

Б. Г. КОНОПЕЛЬЧЕНКО, Ю. Б. РУМЕР

КЛАССИФИКАЦИЯ КОДОНОВ В ГЕНЕТИЧЕСКОМ КОДЕ

(Представлено академиком С. Т. Беляевым 3 III 1975)

В работе ⁽¹⁾ была подчеркнута целесообразность отделения в трехбуквенном кодоне ($xy|z$) последней буквы $|z$) (окончания кодона) от первых двух букв ($xy|$ (корня кодона). При этом 16 корней 64 кодонов распадаются на два октета. Первый октет включает восемь сильных корней, каждый из которых при любом из окончаний Ц, Г, У, А кодирует одну и ту же аминокислоту. Второй октет содержит восемь слабых корней, каждый из которых при окончаниях пиримидиновой пары (Ц, У) кодирует одну аминокислоту, а при окончаниях пуриновой пары (Г, А) — другую.

Сильные корни	ЦЦ	ГЦ	УЦ	АЦ	ЦГ	ГГ	ЦУ	ГУ
Слабые корни	АА	УА	ГА	ЦА	АУ	УУ	АГ	УГ

(1)

Таблица корней (1) содержит два типа аномалий. Во-первых, каждая из трех аминокислот — серин, лейцин и аргенин — кодируется двумя кодами с различными корнями, один из которых является сильным, а другой — слабым ⁽¹⁾. Вторая аномалия состоит в том, что для некоторых кодонов со слабыми корнями, а именно корнями (УА|, (АУ| и (УГ|, и окончаниями пуриновой пары происходит полное снятие вырождения.

В настоящей работе мы рассмотрим свойства двух октетов корней, отвлекаясь от перечисленных выше аномалий. Покажем, что эмпирический материал обнаруживает существование канонической последовательности букв Ц, Г, У, А. Введем также количественную характеристику силы букв и корней и покажем, что в корне ($xy|$ важную роль играют как вторая, так и первая буквы.

Мы также рассмотрим другие возможные классификации корней кодонов и сформулируем гипотезу о распадении 16 корней на два октета по отношению к любому свойству системы кодонов.

Рассмотрим таблицу корней (1). Следуя ⁽¹⁾, мы предположим, что сила корня определяется буквенным содержанием. Поэтому и буквы естественно классифицировать на сильные и слабые. При этом, поскольку количество сильных и количество слабых корней совпадают, то мы имеем также равные количества сильных и слабых букв, т. е. две сильные буквы и две слабые.

Далее, буква Ц, как наиболее часто встречающаяся в сильных корнях, является самой сильной буквой. Аналогично, буква А является самой слабой. Из двух букв Г и У, промежуточной силы, естественно считать сильной букву Г, так как корень (ГГ| является сильным, а корень (УУ| — слабым. В результате получаем последовательность букв

Ц	Г	У	А
самая сильная	сильная	слабая	самая слабая

которую можно назвать канонической с точки зрения классификации корней на сильные и слабые.

Выбранная последовательность букв дает возможность сформулировать правила, определяющие силу корня: а) сила корня, содержащего в качестве

ве второй буквы Ц или А, определяется силой второй буквы; б) сила корня, содержащего в качестве второй буквы Г или У, определяется силой первой буквы.

Мы видим, что сила корня определяется не только второй буквой y , но и первой буквой x . Отметим, что простота правил тесно связана с каноническим порядком букв.

Расположим буквы Ц, Г, У, А в виде двурядной матрицы

$$\begin{pmatrix} \text{ЦГ} \\ \text{УА} \end{pmatrix}$$

Первая строка этой матрицы содержит сильные буквы, вторая — слабые. Первый столбец содержит пиримидины, второй — пурины. Если образовать от этой матрицы тензорный квадрат, то получим четырехрядную матрицу корней (xy) :

$$\begin{pmatrix} \text{ЦЦ} & \text{ГЦ} & \text{ЦГ} & \text{ГГ} \\ \text{УЦ} & \text{АЦ} & \text{УГ} & \text{АГ} \\ \text{ЦУ} & \text{ГУ} & \text{ЦА} & \text{ГА} \\ \text{УУ} & \text{АУ} & \text{УА} & \text{АА} \end{pmatrix}$$

Мы видим, что все сильные корни располагаются над боковой диагональю, а слабые — снизу от нее. Корни, лежащие на самой диагонали, разумно назвать переходными. Отметим, что приведенная матрица корней отличается от аналогичной матрицы в работе (1). В связи с этим отличаются и совокупности переходных корней.

