

УДК 577.21

Ш. А. АТАМБАЕВА, В. А. ХАЙЛЕНКО

СВЯЗЬ ИЗМЕНЕНИЙ ДЛИНЫ ЭКЗОНОВ И ИНТРОНОВ В ГЕНАХ ХРОМОСОМ 1 И 2 ЧЕЛОВЕКА

(Казахский национальный университет им. аль-Фараби)

Определены участки ДНК с минимальным, средним и максимальным количеством генов, в которых изучалась экзон-интронная организация генов. Показано, что в хромосомах 1 и 2 человека существует зависимость экзон-интронной организации генов от числа интронов в них и плотности генов в участках ДНК.

Число экзонов, интронов и их длина в генах эукариот значительно изменяются [1–5]. Установлены предпочтения длин экзонов и интронов в некоторых организмах [6]. В генах человека диапазон изменения длины интронов составляет от нескольких десятков до десятков тысяч нуклеотидов и в среднем экзоны в 22 раза короче интронов [7]. Огромное разнообразие длин интронов и экзонов дает возможность выяснить причины вариабельности экзон-интронной организации генов.

Цель работы заключалась в выявлении связи между изменениями длины экзонов и интронов в генах различных хромосом человека, в том числе в участках ДНК, отличающихся по плотности генов.

МАТЕРИАЛЫ И МЕТОДЫ

Нуклеотидные последовательности генов и ДНК хромосом 1 и 2 заимствованы из GenBank (<http://www.ncbi.nlm.nih.gov>). ДНК каждой хромосомы последовательно делили на участки длиной 1 Мбр и в каждом находили число генов. Участки по числу генов в них распределяли в группы с различной плотностью генов: 1–11, 12–20, 21 и более генов/Мбр. В каждой группе гены распределяли в выборки с 1–2, 3–5, 6–9, 10–14, 15 и более интронами в гене. Определяли длину экзонов, интронов, сумму длин экзонов в гене, длину гена и долю длины экзонов в гене. Анализировали количество интронов и экзонов с длиной в интервалах 1–20, 21–40, 41–60 н и т. д. до 400 н, а также с длиной более 400 н.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В геноме человека проявляется большая гетерогенность плотности распределения генов вдоль каждой из хромосом. Число генов на 1 Мбр ДНК хромосомы 1 изменялось от нуля (15 участ-

тков) до 59 генов/Мбр и в среднем составляло 10 генов/Мбр. В хромосоме 2 число генов на 1 Мбр изменялось от нуля (53 участка) до 36 генов/Мбр и в среднем было 5 генов/Мбр.

В генах каждой из выборок длина экзонов варьировала. Распределение экзонов по длине генов первой группы (4 гена/Мбр) из хромосомы 1 приведено в табл. 1. В генах с 1–2 интронами доля экзонов с длиной более 400 н равнялась 19,2%, и в генах с 15 и более интронами таких экзонов было только 1,3% (рис. 1.1). Длина экзонов при увеличении числа интронов в генах изменялась с образованием оптимума их длины в интервале 60–160 н. Такой же характер изменения длины экзонов при увеличении числа интронов в гене наблюдался в третьей группе генов (32 гена/Мбр), т. е. не зависел от плотности генов в участках ДНК (рис. 1.3).

Во второй хромосоме в генах с 1–2 интронами доля экзонов с длиной более 400 н была 16,4%, а в генах с числом интронов 15 и более равнялась 2,0%. Изменение длины экзонов в генах разных групп происходило так же, как и в генах хромосомы 1.

Интроны с длиной менее 400 н входили в состав генов первой группы хромосомы 1 редко, и доля интронов с длиной более 400 н составляла 78,4% (рис. 1.2). При увеличении числа интронов в генах это соотношение между короткими и длинными интронами мало изменялось, хотя средняя длина интронов в выборках генов имела тенденцию к уменьшению. В третьей группе генов из выборки генов с 1–2 интронами наряду с уменьшением средней длины интронов появились интроны с длиной в интервале от 100 до 120 н (рис. 1.4).

В выборках генов с различным числом интронов средняя длина экзонов отличалась. В генах с 1–2 интронами длина экзонов была наиболь-

