

## АНАЛИЗ СИММЕТРИЙ ГЕНЕТИЧЕСКОГО КОДА И СТЕПЕНЬ ДЕТЕРМИНАЦИИ КОДОНОВ

Д. Р. Дуплий, С. А. Дуплий

Харьковский национальный университет им. М. Н. Каразина, пл. Свободы, 4, Харьков 61077, Украина

E-mail: Steven.A.Duplij@univer.kharkov.ua. Internet: <http://gluon.physik.uni-kl.de/~duplij>

Поступила в редакцию 20 июня 2000 г.

Предлагается трехмерная модель генетического словаря в терминах введенной числовой характеристики нуклеотидов — степени детерминации, которая отражает абсолютную разность содержания пиримидиновых и пуриновых оснований в одной нити ДНК. В рамках этой модели прослеживаются выраженные симметричные закономерности и групповые свойства, которые описаны. Предлагается использовать степень детерминации для анализа генетических текстов а также предсказания структур и значения различных функциональных участков ДНК.

**КЛЮЧЕВЫЕ СЛОВА:** генетический код, кодон, кубическая матрица, пурин-пиримидиновая инверсия, степень детерминации, симметрия, группа вращений

Понимание современных концепций молекулярной генетики невозможно без дальнейшего всестороннего изучения аппарата экспрессии генов [1, 2]. Весьма актуальным в этом направлении являются алгебраические подходы к проблеме эволюции эукариотического генетического кода [3, 4], заключающиеся в применении теоретико-групповых методов [5], поиске симметрий [6, 7] и суперсимметрий [8, 9, 10], техники векторных пространств [11]. Однако внутренняя структура генетического кода и топология ДНК требуют дальнейшего исследования [12, 13, 14].

В работе предлагается трехмерная модель генетического словаря, на основании разной степени детерминации нуклеотидов  $d$ , которая описывается количественно. Внутри полученной модели прослеживаются выраженные симметричные закономерности и групповые свойства. Введенная степень детерминации нуклеотида  $d$  отражает абсолютную разность содержания пиримидиновых и пуриновых оснований в наборах нуклеотидов и находится с ней в периодической зависимости. Обсуждается возможность использования этой числовой характеристики для анализа и чтения генетических текстов, а также предсказание строения различных функциональных участков ДНК.

### ВЫРОЖДЕННОСТЬ КОДА И СПЕЦИФИЧНОСТЬ НУКЛЕОТИДОВ

Хорошо известно, что генетический код имеет триплетный характер с разной степенью специфичности оснований [15]. Так первые два основания кодона в большей степени детерминируют образование определенной аминокислоты, чем третье основание. Например, любая из аминокислот — глицин, валин, пролин, аланин и треонин — кодируется четырьмя кодонами, и в каждом случае эти четыре кодона различаются только нуклеотидами в третьей позиции. Другими словами 32 кодона, то есть половина всего количества, имеет полное вырождение по третьему основанию  $z$ . Аминокислота полностью задана первыми нуклеотидами  $x$  и  $y$  независимо от нуклеотида  $z$ . Для двукратно вырожденных кодонов характер аминокислоты определяется пуринами ( $z = A, G$ ) или пиримидинами ( $z = C, U$ ) находящимися в третьей позиции. Около двух третей общего количества оснований, присутствующих в ДНК имеют примерно постоянный характер у всех организмов — это основания, занимающие первое и второе положение в триплете, изменчивость состава ДНК определяется основаниями в третьем положении [16, 15].

Корреляция между количеством определенных аминокислот в белке и относительным содержанием гуанина и цитозина в соответствующей ДНК была найдена еще до окончательного выяснения генетического словаря [12]. Увеличение содержания  $G + C$  в ДНК различно отражается на

составе аминокислот в белке и характеризуется следующими тремя зависимостями. Содержание одних аминокислот увеличивается при увеличении количества  $G + C$  в ДНК, содержание других уменьшается, а содержание третьих остается индифферентным относительно количества цитозина и гуанина. В зависимости от содержания гуанина и цитозина дуплеты нуклеотидов, кодирующих аминокислоты разделяются на три класса: не включающие ни гуанин, ни цитозин; содержащие исключительно гуанин и цитозин и промежуточные [12]. Очевидно, что присутствие в кодоне гуанина или цитозина в большой степени детерминирует образование совершенно определенной аминокислоты, в то время как наличие аденина или урацила не является высоко специфичным: например, дуплет  $AA$  соответствует  $Lys$ , и  $Asn$ , дуплет  $UU$  кодирует  $Phen$  и  $Leu$ , в то время как  $CG$  всегда детерминирует  $Arg$ ).

