КВАНТОВО-ЛОГИЧЕСКАЯ БИОИНФОРМАТИКА, ГЕНЕТИЧЕСКИЙ АЛФАВИТ УНИТАРНЫХ ОПЕРАТОРОВ АДАМАРА И ЦИКЛИЧЕСКИЕ ГРУППЫ

С.В. Петухов

Российская Инженерная Академия

spetoukhov@gmail.com

Физиологические системы организма для их передачи по наследству генетически кодируются с использованием информационных последовательностей в ДНК и РНК. Формирование информационных сообщений или программ базируется на том или ином алфавите. Например, все компьютерные программы опираются на соответствующие алфавиты программирования. Наука о генетической информатике сегодня базируется на знании алфавитов 4 нуклеотидов ДНК и РНК: аденина A, гуанина G, цитозина C и тимина T (в РНК вместо тимина – урацил U). Но, как подчеркивает Нобелевский лауреат по химии Т. Стайц, все знания о биохимической структуре белков и нуклеиновых кислот, закодированных в геноме, не расскажут нам, например, как летает бабочка [1]. И они не расскажут, как черепашки, вылупившись из яйца, сразу без всякого обучения начинают координированными движениями конечностей ползти к воде, что требует логически согласованной активности миллионов их нервных и мышечных клеток. Подобные генетически наследуемые логические формы коллективного поведения в биосистемах требуют поиска соответствующей скрытой в живой природе биоинформационной системы, базирующейся на соответствующем алфавите. Эта скрытая информатика и ее алфавит, видимо, сопряжены с квантовой механикой и квантовой информатикой, поскольку генетические молекулы принадлежат к микромиру квантовой механики и многие авторы давно подозревают, что живые организмы являются квантово-подобными сущностями. Доклад посвящен найденной автором генетической системе алфавитов 4 унитарных операторов Адамара и развитию квантово-логической биоинформатики на ее основе.

Для поиска алфавитов такой скрытой биоинформационной системы автор обратился к известной генетической закономерности, которая универсальна для высших и низших организмов и которая, в тоже время, не выводится из знания алфавита 4 нуклеотидов ДНК. Речь идет об универсальном правиле статистического строения длинных (более 100 кбит) однонитевых молекул ДНК, известном как феноменологическое второе правило Чаргаффа. Оно до сих пор не нашло общепринятого объяснения в науке [2, 3]. Это правило гласит: в длинных одиночных нитях ДНК процентное содержание аденина и тимина практически одинаково (%A ≃ %T), как и процентное содержание цитозина и гуанина (%C ≃ %G). Отсюда следует, что в длинных одиночных нитях ДНК процентное содержание пуринов A и G практически равно процентному содержанию пиримидинов C и T: %A+%G ≃ %C+%T.

Описываемое ниже использование этих данных второго правила Чаргаффа в генетических матрицах привело к набору 4 унитарных матриц Адамара, предлагаемому автором в качестве генетического алфавита для развития искомой системы квантово-логической биоинформатики. Унитарные операторы являются основой всех вычислений в квантовых компьютерах, где они служат логическими вентилями (гейтами) [4]. В докладе приводятся примеры моделирования биологических явлений с позиций этой новой биоинформационной системы, опирающейся на названный алфавит унитарных операторов, его алгебраические расширения, циклические группы степеней этих унитарных операторов и формализмы квантовой логики. Основоположник квантовой информатики Ю.И.Манин ввел в своей книге [5, с.15] понятие квантового компьютера именно при анализе особенностей скоростной обработки «*генетическими автоматами*» информации ДНК в хромосомах, пророчески указывая при этом на важную роль унитарных операторов и тензорных произведений.

В представляемом исследовании используется понятие генетических матриц, связанное с тем фактом, что ДНК-алфавит 4 нуклеотидов является носителем системы бинарно-оппозиционных индикаторов, включая следующие два субалфавита по два нуклеотида в каждом: 1) два из этих нуклеотидов являются кето-молекулами (T и G), а два других (C и A) – амино-молекулами, что дает бинарное представление **C = A = 0, T = G = 1**; 2) пары комплементарных нуклеотидов A-T и C-G связаны соответственно 2 и 3 водородными связями (называемыми в генетике слабыми и сильными водородными связями), что дает бинарное представление **C = G = 0, A = T = 1**.