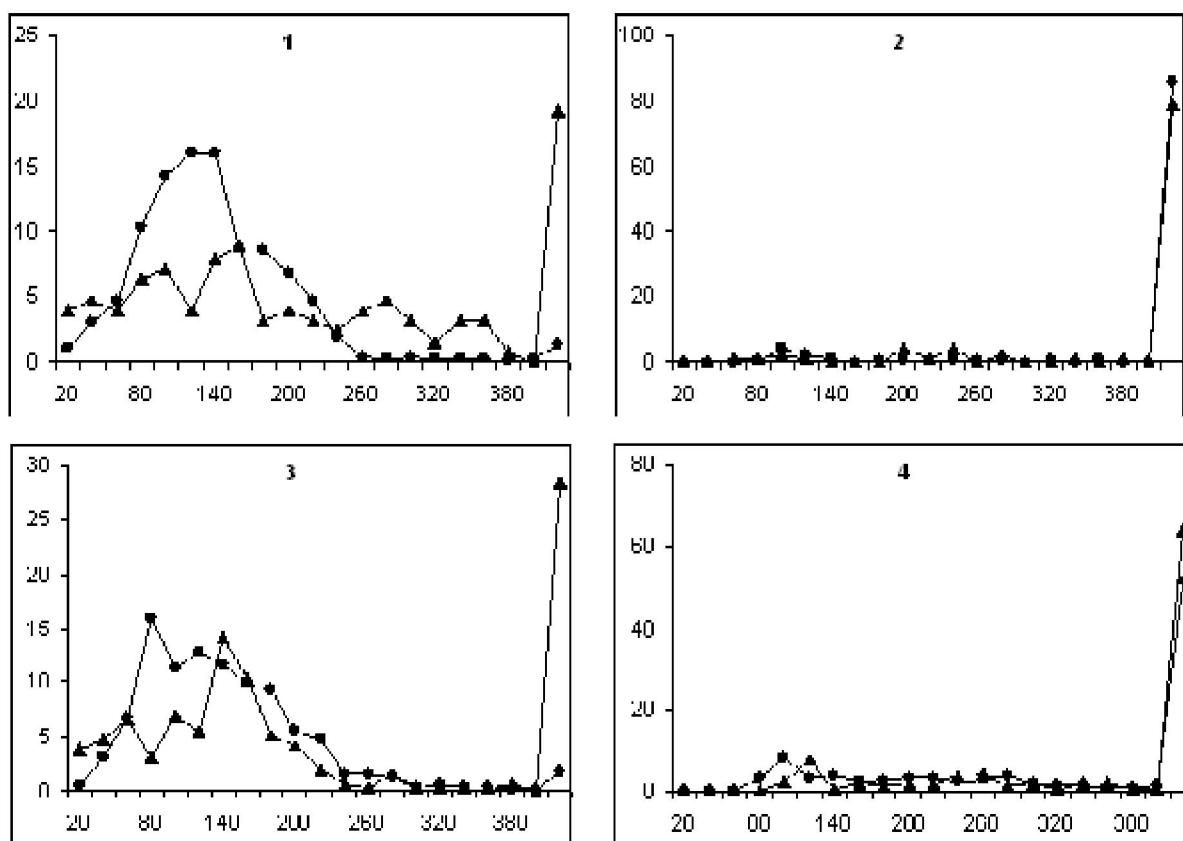


Рис. 1. Длина экзонов (1, 3) и интронов (2, 4) в генах с 1–2 ($\frac{1}{10}$) и в генах с 15 и более интронами ($\frac{2}{21}$) хромосомы 1 в участках с плотностью 4 гена/Мр (1,2) и с плотностью 32 гена/Мр (3,4) Ось x – длина экзонов и интронов (н); ось y – доля экзонов (интронов), %

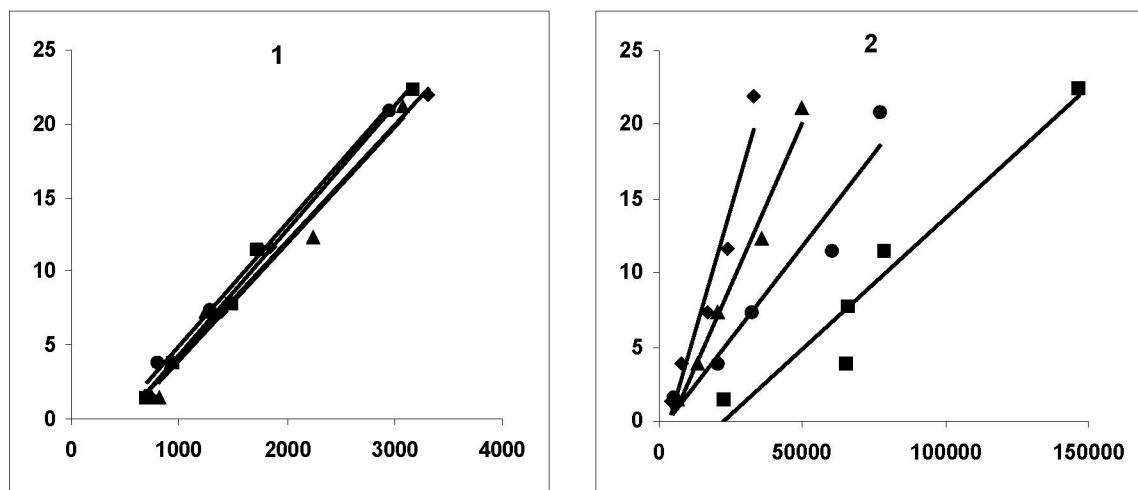


Рис. 2. Зависимость суммарной длины экзонов (1) и длины гена (2) от числа интронов в гене хромосомы 1. В участках ДНК с плотностью генов: 4 гена/Мб – \square , 16 генов/Мб – \square , 26 генов/Мб – \blacktriangle и 32 гена/Мб – \blacktriangle
Ось x – суммарная длина (н) экзонов (1) и генов (2), т.н; ось y – число интронов в гене