### СТРУКТУРА МАТРИЦЫ ДУПЛЕТОВ

Поскольку основание в третьем положении допускает большую вариабельность для одной и той же аминокислоты, целесообразно рассмотреть первые два основания кодонов  $x, y$  отдельно от третьего  $z$ . При этом шестнадцать возможных дуплетов по способности детерминировать аминокислоту распадаются на два октета. Восемь дуплетов (более “сильных”) однозначно кодирующих аминокислоты независимо от третьего основания, и восемь (“слабых”), у которых третье основание определяет значение кодона. Дуплеты  $xy$  первого и второго октетов резко различаются по составу. В первом октете  $A$  встечается лишь один раз, во втором — один раз  $C$ . Причем, переход от дуплетов “сильного” октета в дуплеты “слабого” октета можно получить следующей заменой [17]

$$C \overset{*}{\longleftrightarrow} A, G \overset{*}{\longleftrightarrow} U, \quad (1)$$

которую мы будем обозначать “операция (\*)” и называть пурин-пиримидиновой инверсией.

Четыре нуклеотида можно расположить по способности однозначно детерминировать аминокислоты в порядке убывания следующим образом [18]:

Пиримидин	Пурин	Пиримидин	Пурин	
<b>C</b>	<b>G</b>	<b>U</b>	<b>A</b>	(2)
очень сильная	сильная	слабая	очень слабая	

Эти нуклеотиды отличаются и по числу водородных связей, которые они могут образовывать с комплементарными нуклеотидами антикодона: каждая из сильных букв ( $C$  и  $G$ ) образует по три водородные связи, в то время как каждая из слабых букв ( $U$  и  $A$ ) образует лишь по две водородные связи [18, 19]. Можно предположить, что, чем больше водородных связей в дуплете, тем меньшее значение имеет взаимодействие  $z \rightarrow z'$  третьего основания с антикодоном.

Для того, чтобы перейти от качественного описания структуры генетического кода (относительно способности кодировать аминокислоту) к количественному, введем числовую характеристику эмпирической “силы” — “степень детерминации” нуклеотида  $d$ . Исходя из (2) будем считать, что степень детерминации  $d$  может принимать значения от одного до четырех соответственно возрастанию этой силы. Если обозначить степень детерминации кодона верхним индексом в скобках, то четверку оснований (2) можно представить в виде вектора-столбца

$$\mathbb{V} = \begin{pmatrix} C^{(4)} \\ G^{(3)} \\ U^{(2)} \\ A^{(1)} \end{pmatrix} \quad (3)$$

и соответствующей вектор-строки

$$\mathbb{V}^T = ( C^{(4)} \quad G^{(3)} \quad U^{(2)} \quad A^{(1)} ). \quad (4)$$

Операция (\*) (1) действует на вектор-столбец  $\mathbb{V}$  следующим образом

$$\mathbb{V}^* = \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{U}^{(2)} \\ \mathbf{G}^{(3)} \\ \mathbf{C}^{(4)} \end{pmatrix} \quad (5)$$

Рассмотрим внешнее произведение [20] вектора-столбца (3) и вектора-строки (4)

$$\begin{aligned} \mathbb{M} &= \mathbb{V} \times \mathbb{V}^T = \begin{pmatrix} \mathbf{C}^{(4)} \\ \mathbf{G}^{(3)} \\ \mathbf{U}^{(2)} \\ \mathbf{A}^{(1)} \end{pmatrix} \left( \mathbf{C}^{(4)} \quad \mathbf{G}^{(3)} \quad \mathbf{U}^{(2)} \quad \mathbf{A}^{(1)} \right) = \\ &= \begin{pmatrix} \mathbf{C}^{(4)}\mathbf{C}^{(4)} & \mathbf{C}^{(4)}\mathbf{G}^{(3)} & \mathbf{C}^{(4)}\mathbf{U}^{(2)} & \mathbf{C}^{(4)}\mathbf{A}^{(1)} \\ \mathbf{G}^{(3)}\mathbf{C}^{(4)} & \mathbf{G}^{(3)}\mathbf{G}^{(3)} & \mathbf{G}^{(3)}\mathbf{U}^{(2)} & \mathbf{G}^{(3)}\mathbf{A}^{(1)} \\ \mathbf{U}^{(2)}\mathbf{C}^{(4)} & \mathbf{U}^{(2)}\mathbf{G}^{(3)} & \mathbf{U}^{(2)}\mathbf{U}^{(2)} & \mathbf{U}^{(2)}\mathbf{A}^{(1)} \\ \mathbf{A}^{(1)}\mathbf{C}^{(4)} & \mathbf{A}^{(1)}\mathbf{G}^{(3)} & \mathbf{A}^{(1)}\mathbf{U}^{(2)} & \mathbf{A}^{(1)}\mathbf{A}^{(1)} \end{pmatrix}. \end{aligned} \quad (6)$$

Отсюда видно, что  $\mathbb{M}$  (6) представляет собой матрицу дуплетов, внутренняя структура которой определяется внешним произведением векторов. Именно этот факт обуславливает высокую степень симметрии матрицы  $\mathbb{M}$  и позволяет исследовать свойства математической модели генетического словаря в терминах абстрактной теории групп.

