейсерной области, мы обнаруживали наномеры, в большей степени соответствующие структуре ACAAAAACC, чем при стандартных размерах спейсеров. Кроме 12RSS и 23RSS, это позволило идентифицировать 11RSS, 13RSS, 22RSS, 24RSS и другие. В ходе исследования не рассматривались RSS, состоящие только из одного гептамера или наномера.

Консервативность нуклеотидных позиций исследуемых элементов, определяли на основании процента встречаемости одного из четырех оснований в определенном участке последовательности от общего количества анализируемых элементов.

Результаты исследования

Проанализировав у генов TCRA и TCRB последовательности известных RSS, мы обнаружили, что в структуре некоторых из них неверно указан нуклеотидный состав гептамеров и наномеров. Так, у сегментов TRAV26-1, TRAV26-2, TRBV11-1, TRBV13 генов TCRA и TCRB вместо наномеров GCAATATCT, GCAATATCT, ACAAAAACCT, ACCCAAACC указаны последовательности CAATATCTC, CAATATCTC, CACAAAACCT, TACCCAAAC соответственно. Как мы определили, это было связано с неверно установленным размером спейсерной области. Например, правильнее считать, что длина спейсера RSS сегмента TRBV11-1 равняется 24 основаниям, а не 23, поскольку следующий за ним девятинуклеотидный фрагмент ДНК в большей степени соответствует последовательности ACAAAAACC, так как отличается от нее всего по двум основаниям. В случае 23- нуклеотидного спейсера отличие составило

бы четыре основания, ставя под сомнение точность нуклеотидного состава определяемого элемента. В структуре RSS, описанных в литературе, мы обнаружили ошибки такого же характера [5]. У сегмента IGHV3-54 гена IGH спейсер RSS представлен 25, а не 22 нуклеотидами, ибо именно последовательности CACCAGG и ACACAGAAT больше подходят на роль гептамера и наномера.

Учитывая эти факты, мы отказались от использования электронных баз данных при исследовании структурного полиморфизма сигнальных последовательностей рекомбинации. Проанализировав имеющиеся данные о первичной структуре ДНК трех IG и четырех TCR генов человека, мы самостоятельно установили последовательности гептамеров, спейсеров и наномеров RSS (см. материалы и методы).

У 515 V-, D-, J- сегментов мы обнаружили 547 RSS, содержащих как гептамеры, так и наномеры. У 36 сегментов мотивы рекомбинации отсутствовали или состояли из одного элемента. Мы установили, что половина всех гептамеров и большая часть наномеров (50 % и 89 %) не соответствуют каноническим структурам CACAGTG и CAAAAACC (табл. 1).

Таблица 1

Количественная оценка структурного полиморфизма гептамеров и наномеров сигнальных последовательностей рекомбинации IG и TCR генов человека

Элементы RSS	Общ. Кол-во	Моно-морфные	Поли-морфные	Количество нуклеотидных замен в полиморфных гептамерах и наномерах				
				1	2	3	4	5
Гептамеры (IG, TCR)	547	273 (50 %)	274 (50 %)	148 (54 %)	90 (33 %)	30 (11 %)	6 (2 %)	0
Наномеры (IG, TCR)	547	60 (11 %)	487 (89 %)	192 (39 %)	155 (32 %)	96 (20 %)	37 (8 %)	7 (1 %)
Гептамеры (IG)	325	192 (59 %)	133 (41 %)	83 (63 %)	33 (25 %)	17 (12 %)	0	0
Гептамеры (TCR)	222	81 (36 %)	141 (64 %)	65 (46 %)	57 (41 %)	13 (9 %)	6 (4 %)	0
Наномеры (IG)	325	43 (13 %)	282 (87 %)	150 (53 %)	66 (23 %)	48 (17 %)	14 (5 %)	4 (2 %)
Наномеры (TCR)	222	17 (8 %)	205 (92 %)	42 (20 %)	89 (43 %)	48 (24 %)	23 (11 %)	3 (2 %)

Значительная часть полиморфных последовательностей отличается от них на одно основание (54 % гептамеров и 39 % наномеров), реже на большее их количество. Нужно отметить, что полиморфизм структурных элементов RSS иммуноглобулинов выражен слабее, чем у T-клеточных рецепторов. Из всех трех генов наибольшее количество канонических гептамеров и наномеров обнаружено в гене κ-легкой цепи иммуноглобулина (рис. 1). В целом, нуклеотидный состав наномеров менее консервативен, чем гептамеров. Мы выявили, 86 поли-

морфных типов гептамеров и 165 типов наномеров. Из них в структуре RSS наиболее часто обнаруживались последовательности CACTGTG и ACACAAAACC (32 и 64 раза).