Сформулируем правила, приведенные выше, количественно. Для этого сопоставим каждую букву Ц, Г, У, А с «зарядом» Q , который зависит от того, какое место занимает буква в корне.

Первая буква корня $Q_1(\text{Ц})=2$; $Q_1(\text{Г})=2$; $Q_1(\text{У})=-2$; $Q_1(\text{А})=-2$.

Вторая буква корня $Q_2(\text{Ц})=4$; $Q_2(\text{Г})=1$; $Q_2(\text{У})=-1$; $Q_2(\text{А})=-4$. (2)

Отметим, что при замене $\text{Ц} \leftrightarrow \text{А}$, $\text{Г} \leftrightarrow \text{У}$ $Q \rightarrow -Q$. «Заряд» является количественной характеристикой силы буквы. Сила корня определяется формулой

$$Q(|xy|) = Q_1(x) + Q_2(y).$$

Для сильных корней $Q > 0$, для слабых $Q < 0$ (см. (3)).

$$\begin{pmatrix} 6 & 6 & 3 & 3 \\ 2 & 2 & -1 & -1 \\ 1 & 1 & -2 & -2 \\ -3 & -3 & -6 & -6 \end{pmatrix} \quad (3)$$

Отметим, что значения зарядов, предложенные выше, не являются единственно возможными. Общие ограничения на заряды Q имеют вид ($|Q|$ означает абсолютную величину заряда Q):

$$Q_{1,2}(\text{Ц}) = -Q_{1,2}(\text{А}), \quad Q_{1,2}(\text{Г}) = -Q_{1,2}(\text{У});$$

$$Q_2(\text{Ц}) > |Q_2(\text{Г}, \text{У})|, \quad |Q_1(\text{Ц}, \text{Г}, \text{У}, \text{А})|;$$

$$|Q_2(\text{Г}, \text{У})| < |Q_1(\text{Ц}, \text{Г}, \text{У}, \text{А})|.$$

Мы выбрали значения зарядов так, чтобы подчеркнуть отмеченное в (1) распадение каждого из октетов на квартет (например, ЦЦ, ГЦ, УЦ, АЦ) и два дублета (например, ЦГ, ГГ и ЦУ, ГУ). Квартет, в свою очередь, распадается на два дублета.

Обращаясь к зарядам первой буквы x (см. (2)), можно усмотреть связь заряда с количеством водородных связей соответствующей буквы (три — для сильных букв Ц, Г, два — для слабых букв А, У (1)).

Рассмотрим теперь другие возможные классификации корней. В работе ⁽²⁾ было показано, что все 20 аминокислот можно разбить на две группы: полярные и неполярные. Оказалось, что полярность или неполярность аминокислот не зависит от окончания $|z\rangle$ кодона $(xy|z)$, т. е. полностью определяется корнем $(xy|)$. В результате имеем два октета корней ⁽²⁾:

Неполярные	ЦЦ	ГЦ	⋮	ГГ	УГ	⋮	ЦУ	ГУ	⋮	УУ	АУ
Полярные	АА	УА	⋮	ГА	ЦА	⋮	АГ	ЦГ	⋮	АЦ	УЦ

(4)

Корни в (4) расположены таким образом, чтобы подчеркнуть распадение каждого из октетов (по последней букве) на квартет и два дублета. Отметим, что в неполярном октете пять корней являются сильными, три — слабыми; в полярном октете — три сильных и пять слабых корней.

Распадение 16 корней на два октета обнаруживается и при рассмотрении такой характеристики аминокислот, как полярная зависимость ⁽³⁾.

1	ЦЦ	ГЦ	УЦ	АЦ	⋮	ЦУ	ГУ	УУ	АУ	
Полярная зависимость R_m	6,6	7,0	7,5	6,6	:	4,9	5,6	5,0	5,1	(5)
2	ЦА	ГА	УА	АА	⋮	ЦГ	ГГ	УГ	АГ	
Полярная зависимость R_m	8,5	12,7	5,4	10,0	:	9,1	7,9	5,0	8,3	

Значение R_m для корня $(xy|)$ мы определяем как среднее арифметическое R_m всех кодонов с корнем $(xy|)$.