### СВОЙСТВА МАТРИЦЫ ДУПЛЕТОВ И СТЕПЕНИ ДЕТЕРМИНАЦИИ

Будем полагать, что в первом приближении степень детерминации кодона  $d$  является аддитивной характеристикой, то есть степень детерминации дуплета складывается из степеней детерминации составляющих его нуклеотидов (например,  $d_{\mathbf{GA}} = d_{\mathbf{G}} + d_{\mathbf{A}}$ ). Тогда из матрицы дуплетов  $\mathbb{M}$  можно получить соответствующую (симметрическую) матрицу  $D$  степеней детерминации

$$D = \begin{pmatrix} 8 & 7 & 6 & 5 \\ 7 & 6 & 5 & 4 \\ 6 & 5 & 4 & 3 \\ 5 & 4 & 3 & 2 \end{pmatrix}, \quad (7)$$

в которой явно прослеживаются симметричные закономерности. Отметим, что симметрия матрицы (7) настолько высока, что матрица сингулярна, т. е. ее детерминант равен нулю  $\det D = 0$ , и ранг матрицы равен двум  $\text{rank } D = 2$ , ее дефект также равен двум. Это является следствием того факта, что она есть внешнее произведение (6). Примечательно, что след матрицы равен  $\text{tr } D = 20$  и совпадает с суммой элементов боковой диагонали. Видно также, что по боковой диагонали матрицы (7) и параллельно ей находятся “равносильные” дуплеты. Операция (\*) (1) соответствует отражению матрицы степеней детерминации (7) относительно боковой диагонали

$$D^* = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \\ 5 & 6 & 7 & 8 \end{pmatrix}. \quad (8)$$

и, следовательно, пурин-пиримидиновая инверсия не меняет основных свойств матрицы  $D$ , то есть имеем  $\det D^* = \det D = 0$ ,  $\text{rank } D^* = \text{rank } D = 2$  и  $\text{tr } D^* = \text{tr } D = 20$ .

Если расположить боковую диагональ матрицы  $\mathbb{M}$  горизонтально, то мы получаем ромбическую структуру дуплетов

			CC				= 8	
			GC		CG		= 7 сильные	
		CU		GG		CU	= 6	
AC			UG		GU		CA	= 5 переходные (9)
		AG		UU		GA	= 4	
			AU		UA		= 3 слабые	
			AA				= 2	

соответствующую ромбическому варианту генетического словаря [18, 21], в которой определена “сила” каждого дуплета в терминах его степени детерминации так, что горизонтальные ряды состоят из равносильных дуплетов. Пури-пиримидиновая инверсия (переход (1)) соответствует отражению ромбической структуры относительно диагонали. Дуплеты со степенью детерминации больше 5 кодируют одну аминокислоту, со степенью детерминации меньше 5 кодируют две аминокислоты, значение 5 — промежуточное: из этого ряда **AC** и **GU** детерминируют по одной аминокислоте, а **UG** и **CA** по две. Аминокислоты, кодируемые дуплетами с низкой степенью детерминации (меньше 5) принадлежат к разным классам по химическому типу радикалов, кроме дуплета **GA**, детерминирующего Asp и Glu одного химического класса. В промежуточном случае степень детерминации дуплета пропорциональна числу водородных связей. Пуриновые основания (**G** и **A**) образуют по две водородных связи, а пиримидиновые (**C** и **U**) по три связи.

### ТРЕХМЕРНАЯ МАТРИЦА КОДОНОВ

От матрицы дуплетов  $\mathbb{M}$  (6) можно перейти к структуре триплетного кода следующим образом. Представим, что матрица дуплетов  $\mathbb{M}$  находится в плоскости координат  $xy$ . Умножим эту матрицу на вектор столбец  $\mathbb{V}$  (3), лежащий на оси, перпендикулярной плоскости  $xy$ , то есть построим тройное внешнее произведение  $\mathbb{K} = \mathbb{V} \times \mathbb{M}$ . Таким образом, по аналогии с (6) мы получили трехмерную матрицу четвертого порядка, или кубическую матрицу над множеством триплетов.