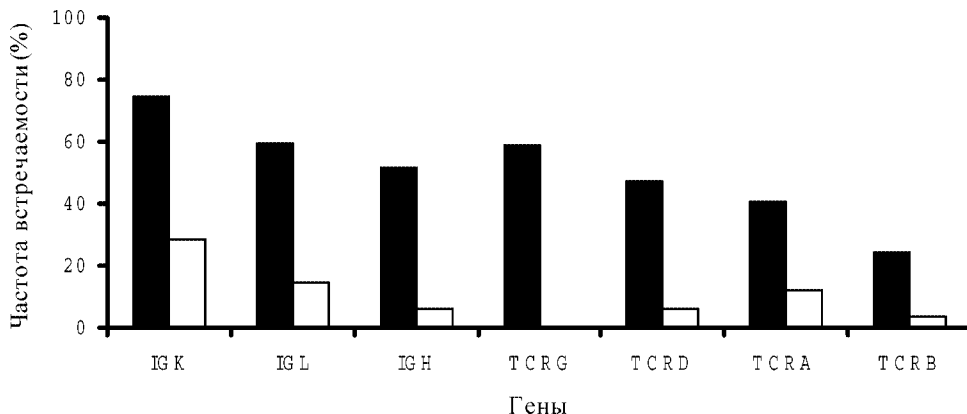


Рис. 1. Частота встречаемости канонических гептамеров (CACAGTG) и наномеров (ACAAAAACC) в генах иммуноглобулинов и Т-клеточных рецепторов человека. Гептамеры отмечены черным, а наномеры — белым

Мы установили, что в основном замены в гептамерах происходят по последним четырем позициям консенсуса (AGTG) - соответственно с 84 %, 87 %, 79 % и 82 % частотой. Кроме описанного в литературе тринуклеотида AAA, высокую степень консервативности проявляет также второй цитозин наномера. Его замена на другое основание происходит с 9 % частотой. Нужно отметить, что консервативность нуклеотидного состава как гептамеров, так и наномеров иммуноглобулинов выше, чем у Т-клеточных рецепторов (табл. 2).

Таблица 2

Консервативность нуклеотидных позиций гептамеров и наномеров сигнальных последовательностей рекомбинации IG и TCR генов человека (в процентах)

Гептамеры генов	Нуклеотидные позиции консенсуса							Наномеры генов	Нуклеотидные позиции консенсуса								
	C	A	C	A	G	T	G		A	C	A	A	A	A	A	C	C
IG, TCR	<u>97</u>	<u>96</u>	<u>95</u>	84	87	79	82	IG, TCR	74	<u>91</u>	82	44	<u>90</u>	<u>96</u>	<u>90</u>	82	73
IG	<u>96</u>	<u>94</u>	<u>93</u>	85	<u>91</u>	<u>90</u>	<u>91</u>	IG	77	<u>92</u>	86	41	<u>90</u>	<u>95</u>	<u>91</u>	<u>90</u>	83
TCR	<u>99</u>	<u>99</u>	<u>97</u>	84	81	63	68	TCR	69	89	75	47	<u>90</u>	<u>97</u>	<u>90</u>	72	59

Примечание: подчеркнута частота встречаемости нуклеотидов в наиболее консервативных позициях гептамеров и наномеров

Проанализировав размер 547 спейсерных участков RSS, мы обнаружили, что только 461 из них (84 %) состоят из 12 или 23 нуклеотидов. Кроме этих общеизвестных и их производных, встречаются спейсеры размером 14, 15, 16, 17, 19, 20, 21 и 25 нуклеотидов (табл. 3).