В силу этого ДНК-алфавиты 4 нуклеотидов, 16 дуплетов и 64 триплетов представляются в форме квадратных генетических матриц, столбцы которых нумеруются бинарными символами одного из этих суб-алфавитов, а строки – бинарными символами другого [3, 6]. Рисунок показывает пример размещения алфавита 4 нуклеотидов в генетической (2•2)-матрице, строки которой нумеруются бинарными символами суб-алфавита «амино или кето», а столбцы – бинарными символами суб-алфавита «сильных или слабых водородных связей». В этой показанной матрице пиримидины C и T расположены на главной диагонали (по Чаргаффу, их общая вероятность равна 0.5 в длинных однонитевых ДНК), а оба пурина A и G– на второй диагонали (по Чаргаффу, их общая вероятность тоже 0.5). В квантовой механике результаты измерений не предсказуемы и можно говорить только о вероятностях различных исходов. Для их оценки традиционно используется понятие математического ожидания или средневзвешенное значение наблюдаемой величины; в силу этого имеем равенство усредненных вероятностей в паре пиримидинов (%C ≃ %T = 0.25) и в паре пуринов (%A ≃ %G = 0.25). Представляя каждый нуклеотид его усредненной вероятностью 0.25, получаем генетическую матрицу вероятностей, удвоение которой дает матрицу $\mathbf{W_D}$:

	0	1
0	C	A
1	G	T

→

	0	1
0	%C	%A
1	%G	%T

→

	0	1
0	0.25	0.25
1	0.25	0.25

→ $\mathbf{W_D} =$

0.5	0.5
0.5	0.5

Эта матрица $\mathbf{W_D}$ является дважды стохастической, поскольку в каждой строке и столбце суммы вероятностей равны 1. Но в отношении дважды стохастических матриц известна следующая теорема [7]: если матрица $\mathbf{V} = ||\mathbf{v_{ij}}||^n$ унитарная, то матрица $\mathbf{W} = ||\mathbf{w_{ij}}||^n$, где $\mathbf{w_{ij}} = |\mathbf{v_{ij}}|^2$, является дважды стохастической. Согласно этой теореме, дважды стохастической генетической матрице $\mathbf{W_D}$ вероятностей соответствуют 4 унитарные генетические матрицы Адамара $\mathbf{H_C}$, $\mathbf{H_A}$, $\mathbf{H_T}$ и $\mathbf{H_G}$:

$$\mathbf{H_C} = 2^{-0.5} \begin{vmatrix} 1, & -1 \\ 1, & 1 \end{vmatrix}; \quad \mathbf{H_T} = 2^{-0.5} \begin{vmatrix} 1, & 1 \\ 1, & -1 \end{vmatrix}; \quad \mathbf{H_G} = 2^{-0.5} \begin{vmatrix} 1, & 1 \\ -1, & 1 \end{vmatrix}; \quad \mathbf{H_A} = 2^{-0.5} \begin{vmatrix} -1, & 1 \\ 1, & 1 \end{vmatrix}$$

Два из этих унитарных операторов - $\mathbf{H_A}$ и $\mathbf{H_T}$ – являются операторами отражения и при возведении в целые степени образуют циклическую группу с периодом 2. Два другие унитарные оператора - $\mathbf{H_C}$ и $\mathbf{H_G}$ – являются операторами поворотов и при возведении в целые степени образуют циклическую группу с периодом 8.