Для определения степени детерминации триплетов также воспользуемся предположением об аддитивности, как и для дуплетов (например,  $d_{CGA} = d_C + d_G + d_A$ ). Тогда каждый из 64 элементов (кодонов) кубической матрицы будет иметь числовую характеристику — степень детерминации кодона от 3 до 12, которую для краткости будем называть силой кодона, а суммы степеней детерминации кодонов, лежащих на одной грани — силой грани. Кодоны с минимальной силой 3 — (**AAA**) и максимальной 12 — (**CCC**) лежат в противоположных вершинах куба, причем **CCC** находится в плоскости верхней грани, а **AAA** — в плоскости нижней грани куба. Назовем **AAA(3)-Lys** минимальным полюсом, а **CCC(12)-Pro** — максимальным<sup>1</sup>. Сумма сил по всем четырем диагоналям куба равна 30, что указывает на высокую степень симметрии матрицы. Сумма сил верхней грани куба

<b>CCC(12)</b> –Pro	<b>CCG(11)</b> – Pro	<b>CCU(10)</b> –Pro	<b>CCA(9)</b> –Pro	(10)
<b>CGC(11)</b> –Arg	<b>CGG(10)</b> –Arg	<b>CGU(9)</b> –Arg	<b>CGA(8)</b> –Arg	
<b>CUC(10)</b> –Leu	<b>CUG(9)</b> –Leu	<b>CUU(8)</b> – Leu	<b>CUA(7)</b> –Leu	
<b>CAC(9)</b> –His	<b>CAG(8)</b> –Gln	<b>CAU(7)</b> –His	<b>CAA(6)</b> –Gln	

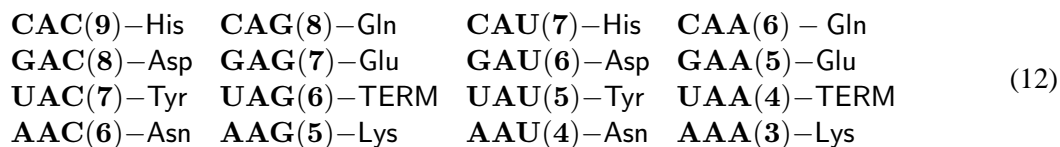
равна 144, а нижней грани

<b>ACC(9)</b> –Thr	<b>ACG(8)</b> –Thr	<b>ACU(7)</b> – Thr	<b>ACA(6)</b> –Thr	(11)
<b>AGC(8)</b> –Ser	<b>AGG(7)</b> –Arg	<b>AGU(6)</b> –Ser	<b>AGA(5)</b> –Arg	
<b>AUC(7)</b> –Ile	<b>AUG(6)</b> –Met	<b>AUU(5)</b> –Ile	<b>AUA(4)</b> –Ile	
<b>AAC(6)</b> –Asn	<b>AAG(5)</b> –Lys	<b>AAU(4)</b> –Asn	<b>AAA(3)</b> –Lys	

равна 96. Боковые грани (наружные плоскости кубической матрицы  $\mathbb{K}$ ) по силе различаются относительно диагонального сечения, построенного через элементы **CAC(9)**, **CCA(9)**, **ACA(6)**,

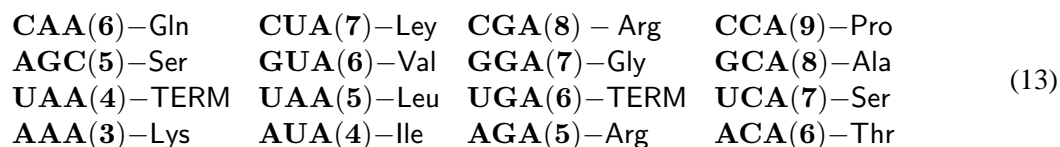
<sup>1</sup>В скобках приводится сила триплета.

**ААС(6)**. Грани, имеющие общее ребро **САА— ААА**



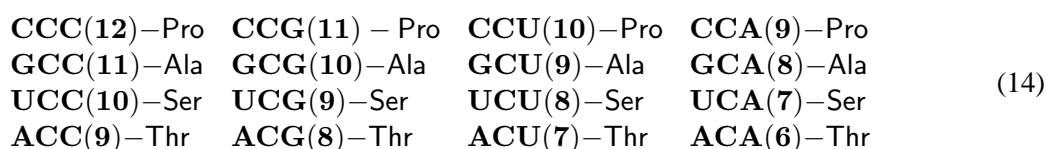
(12)

и



(13)

имеют силу, равную **96**, а грани с общим ребром **ССС— АСС**



(14)

и



(15)

имеют силу, равную **144**. Следовательно, грани имеющие общую точку **ААА(3)** характеризуются силой **96**, а грани, пересекающиеся в точке максимального полюса **ССС(12)**, имеют силу **144** каждая. Это подтверждает высокую симметрию кубической матрицы  $\mathbb{K}$ .

