Алгебраическая биология, гештальт-генетика, наследственный и искусственный интеллект

С.В. Петухов

Лаборатория биомеханических систем Института машиноведения РАН,
http://petoukhov.com/

Тезисы. Представлены результаты анализа структурных особенностей молекулярной системы генетического кодирования. Выявлены закономерности ДНК-последовательностей нуклеотидов в геномах высших и низших организмов. Их обнаружение основано на представлении геномных ДНК-последовательностей в виде многослойных алгебраических структур и сравнительном анализе процентного содержания *n*-плетов в разных слоях. Описаны правила постоянства сумм процентов в специальных тетрагруппировках *n*-плетов в разных текстовых слоях геномной ДНК. Данные постоянства сумм реализуются при очень разных количествах и величинах слагаемых. Эти правила постоянства сумм процентов сопровождаются гиперболическими правилами взаимосвязи количеств в группировках *n*-плетов в разных слоях геномных ДНК-текстов. Обнаруженные правила являются кандидатами на роль универсальных геномных правил. В исследовании используется тензорное семейство матричных представлений ДНК-алфавитов для выявления алгебраических особенностей упомянутых правил. Отмечаются аналогии обнаруженных генетических явлений с явлениями гештальт психологии, что позволяет развивать идеи гештальт генетики. Представляемые результаты свидетельствуют: 1) живые организмы можно трактовать как квантово-информационные алгебро-гармонические сущности; 2) биоинформатика базируется на вероятностных методах многослойного кодирования и обработки информации.

Ключевые слова: ДНК алфавиты, бинарные оппозиции, геномы, многослойные тексты, вероятности, тетрагруппировки, гештальт, матрицы, тензорное произведение, гиперболические числа, проекторы, матричная генетика, алгебраическая и квантовая биология, интеллект.

Содержание

1. Введение.
2. Бинарные-оппозиции в алфавитах ДНК и матрицы.
3. Гештальт правила процентного состава композитных ДНК-текстов геномов, связанные с *m*-позиционными тетрагруппировками *n*-плетов.
4. О гештальт-генетике и наследуемых феноменах гештальт-биологии.
5. Геномные ДНК-тексты и квантовая информатика.
6. Матричные алгебры и матрицы *m*-позиционных тетрагруппировок.
7. Наклонные проекторы и проекторные тетрагруппировки *n*-плетов в слоях ДНК-текстов геномов.
8. Гиперболические правила количеств *n*-плетов в геномах и гармоническая прогрессия.
9. О геометриях, физиологии и искусственном интеллекте.
10. Теория Фрёлиха о коллективных квантовых эффектах в биологических системах по аналогии с квантовым конденсатом Бозе-Эйнштейна.

Заключение.
Выводы.
Приложение I. Пример гештальт правил суммарных процентов *n*-плетов в бактериальном геноме.
Приложение II. О Инь-Ян схемах древнекитайской “Книги перемен” и 4-секционных мандалах в древних традициях.

1. Введение.

Все живые организмы от рождения наделены способностями совершать действия интеллектуального характера. Например, паук плетет паутину с креплением ее концов к случайным опорам, птицы-ткачи ажурно сплетают гнезда сложной формы из сотен соломинок, пчелы строят соты (Рис. 1.1). Примитивными интеллектуальными способностями обладают также организмы, лишенные нервных клеток. Например, впечатляющих пример опубликован в журнале Nature о способностях плесневых грибов творческим образом находить оптимальные пути в лабиринтах [Nakagaki, Yamada, Toth, 2000; Орен Гоберман, 2011]. При создании систем искусственного интеллекта следует подсматривать у живой природы алгоритмы и принципы генетического наследования таких способностей.

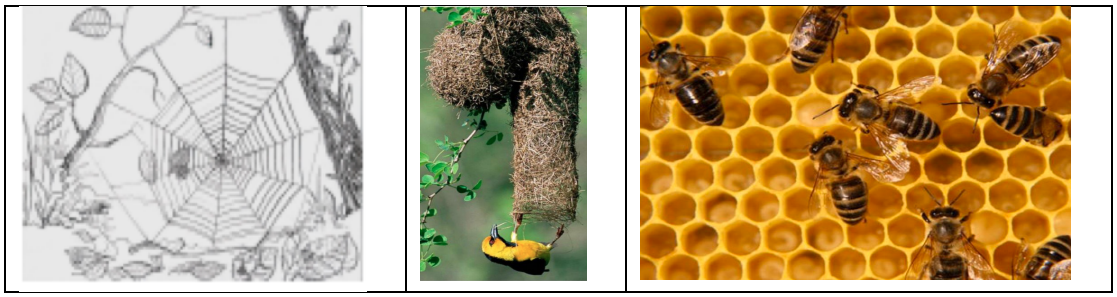


Рис. 1.1. Примеры наследуемой способности к действиям интеллектуального характера у живых организмов: плетение паутины, плетение сложных гнезд птицами-ткачами и постройка сот пчелами.

Существует фундаментальное отличие живых тел от неодушевленных, которое отмечали основоположники квантовой механики П.Йордан и (позже) Э.Шредингер: неодушевленные объекты управляются средним случайным движением их миллионов частиц и движение отдельных частиц не существенно для целого; напротив, в живом организме избранные – генетические - молекулы обладают диктаторским влиянием на весь организм за счет квантового усиления (см. историю квантовой биологии в [McFadden, Al-Khalili, 2018].