Таблица 3

Количество разных типов спейсеров, обнаруженных в структуре сигнальных последовательностей рекомбинации IG и TCR генов человека

Гены	Типы длин спейсеров (в нуклеотидах)														Общее количество спейсеров
	11	12	13	14	15	16	17	19	20	21	22	23	24	25	
IG, TCR	6	210	3	1	2	1	2	3	4	6	40	251	12	6	547
IG	4	130	2	0	2	1	2	3	2	5	19	140	10	5	325
TCR	2	80	1	1	0	0	0	0	2	1	21	111	2	1	222

Основной вывод работы состоит в том, что сигнальные последовательности рекомбинации — это полиморфные структуры, в которых нуклеотидный состав гептамеров и наномеров может значительно отличаться от общеизвестных канонических последовательностей.

Литература

1. David G. Schatz. V(D)J recombination // Immunological Reviews. — 2004. — V. 200. — P. 5–11.
2. Bassing C. H., Swat W. and Alt F. W. The mechanism and regulation of chromosomal V(D)J recombination // Cell. — 2002. — V. 109. — P. 45–55.
3. Yoshiko Akamatsu, Naoyo Tsurushita, Fumikiyo Nagawa, Masao Matsuoka, Kenji Okazaki, Mutsuo Imai, and Hitoshi Sakano. Essential residues in V(D)J recombination signals // The Journal of Immunology. — 1994. — V. 153. — P. 4520–4529.
4. Dale A. Ramsden, Kristin Baetz and Gillian E. Wu. Conservation of sequence in recombination signal sequence spacers // Nucleic Acids Research. — 1994. — V. 22. — P. 1785–1796.
5. Matsuda F., Ishi K., Bourvagnet P., Kumma I., Hayashida H., Miyata T. and Honjo T. The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus // J. Exp. Med. — 1998. — V. 188. — P. 2151–2162.

А. Ю. Губський¹, О. В. Жук²

¹ Одеський національний університет ім. І. І. Мечникова, кафедра фізіології людини і тварин, вул. Дворянська, 2, Одеса, 65026, Україна

² Опольський університет, кафедра молекулярної та експериментальної біології, вул. Олеска, 48, Опольце, 45052, Польща

ПОЛІМОРФІЗМ СТРУКТУРНИХ ЕЛЕМЕНТІВ СИГНАЛЬНИХ ПОСЛІДОВНОСТЕЙ РЕКОМБІНАЦІЇ ГЕНІВ ІМУНОГЛОБУЛІНІВ І Т-КЛІТИННИХ РЕЦЕПТОРІВ ЛЮДИНИ

Резюме

Встановлена структура 547 сигнальних послідовностей рекомбінації трьох генів імуноглобулінів і чотирьох генів Т-клітинних рецепторів людини. Нуклеотидний склад 50 % гептамерів і 89 % наномерів не відповідає канонічним послідовностям CACAGTG і ACAAAAACC. Тільки 84 % RSS містять 12- або 23- нуклеотидні спейсери. Структура послідовностей рекомбінації імуноглобулінів менш поліморфна, чим Т-клітинних рецепторів.

Ключові слова: гептамери, наномери, RSS.

A.Yu. Gubsky¹, O. V. Zhuk²

¹ Odessa National University, Department of Human and Animals Physiology, Dvoryanskaya St., 2, Odessa, 65026, Ukraine

² University of Opole, Department of Molecular and Experimental Biology, Oleska 48, Opole, 45052, Poland

STRUCTURAL POLYMORPHISM RECOMBINATION SIGNAL SEQUENCES OF HUMAN IMMUNOGLOBULIN AND T-CELL RECEPTOR GENES

Summary

The structure of 547 recombination signal sequences of three human immunoglobulin genes and four human T-cell receptor genes have been determined. Nucleotide content of 50 % heptamers and 89 % of nanomers does not correspond to CACAGTG and ACAAAAACC canonical sequences. Only 84 % of RSS contains 12- or 23- nucleotide spacers. The RSS structure of immunoglobulin genes is more conservative than T-cell receptors.

Keywords: heptamer, nanomer, RSS.