Выявление данной связи статистических универсалий геномных ДНК с этими 4 унитарными матрицами Адамара важно в силу значимости унитарных преобразований для квантовой механики, квантового компьютинга и пр. Один из этих 4 унитарных операторов Адамара – $\mathbf{H_T}$ – традиционно используется в квантовых компьютерах для операций над кубитами, являясь ключевым гейтом во многих квантовых алгоритмах. Он обеспечивает в квантовых алгоритмах квантовую запутанность и принцип суперпозиции, благодаря чему демонстрируется квантовое превосходство, то есть их значительно более эффективная работа в сравнении с классическими алгоритмами [4]. Квантовая логика – это

алгебраическая система для описания с помощью квантовых гейтов того, как кубиты работают и взаимодействуют и как извлекать из них информацию [8]. В квантовой логике «логика» заключена не в рассуждениях, а в математическом описании состояний и операций. Квантовая логика работает с кубитами, векторами и матрицами, а не со множествами. Она основоположена работой Биркгоффа и фон Неймана (1936г), которые видели в квантовой логике возможное основание физики, и ее математической основой служит теория гильбертовых пространств, унитарных и проективных операторов.

Многие генетически наследуемые структуры в организмах очевидным образом связаны с унитарными преобразованиями поворотов и зеркальных отражений. Например, кинематическая схема тела человека и его локомоций устроена на унитарных преобразованиях поворотов в суставах и зеркальной симметрии левой и правой половин тела, имеющего около 300 суставов. Моторная деятельность человека сводится к искусному управлению нервной системой ансамблями этих унитарных преобразований в кинематике тела, что сопряжено с генетически наследуемой способностью нервной системы оперировать унитарными преобразованиями. При этом само представление человека о схеме его тела носит врожденный характер: люди с отсутствующими от рождения конечностями и не имеющие никакого персонального опыта пользования ими, тем не менее ощущают их реально существующими с фантомными болями в них. При изучении сенсомоторных особенностей человека надо учитывать, что его генетически наследуемая нервная система в своей структурной организации, по всей видимости, родственна генетической системе. Другой биологический пример – конструирование трехмерных форм белков, которые строятся на унитарных преобразованиях поворота соседних сегментов белковых молекул друг относительно друга вокруг углерод-углеродных связей.

Автор рассматривает и использует семейство 4 унитарных операторов качестве базового генетического квантово-логического алфавита для развития на его основе квантово-логической биоинформатики (или грамматики квантово-логической биологии), дающей богатые возможности моделирования множества генетически наследуемых и логически организованных явлений. В частности, генетические унитарные матрицы Адамара сопряжены с соответствующими полными ортогональными системами функций Уолша, являющихся основой особого спектрального анализа сигналов и сопряженных с циклическими кодами Грея, логической голографией, антеннами Уолша, а также фрактальной кривой Гильберта, которой, как известно, соответствует пространственная упаковка хроматина в геноме человека [3]. Тензорное произведение унитарных операторов Адамара всегда дает унитарный оператор Адамара увеличенного порядка. Многоблочные матрицы, построенные на генетических унитарных операторах Адамара, в некоторых случаях дают унитарные операторы, связанные с гиперкомплексными числами (в том числе с гиперболическими или двойными числами) и фракталами. Например, матрица $\mathbf{Q_1} = 2^{-0.5}[\mathbf{H_G}, \mathbf{H_A}; -\mathbf{H_A}, \mathbf{H_G}]$ является унитарным матричным представлением кватерниона Гамильтона («унитарного кватерниона») $\mathbf{Q_1} = 0.5\mathbf{v_0} + 0.5\mathbf{v_1} + 0.5\mathbf{v_2} + 0.5\mathbf{v_3}$, где $\mathbf{v_0}$ – единичная матрица, а $\mathbf{v_1}$, $\mathbf{v_2}$,$\mathbf{v_3}$ – матричные представления мнимых единиц алгебры кватернионов Гамильтона. При возведении этого унитарного кватерниона в целые степени образуется циклическая группа с периодом 6, используемая для квантово-логического моделирования циклических биоструктур. В частности, она используется для моделирования закономерностей цветовосприятия у человека, представленных в 6-секторном цветовом круге Ньютона (понятия красного и других цветов являются психофизиологическими, а не физическими, и относятся к теме особенностей чувственного опыта - «квалиа» -, в которой многие видят ключ к пониманию природы сознания). Унитарная матрица $\mathbf{Q_2} = 2^{-0.5}[\mathbf{H_G}, -\mathbf{H_A}; \mathbf{H_A}, \mathbf{H_G}]$ выступает другим представлением унитарного кватерниона, не коммутирующим с $\mathbf{Q_1}$ (их коммутатор не равен нулю). В развиваемой квантово-логической биоинформатике имеются также другие примеры пар операторов с ненулевыми коммутаторами, тема

которых важна потому, что в квантовой механике ненулевые коммутаторы привели к принципу неопределенности Гейзенберга.