Внутренние плоскости матрицы, параллельные граням, число которых шесть, как и наружных, по сумме сил разделяются на две группы: с силами **112** и **128**. Две вертикальные плоскости [**CAU(7)**, **CCU(10)**, **ACU(7)**, **AAU(4)**] и [**CUC(10)**, **CUA(7)**, **AUA(4)**, **AUC(7)**], линия пересечения которых проходит через **CUU(8)– AUU(5)**, а также плоскость [**UAC(7)**, **UCC(10)**, **UCA(7)**, **UAA(4)**], пересекающая две первые и имеющая с ними общую точку **UUU(6)** имеют силу **112** каждая и находятся ближе к минимальному полюсу. Плоскости [**CGC(11)**, **CGA(8)**, **AGA(5)**, **AGC(8)**] и [**CAC(8)**, **CCG(11)**, **ACG(8)**, **AAC(5)**] пересекаются по следующей прямой **CGG(10)– AGG(7)** и имеют с плоскостью [**GAC(8)**, **GCC(11)**, **GCA(8)**, **GAA(5)**] общую точку **GGG(9)**. Сила каждой из них равна **128**. Отметим что, точками пересечения равносильных плоскостей являются монотонные тринуклеотиды, находящиеся на главной диагонали. Отсюда следует правило:

*Сила грани или плоскости, параллельной грани,  $d_{plane}$  однозначно определяется ее единственным монотонным тринуклеотидом **ААА(3)**, **UUU(6)**, **GGG(9)**, **ССС(12)** и равна **96**, **112**, **128**, **144** соответственно.*

Важно, что график зависимости силы плоскости  $d_{plane}$  от силы принадлежащего ей монотонного тринуклеотида  $d_{mono}$  представляет собой прямую линию

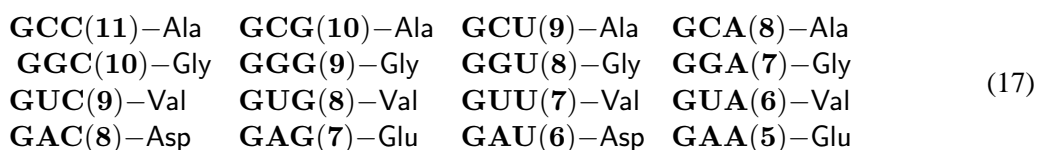
$$d_{plane} = \frac{16}{3}d_{mono} + 80, \quad (16)$$

что является следствием высокой симметрии кубической матрицы.

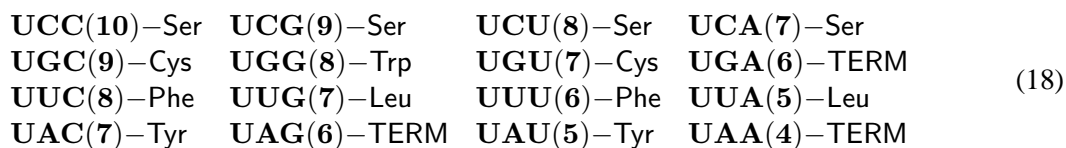
## КУБИЧЕСКАЯ МАТРИЦА КОДОНОВ И ВЫРОЖДЕННОСТЬ КОДА

Рассмотрим особенности расположения кодонов и их значений в построенной кубической матрице  $\mathbb{K}$ . Поскольку каждому кодону, кроме трех терминальных, соответствует аминокислота, то можно говорить о трехмерной модели генетического словаря.

Отметим следующие свойства этой модели. Гидрофобные аминокислоты Leu, Val, Phe, Ile лежат в одной плоскости **CUC(10)**, **CUA(7)**, **AUA(4)**, **AUC(7)**. Оксимоноаминокарбоновые кислоты серин (Ser) и треонин (Thr) находятся на наружной грани **CCC(12)**, **CCA(9)**, **ACA(6)**, **ACC(9)**. На верхней грани куба находятся четыре из шести кодонов лейцина и аргинина, и все кодоны этой грани содержат цитозин. Кодоны пролина занимают строку с наибольшими степенями детерминации **CCC(12)**, **CCG(11)**, **CCU(10)**, **CCA(9)**. Аминокислоты, кодирующиеся четырьмя кодонами, расположены в строку, причем моноаминокарбоновые аминокислоты лежат в одной плоскости



Последнюю строку занимают моноаминодикарбоновые аминокислоты: аспарагиновая и глутаминовая, обладающие гидрофильными свойствами. Уникальные кодоны **AUG(6)**–Met находится в плоскости нижнего основания основания матрицы, а **UGG(8)**–Trp лежит в одной плоскости



с терминальными кодонами **UAA(4)**, **UGA(6)**, **UAG(6)**, которые имеют степени детерминации не больше **6** и находятся на равносильных боковых гранях с  $d = 96$ . В общем случае можно заметить, что, чем больше кодонов кодирует одну аминокислоту, тем больше сила каждого из них (**8-10**), уникальные кодоны, напротив, имеют невысокие степени детерминации (**4-8**).