Корни второго октета характеризуются большими значениями R_m , чем корни первого октета. Аномальные значения R_m для корней $(УА|)$ и $(УГ|)$, по-видимому, связаны с тем, что кодоны $(УА|Г)$, $(УА|А)$ и $(УГ|А)$ являются бессмысленными.

Мы видим, что октеты в (5) обладают регулярным строением. Для первого октета вторая буква является пиримидином (Ц или У), во втором октете — пурином (Г или А). Каждый из октетов разбивается на два квартета. Корни внутри каждого квартета характеризуются близкими значениями полярной зависимости R_m . При замене $Ц \leftrightarrow А$, $Г \leftrightarrow У$ октеты переходят друг в друга. Отметим также, что в первом октете содержится шесть сильных корней и два слабых. Соответственно, во втором октете — два сильных и шесть слабых.

Общей чертой рассмотренных нами классификаций корней кодонов — сильные — слабые: неполярные — полярные, по величине полярной зависимости — является распадение 16 корней на два октета, которые в свою очередь распадаются на квартеты, дублеты и т. п. Распадение 16 возможных корней на два октета представляется нам фундаментальным свойством системы кодонов. В связи с этим естественно сформулировать гипотезу: какое бы свойство системы кодонов мы ни рассмотрели, 16 корней кодонов распадаются по отношению к этому свойству на два октета. Эта гипотеза устанавливает соответствие между свойствами системы кодонов и разбиением корней на октеты. Можно думать, что большинству разбиений всех корней на два октета соответствуют классификации кодонов по некоторым свойствам.

Существует большое число способов, которыми можно разбить 16 корней на два октета, и поэтому мы, по-видимому, имеем большое число соответствующих свойств системы кодонов.

Здесь мы приведем некоторые примеры возможных разбиений корней на октеты.

Пример 1.

1	ЦЦ	ГЦ	УЦ	ЦГ	ГГ	АГ	ГУ	ЦА	
2	АА	УА	ГА	АУ	УУ	ЦУ	УГ	АЦ	(6)

При замене $Ц \leftrightarrow Г$, $А \leftrightarrow У$ каждый из октетов переходит сам в себя.

Так как эта замена букв преобразует кодон в антикодон, то в каждом из октетов содержатся как корни кодонов, так и корни соответствующих антикодонов. При преобразовании $\text{Ц} \leftrightarrow \text{А}$, $\text{Г} \leftrightarrow \text{У}$ октеты переходят друг в друга.

В (6) отмечено также распадение каждого из октетов на два триплета и два синглета.

Пример 2.

1	ЦЦ	ГЦ	УЦ	ЦГ	ГГ	АГ	ЦУ	ГА	(7)
2	АА	УА	ЦА	АУ	УУ	ГУ	УГ	АЦ	

В (7) каждый из октетов при замене $\text{Ц} \leftrightarrow \text{Г}$, $\text{А} \leftrightarrow \text{У}$ преобразуется сам в себя, подобно примеру 1. Октеты в (7) интересны также тем, что при перестановке букв в корне (т. е. $(xy| \rightarrow (yx|)$) каждый из них переходит сам в себя. Кроме того, если сгруппировать корни в любом из этих октетов в триплеты и синглеты по первой букве, то они совпадут с триплетными и синглетными октетами с переставленными буквами.

Кроме приведенных выше двух примеров можно сконструировать много других. Способ построения пар октетов корней с требуемыми свойствами достаточно очевиден. Важным и нетривиальным является вопрос о нахождении соответствующих этим разбиениям на октеты свойств системы кодонов.

Авторы признательны А. А. Прокопенко за многочисленные обсуждения.

Институт ядерной физики
Сибирского отделения Академии наук СССР
Новосибирск

Поступило
27 II 1975

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ Ю. Б. Румер, ДАН, т. 183, 225 (1968). ² М. В. Волькенштейн, Генетика, № 4, 119 (1966). ³ C. R. Woese, D. H. Dugre et al., In: Cold Spring Harbor Symp. Quant. Biol., v. 31, 723 (1966).