шая и уменьшалась с повышением в генах числа интронов (см. табл.1). Эти изменения длины экзонов были сходны во всех группах генов из участков ДНК с различной плотностью генов, и в среднем длина экзонов в генах с 15 и более

интронами была в 2 раза меньше таковой в генах с 1–2 интронами (см. табл.1).

Средняя длина интронов в группах генов заметно отличалась (см. табл.1). В участках с плотностью генов в интервале 1–11 генов/Мбp в

Таблица 1. Характеристики экзон-интронной структуры генов хромосом 1 и 2 *H. sapiens*

| Среднее число интронов | Средняя длина экзона, н | Средняя длина интрона, н | Сумма длин экзонов, н | Средняя длина гена, н | Доля длины экзонов, % | Число генов |
|------------------------|-------------------------|--------------------------|-----------------------|-----------------------|-----------------------|-------------|
| Хромосома 1 | | | | | | |
| 4 гена/Мбр | | | | | | |
| 1,5 | 282 | 15020 | 691 | 22485 | 3,1 | 51 |
| 3,9 | 191 | 16447 | 936 | 65112 | 1,4 | 51 |
| 7,8 | 169 | 8347 | 1478 | 66171 | 2,2 | 72 |
| 11,5 | 136 | 6666 | 1711 | 78642 | 2,2 | 50 |
| 22,4 | 135 | 6399 | 3163 | 146296 | 2,2 | 49 |
| 16 генов/Мбр | | | | | | |
| 1,6 | 274 | 2769 | 708 | 5108 | 13,9 | 73 |
| 3,9 | 164 | 5056 | 801 | 20476 | 3,9 | 83 |
| 7,4 | 153 | 4153 | 1289 | 32181 | 4,0 | 73 |
| 11,5 | 139 | 5106 | 1739 | 60532 | 2,9 | 37 |
| 20,9 | 135 | 3556 | 2946 | 77198 | 3,8 | 59 |
| 26 генов/Мбр | | | | | | |
| 1,5 | 323 | 3456 | 807 | 5991 | 13,5 | 62 |
| 3,9 | 186 | 3149 | 920 | 13326 | 6,9 | 83 |
| 7,3 | 152 | 2590 | 1257 | 20141 | 6,2 | 82 |
| 12,3 | 168 | 2739 | 2247 | 36074 | 6,2 | 43 |
| 21,1 | 140 | 2227 | 3082 | 50022 | 6,2 | 50 |
| 32 гена/Мбр | | | | | | |
| 1,4 | 304 | 2192 | 745 | 3918 | 19 | 105 |
| 3,9 | 194 | 1740 | 948 | 7731 | 12,3 | 108 |
| 7,3 | 167 | 2157 | 1385 | 17115 | 8,1 | 85 |
| 11,6 | 147 | 1904 | 1846 | 23935 | 7,7 | 55 |
| 21,9 | 144 | 1349 | 3308 | 32856 | 10,1 | 43 |
| Хромосома 2 | | | | | | |
| 4 гена/Мбр | | | | | | |
| 1,4 | 280 | 12820 | 676 | 18850 | 3,6 | 91 |
| 4,0 | 168 | 10233 | 841 | 41773 | 2,0 | 93 |
| 7,6 | 174 | 10113 | 1488 | 77866 | 1,9 | 76 |
| 12,0 | 136 | 8424 | 1769 | 103262 | 1,7 | 63 |
| 23,6 | 141 | 6513 | 3472 | 157366 | 2,2 | 105 |
| 15 генов/Мбр | | | | | | |
| 1,5 | 318 | 4800 | 780 | 7749 | 10,1 | 73 |
| 3,9 | 166 | 5307 | 808 | 21305 | 3,8 | 116 |
| 7,4 | 155 | 4707 | 1304 | 36211 | 3,6 | 72 |
| 11,7 | 133 | 3180 | 1696 | 38968 | 4,4 | 57 |
| 26,5 | 181 | 4178 | 4967 | 115685 | 4,3 | 58 |
| 29 генов/Мбр | | | | | | |
| 1,4 | 201 | 3425 | 476 | 5174 | 9,2 | 35 |
| 3,8 | 199 | 1638 | 947 | 7108 | 13,3 | 42 |
| 7,6 | 151 | 2672 | 1304 | 21639 | 6,0 | 54 |
| 11,9 | 152 | 1957 | 1961 | 25245 | 7,8 | 30 |
| 22,8 | 139 | 1507 | 3321 | 37741 | 8,8 | 25 |

изученных хромосомах длина интронов была больше длины интронов генов из участков с плотностью 12–20 генов/Мбр. В генах из участков с плотностью 21/Мбр и более интроны были в среднем в 5 раз короче, чем в генах из участков с наименьшей их плотностью.