Определим силу аминокислоты  $d_{AMK}$  как среднее арифметическое сил кодонов  $d_{codon}$ , детерминирующих ее

$$d_{AMK} = \frac{\sum d_{codon}}{n_{deg}}, \quad (19)$$

где  $n_{deg}$  — это ее степень вырожденности. То есть, для каждой из 20 аминокислот мы получили числовую характеристику  $d_{AMK}$ , которая показывает с какой силой детерминируется данная аминокислота. Эта зависимость и зависимость средней силы аминокислоты  $d_{AMK}$  от числа кодонов  $n_{deg}$  представлены в Таблице 1.

AMK	$d_{AMK}$	$n_{deg}$
Lys	4	2
Asn	5	2
Ile	$5\frac{1}{3}$	3
Glu/Met/Tyr	6	2/1/2
Phe/Asp/Gln	7	2/2/2
Val/Thr	$7\frac{1}{2}$	4/4
Leu	$7\frac{2}{3}$	6
Cys/Trp/Ser/His	8	2/1/6/2
Arg	$8\frac{1}{3}$	6
Gly	$8\frac{1}{2}$	4
Ala	$9\frac{1}{2}$	4
Pro	11	4

Это позволяет анализировать различные свойства аминокислот в зависимости от введенной силы (степени детерминируемости  $d_{AMK}$ ).

### СВОЙСТВА РАВНОСИЛЬНЫХ СЕЧЕНИЙ

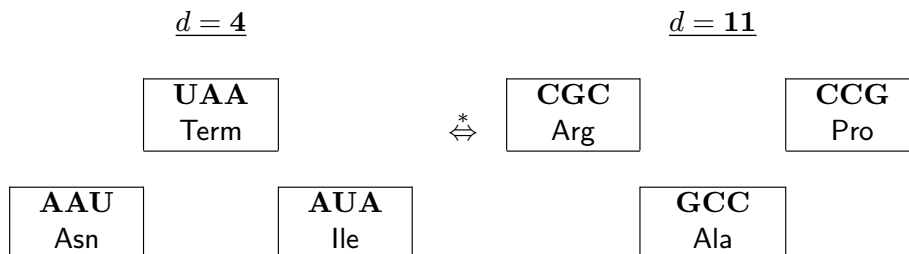
Рассмотрим геометрическое расположение в кубической матрице  $\mathbb{K}$  кодонов, имеющих равную силу. Кодоны с одинаковой силой  $d_{codon} = const$  лежат в плоскостях, перпендикулярных главной диагонали куба **AAA(3)—CCC(12)**. Можно показать, что таких плоскостей (или сечений) десять (с учетом двух угловых одноэлементных), которые представляются в виде различных геометрических фигур. Сила сечения  $d_{section}$  определяется как сила каждого из входящих в него кодонов  $d_{section} = d_{codon} = const$ , а количество элементов в сечении находится в следующей зависимости от  $d_{section}$

Сила сечения $d_{section}$	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>
Число элементов в сечении	1	3	6	10	12	12	10	6	3	1

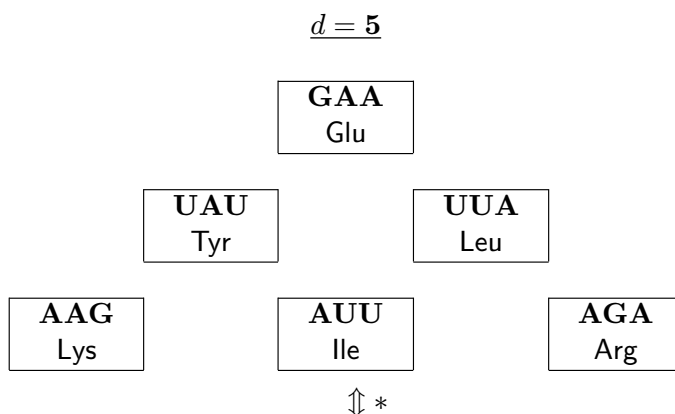
Отсюда видно, что десять сечений распадаются на пять взаимодополнительных пар, имеющих одинаковую форму и количество триплетов и связанных между собой пурин-пиримидиновой инверсией и отражением. Если силу сечения дополнительной пары обозначить  $d_{section}^*$ , то из Таблицы 2 следует формула

$$d_{section} + d_{section}^* = 15. \quad (20)$$

В простейшем случае одноэлементной пары наблюдается монотонная пурин-пиримидиновая инверсия **AAA(3)  $\xleftrightarrow{*}$  CCC(12)**. Следующая пара взаимодополнительных (в смысле формулы (20)) сечений состоит из плоскостей с тремя кодонами