Автором через построение многоблочных матриц на основе унитарных гейтов $\mathbf{H_C}$, $\mathbf{H_A}$, $\mathbf{H_T}$ и $\mathbf{H_G}$ также получены унитарные матричные представления бикватерниона. Возведения «унитарного бикватерниона» в целые степени образуют циклическую группу с периодом 24, которая используется как модель циклических изменений в физиологических системах организмов в соответствие с универсальным суточным биоритмом в 24 часа.

Представляемая квантово-логическая биоинформационная система на алфавитах унитарных операторов полезна также для объяснения, почему виды организмов в эволюции развивались так быстро: сложные органы и ткани формируются не столько благодаря появлению новых генов, сколько благодаря изменению способов использования уже существующих генов под действием операторов квантово-логической информатики (тема квантово-логического дарвинизма). В этой связи автором дополнительно развивается тема квантово-логических операторов, элементы которых являются функциями времени.

В квантово-логическом подходе к биологии с учетом циклических или пульсационных особенностей живых тел автор опирается на модельное представление о биомеханических средах, состоящих из связанных пульсирующих структур; они согласованно меняются во времени и кодируются молекулярно-генетической системой, которая сама циклически структурирована в соответствие с циклическими кодами Грея [9]. Это представление о пульсирующих средах имеет аналогии с архитектурами процессоров, используемых в искусственном интеллекте и известных под названиями систолических или пульсирующих информационных решеток (пульсиров). Это название отражает аналогию их работы с деятельностью сердца. Пульсация в них предстает как волна данных, а процесс вычисления выглядит как распространение волн активности. Данные, поступившие на входы решетки, начинают "пульсировать" сквозь неё, преобразуясь на каждом шаге. Эта архитектура кардинально отличается от архитектуры фон Неймана обычных процессоров, поскольку в ней нет центрального управляющего устройства; все ячейки работают одновременно и синхронно; данные не записываются в классическом смысле, а непрерывно «пульсируют» сквозь структуру процессора, подобно потоку крови через капилляры. Примером служит тензорный процессор Google (TPU), который использует в своей основе большой двумерный систолический массив для выполнения массивных матричных умножений, необходимых для нейронных сетей, с высокой эффективностью.

Литература:

1. **Steitz T.A**. Collecting Butterflies and the Protein Structure Initiative: The Right Questions? - Structure, 2007, Volume 15, Issue 12, 1523 – 1524, DOI 10.1016/j.str.2007.11.005.
2. **Yamagishi M.E.B**. Mathematical Grammar of Biology. Switzerland, Springer, 2017.
3. **Petoukhov S.V., He M**. Algebraic biology, matrix genetics, and genetic intelligence. - Singapore, World Scientific, 2023, 616 p., https://doi.org/10.1142/13468.
4. **Nielsen M.A., Chuang I.L**. Quantum Computation and Quantum Information, (Cambridge Univ. Press, 2010) (https://doi.org/10.1017/CBO9780511976667)
5. **Манин Ю.И**. Вычислимое и невычислимое, М., Советское радио, 1980, 128 с.
6. **Петухов С.В.** Матричная генетика, алгебры генетического кода, помехоустойчивость. – М., РХД, 2008, 316 с.
7. **Прасолов В.В**. Задачи и теоремы линейной алгебры. 2-е изд. - М., РХД, 2008, 536 с.
8. **Васюков В.Л**. Квантовая логика. — М.: ПЕР СЭ, 2005, 191 с. ISBN 5-9292-0142-0 .
9. **Petoukhov S.V**. Genetic code, the problem of coding biological cycles, and cyclic Gray codes. – Biosystems, vol. 246, December 2024, 105349, https://doi.org/10.1016/j.biosystems.2024.105349 .