

С. З. Мальченко

КОМПЬЮТЕРНЫЙ ПОИСК В A₁₁-ПОВТОРАХ ПРОМОТОРОПОДОБНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДЛЯ РНК-ПОЛИМЕРАЗЫ II И ПОТЕНЦИАЛЬНЫХ ПЕПТИДОВ A₁₁

В первом мономере A₁₁-повторов компьютерным анализом обнаружены СААТ- и ТАТА-блоки, характерные для промоторов эукариотических генов, транскрибируемых РНК-полимеразой II. Во втором мономере A₁₁-повторов обнаружены открытые рамки считывания, кодирующие короткие пептиды. Два мажорных класса этих пептидов составляют около 75% всех потенциальных пептидов A₁₁, они названы соответственно ЕАЕ- и GG-потенциальными пептидами A₁₁.

Функциональная роль A₁₁-повторов в настоящий момент неочевидна [1]. Показано, однако, наличие в A₁₁ энхансероподобных структур, промотора РНК-полимеразы III и ряда других функциональных сайтов [1-4]. В цитоплазме клеток идентифицированы poly(A)⁺ и poly(A)⁻ A₁₁+ РНК размерами от 100 до 1500 п. н., функциональная роль которых пока не выяснена [3-5]. В связи с этим представляет определенный интерес изучение возможности функционирования A₁₁-повторов как матрицы для РНК-полимеразы II.

Материалы и методы. В работе использован минибанк A₁₁-повторов, любезно предоставленный В. В. Капитоновым (Ин-т цитологии и генетики Сиб. отд-ния АН СССР). Статистическая обработка заключалась в оценке неслучайного сходства с функциональным сайтом, предложенной в работе [6].

Результаты и обсуждение. В генах, транскрибируемых РНК-полимеразой II, ранее обнаружены два консенсусных блока: СААТ-блок (GGC/TCAATCT) и ТАТА-блок (ТАТАА/ТАА/Т) с относительным расстоянием между ними 45—55 п. н. [1]. На выборке из 65 A₁₁-повторов компьютерным анализом нами определены следующие блоки: GG/AC/TCAA/GCCT (уровень совпадения с СААТ-блоком 80%) и ТАААААТАСАААА (уровень совпадения с ТАТА-блоком 90%), находящиеся соответственно в средних позициях 81 и 121 п. н. (относительное расстояние между блоками 40 п. н.). Определен достоверный интерес для встречаемости СААТ- и ТАТА-блоков в A₁₁, он соответственно равен P_{СААТ} = 0,85 ± 0,15 и P_{ТАТА} = 0,89 ± 0,14, т. е. от 70 до 100% A₁₁ в генеральной совокупности A₁₁-повторов содержат СААТ-блок и от 75 до 100% — ТАТА-блок. Оценка сходства ТАТА-блока A₁₁ с консенсусом ТАТА-блока показала его неслучайность.

С первого (после ТАААААТАСАААА-блока) АТG-кодона (средняя позиция 165 п. н.) считаны потенциальные пептиды (ПП). Определены девять классов ПП, среди которых выделены два мажорных, составляющих 75% всех ПП A₁₁-повторов. После выравнивания (вручную) аминокислотных последовательностей выведен консенсус для этих двух классов ПП, названных нами соответственно ЕАЕ-пептидом и GG-пептидом:

MPVI/VPAT/IREAEAGESLEPR/GRQ/LRLQ/R	ЕАЕ-ПП
MVAIIACSPSYS/LGG	GG-ПП

Для выявления потенциальных функций A₁₁-пептидов проведен поиск сходства с известными последовательностями пептидов. В геноме человека обнаружен ген *son3*, кодирующий ДНК-связывающий белок [4]. В его структуре найдены tandemные повторы, обозначенные от А до Е и представляющие собой кластеры аргининов. Сравнение с С'-концом ЕАЕ-пептида показало большое (>60%) сходство данных структур. Во многих эукариотических мРНК имеются длинные лидерные последовательности, которым характерно наличие коротких открытых рамок

© с. з. мальченко, 1991

считывания (ОРС). Предполагается, что они вовлечены в контроль генной экспрессии на разных уровнях [8]. В этой же работе определены последовательности коротких ОРС в лидерных районах мРНК *Caulimoviruses*. Сравнение N'-конца D'-короткой ОРС *Caulimoviruses* с C'-концом ЕАЕ-пептида показало высокую степень сходства данных структур.