Отмеченные изменения длины интронов и экзонов в генах различных групп и выборок не произвольны, а подчиняются определенным за-

кономерностям. Установлено, что увеличение суммы длин экзонов в генах связано с ростом числа интронов в них (рис. 2.1). Эта зависимость описывается уравнением: $N_{in} = aL_{ex} + b$, где a и b – коэффициенты линейной регрессии. Для генов всех групп хромосом 1 и 2 параметры установленных зависимостей приведены в табл. 2.

Полученные данные свидетельствуют о том, что длины экзонов, интронов и генов изменяют-

Таблица 2. Параметры линейной регрессии зависимостей длины гена и суммарной длины экзонов от числа интронов в гене хромосом 1 и 2

| Гены/1 Мбр | Параметры линейной регрессии | | | | | | Количество генов |
|-------------|------------------------------|----------|----------|----------|----------|----------|------------------|
| | <i>a</i> | <i>b</i> | <i>r</i> | <i>c</i> | <i>d</i> | <i>r</i> | |
| Хромосома 1 | | | | | | | |
| 4 | 0,0085 | -4,06 | 0,997 | 0,00018 | -3,96 | 0,966 | 273 |
| 16 | 0,0083 | -3,40 | 0,997 | 0,00025 | -0,70 | 0,967 | 325 |
| 26 | 0,0079 | -3,88 | 0,989 | 0,00043 | -1,64 | 0,991 | 320 |
| 32 | 0,0079 | -3,74 | 0,997 | 0,00066 | -2,12 | 0,971 | 396 |
| Хромосома 2 | | | | | | | |
| 4 | 0,0078 | -3,15 | 1,000 | 0,00016 | -2,93 | 0,984 | 428 |
| 4 | 0,0072 | -3,26 | 0,998 | 0,00013 | -0,48 | 0,991 | 525 |
| 15 | 0,0058 | -0,71 | 0,983 | 0,00024 | -0,39 | 0,985 | 376 |
| 29 | 0,0076 | -2,70 | 0,998 | 0,00060 | -2,16 | 0,964 | 186 |

ся связано и эта связь зависит от плотности генов в участках хромосом 1 и 2. Из этих результатов следует, что интроны выполняют в генах функцию, регулирующую структуру гена. Поскольку экзон-интронная организация генов сильно зависит от плотности генов на участке ДНК, то, вероятно, кластеры генов таких участков выполняют общие функциональные задачи.

ЛИТЕРАТУРА

1. Sulston J. The *Caenorhabditis elegans* genome sequencing project a beginning // Nature. 1992. V. 356. P. 37.
2. Long M., De Souza J.S., Gilbert W. The role of introns in regulated splicing // Cell. 1997. V. 91. P. 739-740.
3. Comeron J.M., Kreitman M. The correlation between intron length and recombination in *Drosophila*: dynamic equilibrium between mutational and selective forces // Genetics. 2000. V. 156. P. 1175-1190.
4. Boudet N., Aubourg S., Toffano-Nioche C., Martin H. Evolution of intron/exon structure of DEAD helicase family in *Arabidopsis*, *Caenorhabditis*, and *Drosophila* // Genome Res. 2002. V. 11. P. 2101-2114.

5. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana* // Nature. 2000. V. 408. P. 796-813.

6. Deutsch M., Long M. Intron-exon structures of eukaryotic model organisms // Nucl. Acids Res. 1999. V. 27. P. 3219-3229.

7. Venter J.C., Adams M.D., Myers E.W. et al., The sequence of the human genome // Science. 2001. V. 291. P. 1304-1351.

Резюме

ДНК-ның гендер саны минималды, орташа және максималды болатын бөліктерінде гендердің экзон-интрондық ұйымдасуы анықталды. Адам 1-ші және 2-ші хромосомаларындағы гендердің экзон-интрондық ұйымдасуы олардағы интрондардың санына және ДНК-ның бөліктеріндегі гендердің тығыздығына тәуелді екені көрсетілді.

Summary

DNA sites with minimum, medium and a maximum quantity of genes have been determined, were exon-intron organization of genes was studied. It was shown dependence exon-intron organization of genes on number of intron in them and density of genes in DNA sites in 1 and 2 human chromosomes.