Отметим, что сечение  $d = 4$  не содержит “сильных” дуплетов, а сечение  $d = 11$  не содержит “слабых” дуплетов. В остальных парах происходит смешение “слабых” и “сильных” дуплетов. Например, пара сечений с 6 элементами имеет вид







стрелке, если смотреть вдоль оси в положительном направлении. Обозначим операцию вращения треугольника ААУ на угол  $\frac{2}{3}\pi$  как  $R_1$ , а  $R_2$  — поворот на  $\frac{4}{3}\pi$ , поворот ААУ на нулевой угол обозначим  $E$ . Можно показать, что совокупность операций вращения  $E, R_1, R_2$  вокруг оси  $z$  образует группу. Для доказательства этого утверждения составим таблицу умножения (таблицу Кэли). Согласно определению [22] совокупность  $\Gamma$  элементов  $G_1, G_2, G_3 \dots$  называется группой, если задан закон “умножения” или операция, удовлетворяющая следующим требованиям. Результат умножения двух элементов  $G_a$  и  $G_b$  называется произведением и должен принадлежать группе. В нашем случае элементами группы являются повороты  $E, R_1, R_2$ , а произведением — последовательное выполнение поворотов. Если поворот  $R_1$  переводит систему из положения А в положение В, а поворот  $R_2$  из положения В в положение С, то произведение  $R_1 R_2$  переводит систему из А в С (закон умножения операций вращения). При повороте на нулевой угол треугольник переходит сам в себя, то есть  $E$  — операция тождественная, полностью соответствующая требованиям единичного элемента

$$R_1 \times E = R_1 ; R_2 \times E = R_2.$$

Двойное выполнение поворота на  $\frac{2}{3}\pi$  идентично повороту на  $\frac{4}{3}\pi$ , поэтому  $R_1 \times R_1 = R_2$ . Аналогично находим произведения остальных элементов.

$G_a \setminus G_b$	$E$	$R_1$	$R_2$
$E$	$E$	$R_1$	$R_2$
$R_1$	$R_1$	$R_2$	$R_1$
$R_2$	$R_2$	$E$	$R_1$

Видно, что совокупность элементов  $\Gamma$  является группой, поскольку выполняются условия:

1. Произведение любых двух элементов принадлежит совокупности.
2. Условие ассоциативности, то есть при перемножении трех элементов не должно иметь значения в какой последовательности выполняется это умножение. Поскольку элемент группы — поворот, то это условие выполняется.
3. Единичный элемент оговорен выше, а существование обратного вытекает из таблицы умножения. Как произведение двух операций совмещения, так и действие обратное любой из них, естественно, являются операциями совмещения.

Таким образом мы доказали, что сечение со степенью детерминации  $d = 4$  обладает свойствами группы. Аналогично можно доказать групповые свойства сечения  $d = 11$ , также содержащее три кодона

$$\mathbf{CCG} \rightarrow \mathbf{CGC} \rightarrow \mathbf{GCC} \tag{23}$$

В каждом из сечений можно найти общее число пуриновых и пиримидиновых оснований.

$d_{section}$	$n_C$	$n_G$	$n_U$	$n_A$	$n_C + n_U$	$n_G + n_A$	$\Delta n$
<b>3</b>	0	0	0	3	0	3	-3
<b>4</b>	0	0	3	6	3	6	-3
<b>5</b>	0	3	6	9	6	12	-6
<b>6</b>	3	6	9	12	12	18	-6
<b>7</b>	6	9	12	9	18	18	0
<b>8</b>	9	12	9	6	18	18	0
<b>9</b>	12	9	6	3	18	12	6
<b>10</b>	9	6	3	0	12	6	6
<b>11</b>	6	3	0	0	6	3	3
<b>12</b>	3	0	0	0	3	0	3

График зависимости  $\Delta n$  от  $d$  представляет собой периодическую функцию с областью определения  $\{3;12\}$  и областью значений на отрезке  $\{-6;6\}$ . Эта функция описывает колебания разности количества пиримидинов и пуринов в кодонах с одинаковыми степенями детерминации, что позволяет говорить об определенном биологическом смысле введенной нами степени детерминации нуклеотида  $d$ .

Отметим, что речь идет о зависимостях внутри генетического словаря, применимых и к одноцепочечным нуклеотидным последовательностям в отличие от Чаргаффа[23], описавшего пуринопиримидиновые соотношения в молекуле ДНК. Таким образом, благодаря введенной степени детерминации нуклеотида возможно анализировать не только относительное и качественное, но и абсолютное содержание определенных нуклеотидов в любом генетическом тексте.