А-повтор <i>son3</i>	TPSRRSR
ЕАЕ-ПП	EP-RRLRI R
С-повтор <i>son3</i>	SPRLRRSR
ЕАЕ-ПП	SPRLLRL
ОРС D'	MRRLNLEPRRRRLE
ЕАЕ-ПП	GESLEPRRLRLR

На C'-конце РНК-полимеразы II дрожжей обнаружен блок из семи аминокислот, тандемно повторяющийся 26 раз. Предполагается, что он участвует в точном узнавании промотора за счет взаимодействия с регуляторным белком типа актина [9]. Сравнение блока с GG-пептидом Alu также выявило высокую степень сходства (>50%) данных структур.

7 АК-блок	PTSPSYS
GG-ПП	MVAHACS PSYSGG

Резюме

У першому мономері Alu-повторів комп'ютерним аналізом виявлені СААТ- і ТАТА-блоки, які характерні промоторам еукаріотичних генів, що транскрибуються РНК-полімеразою II. У другому мономері Alu-повторів знайдені відкриті рамки зчитування, які кодують короткі пептиди. Два мажорних класи цих пептидів складають біля 75% усіх потенціальних пептидів Alu, їх названо відповідно ЕАЕ- та GG-потенційними пептидами.

Summary

CAAT and TATA boxes typical on many eukaryotic RNA-polymerase II promoters are observed in the first monomer of Alu-repeats. The computer analysis has shown that these boxes are homological to correspondent functional sites. Open reading frames coding short peptides are found in the second monomer of Alu-repeats. Two major types of peptides are distinguished among these potential peptides. They contain up to 75% of all the peptides and are called EAE and GG-peptides, respectively.

СПИСОК ЛИТЕРАТУРЫ

1. Льюин Б. Гены.— М.: Мир, 1987.— 544 с.
2. Эволюция повторов Alu / И. А. Шахмурадов, Н. А. Колчанов, Л. В. Омеляничук и др.— Новосибирск, 1987.— 55 с.— (Препринт / АН СССР, Ин-т цитологии и генетики).
3. Корнеев С. А. Повторяющиеся последовательности генома человека // Генетика.— 1988.— 24, № 6.— С. 965—979.
4. Elder J. T., Pan J., Weissman S. M. Transcriptional analysis of interspersed repetitive polymerase III transcription units in human DNA // Nucl. Acids Res.— 1981.— 9, N 9.— P. 1171—1189.
5. Limborska S. A., Korneev S. A., Mateeva N. E. Cloning of ALU-containing cDNAs from human fibroblasts and identification of small ALU+poly(C)+RNAs in a variety of human normal and tumor cells // FEBS Lett.— 1987.— 212, N 2.— P. 208—212.
6. Введение в теорию генетических текстов / В. В. Соловьев, Ф. Э. Кель, И. Б. Рогозин, Н. А. Колчанов.— Новосибирск, 1987.— 92 с.
7. Бердичевский Ф. Б., Чумаков И. М., Киселев Л. Л. Расшифровка первичной структуры участка *son3* генома человека: идентификация нового белка, имеющего необычную структуру и гомологию с ДНК-связывающими белками // Молекуляр. биология.— 1988.— 22, № 3.— С. 794—801.
8. The leading sequence of caulimovirus large RNA can be folded into a large step-loop structure / J. Futterer, K. Gordon, J. M. Bonneville et al. // Nucl. Acids Res.— 1988.— 16, N 17.— P. 8377—8390.
9. Extensive homology among the larges subunits of eukaryotic and prokaryotic RNA polymerases / L. A. Allison, M. Moyle, M. Shales, C. J. Ingles // Cell.— 1985.— 42, N 2.— P. 599—610.

Ин-т молекуляр. биологии и генетики АН УССР, Киев

Получено 27.08.90