## ВЫВОДЫ

Таким образом, алгебраический подход дает возможность по новому взглянуть на проблему генетического кода. Симметрии, наблюдаемые в коде проявляют себя в процессах выбора кодонов для определения различных аминокислот. Предложенная модель построена на основе различной способности нуклеотидов С, G, U, A однозначно детерминировать аминокислоту. Введенное понятие степени детерминации нуклеотида позволяет представлять генетические тексты как последовательность чисел от 1 до 4 (фактически в четверичной системе счисления). Анализ таких последовательностей может привести к более глубокому пониманию процессов транскрипции и, возможно, к формулировке новых принципов конструирования рекомбинантных ДНК, что в настоящее время является неотъемлемым компонентом развития методов клонирования и геной инженерии.

## СПИСОК ЛИТЕРАТУРЫ

1. Сингер М., Берг П. *Гены и геномы*. - М.: Мир, 1998. - 391 с.
2. Стент Г., Кэлиндар Р. *Молекулярная генетика*. - М.: Мир, 1981. - 646 с.
3. Hornos J. E. M., Hornos Y. M. M. *Model for the evolution of the genetic code // Phys. Rev. Lett.* - 1993. - V. **71**. - P. 4401–4404.
4. Jarvis P. D., Bashford J. D. *Systematics of the genetic code and anticodon: history, supersymmetry, degeneracy and periodicity*. - Hobart: 1998. - 3 p. (Preprint / Univ. of Tasmania, physics/9809030).
5. Frappat L., Sorba P., Sciarrino A. *A crystal base for the genetic code*. - Lyon: 1997. - 12 p. (Preprint / Lab. de Phys. Theor. ENS; ENSLAPP-AL-671/97, physics/9801027).
6. Суходолец В. В. *Смысл генетического кода: реконструкция этапа предбиологической эволюции // Генетика*. - 1985. - Т. **21**. - № 10. - С. 1589–1599.
7. Findley G. L., Findley A. M., McGlynn S. P. *Symmetry characteristics of the genetic code // Proc. Natl. Acad. Sci. USA*. - 1982. - V. **79**. - № 22. - P. 7061–7065.
8. Bashford J. D., Tsohantjis I., Jarvis P. D. *A supersymmetric model for the evolution of the genetic code // Proc. Natl. Acad. Sci. USA*. - 1998. - V. **95**. - P. 987–992.
9. Bashford J. D., Tsohantjis I., Jarvis P. D. *Codon and nucleotide assignments in a supersymmetric model of the genetic code // Phys. Lett.* - 1997. - V. **A233**. - P. 481–488.
10. Bashford J. D., Jarvis P. D., Tsohantjis I. *Supersymmetry and the genetic code // Physical Applications and Mathematical Aspects of Geometry, Groups, and Algebras*. - Singapore: World Sci., 1997. - P. 826–831.

11. Bhry T., Cziryk A., Vicsek T., Major B. *Application of vector space techniques to DNA // Fractals*. - 1998. - V. **6**. - № 3. - P. 205–210.
12. Ичас М. *Биологический код*. - М.: Мир, 1971. - 351 с.
13. Ратнер В. А. *Структура и эволюция генетического кода // Итоги науки и техники. Сер. Мол. Биол.* Т. **21**. - М.: ВИНТИ, 1985. - С. 158–197.
14. Maddox J. *The genetic code by numbers // Nature*. - 1994. - V. **367**. - P. 111.
15. Льюин Б. *Гены*. - М.: Мир, 1987. - 544 с.
16. Сингер М., Берг П. *Гены и геномы*. - М.: Мир, 1998. - 373 с.
17. Румер Ю. Б. *О систематизации кодонов в генетическом коде // ДАН СССР*. - 1969. - Т. **187**. - № 187. - С. 937–938.
18. Румер Ю. Б. *Систематизация кодонов в генетическом коде // ДАН СССР*. - 1968. - Т. **183**. - № 1. - С. 225–226.
19. Румер Ю. Б. *О систематизации кодонов в генетическом коде // ДАН СССР*. - 1966. - Т. **167**. - № 6. - С. 1393–1394.
20. Гантмахер Ф. Р. *Теория матриц*. - М.: Наука, 1988. - 548 с.
21. Карасев В. А. *Ромбический вариант генетического словаря на основе комплементарности кодирующих нуклеотидов // Вестн. Ленингр. ун-та*. - 1976. - Т. **1**. - № 3. - С. 93–97.
22. Эллиот Д., Добер П. *Симметрия в физике*. - М.: Мир, 1983. - 364 с.
23. Бреслер С. Е. *Введение в молекулярную биологию*. - Л.: Наука, 1966. - 296 с.