Соответственно, в данной статье автор сосредотачивает внимание на структурных особенностях алфавитов молекул наследственности ДНК и длинных последовательностей нуклеотидов (ДНК-текстах) в ДНК геномов высших и низших организмов. Эти последовательности нуклеотидов в одиночной нити ДНК представлены в открытом доступе в GenBank для каждого генома множества организмов в стандартном виде (Рис. 1.2), создающем впечатление хаотичного ДНК-текста из 4-х букв, занимающего в силу его длины тысячи страниц. Например, ДНК-текст первой хромосомы человека

состоит примерно из 230 миллионов букв, что примерно равносильно тексту 100 романов «Война и мир» Л.Н.Толстого.

```
TCTGACCTGAGGAGAACTGTGCTCCGCCTTCAGAGTACCACCGAAATCTGTGCAGAGGACAACGCAGCTC
CGCCCTCGCGGTGCTCTCCGGGTCTGTGCTGAGGAGAACGCAACTCCGCCGTTGCAAAGGCGCGCCGCGC
CGGCGCAGGCGCAGAGAGGCGCGCCGCGCCGGCGCAGGCGCAGAGAGGCGCGCCGCGCCGGCGCAGGCGC
AGAGAGGCGCGCCGCGCCGGCGCAGGCGCAGAGAGGCGCGCCGCGCCGGCGCAGGCGCAGAGAGGCGCGC
CGCGCCGGCGCAGGCGCAGACACATGCTAGCGCGTCGGGGTGGAGGCGTGGCGCAGGCGCAGAGAGGCGC
GCCGCGCCGGCGCAGGCGCAGAGACACATGCTACCGCGTCCAGGGGTGGAGGCGTGGCGCAGGCGCAGAG
AGGCGCACCGCGCCGGCGCAGGCGCAGAGACACATGCTAGCGCGTCCAGGGGTGGAGGCGTGGCGCAGGC
GCAGAGACGCAAGCCTACGGGCGGGGGTTGGGGGGGCGTGTGTTGCAGGAGCAAAGTCGCACGGCGCCGG
GCTGGGGCGGGGGGAGGGTGGCGCCGTGCACGCGCAGAAACTCACGTCACGGTGGCGCGGCGCAGAGACG
GGTAGAACCTCAGTAATCCGAAAAGCCGGGATCGACCGCCCCTTGCTTGCAGCCGGGCACTACAGGACCC
GCTTGCTCACGGTGCTGTGCCAGGGCGCCCCCTGCTGGCGACTAGGGCAACTGCAGGGCTCTCTTGCTTA
GAGTGGTGGCCAGCGCCCCCTGCTGGCGCCGGGGCACTGCAGGGCCCTCTTGCTTACTGTATAGTGGTGG
CACGCCGCCTGCTGGCAGCTAGGGACATTGCAGGGTCCTCTTGCTCAAGGTGTAGTGGCAGCACGCCCAC
CTGCTGGCAGCTGGGGACACTGCCGGGCCCTCTTGCTCCAACAGTACTGGCGGATTATAGGGAAACACCC
GGAGCATATGCTGTTTGGTCTCAGTAGACTCCTAAATATGGGATTCCTGGGTTTAAAAGTAAAAAATAAA
TATGTTTAATTTGTGAACTGATTACCATCAGAATTGTACTGTTCTGTATCCCACCAGCAATGTCTAGGAA
TGCCTGTTTCTCCACAAAGTGTTTACTTTTGGATTTTTGCCAGTCTAACAGGTGAAGCCCTGGAGATTCT
TATTAGTGATTTGGGCTGGGGCCTGGCCATGTGTATTTTTTTAAATTTCCACTGATGATTTTGCTGCATG
GCCGGTGTTGAGAATGACTGCGCAAATTTGCCGGATTTCCTTTGCTGTTCCTGCATGTAGTTTAAACGAG
ATTGCCAGCACCGGGTATCATTCACCATTTTTCTTTTCGTTAACTTGCCGTCAGCCTTTTCTTTGACCTC
TTCTTTCTGTTCATGTGTATTTGCTGTCTCTTAGCCCAGACTTCCCGTGTCCTTTCCACCGGGCCTTTGA
```

Рис. 1.2. Типичный вид фрагмента данных о геномных последовательностях нуклеотидов в одиночной нити ДНК из международного банка генетических данных GenBank.

Получаемые автором результаты анализа таких геномных текстов говорят об алгебраически закономерном характере ансамблей структур генетической системы и позволяют развивать алгебраическую и квантовую биологию. Выявляемые алгебраические особенности генетических структур представляются полезными для совершенствования систем искусственного интеллекта.

Следует отметить, что генетика и алгебраическая биология фактически начались с открытия Менделя в опытах по скрещиванию организмов того, что наследование их признаков происходит по алгебраическим правилам, несмотря на чудовищную сложность и разнообразие молекулярного состава тел. Эти алгебраические правила оказались связанными с понятием бинарной оппозиции доминантных и рецессивных факторов наследования, введенным Менделем.

2. Бинарные-оппозиции в алфавитах ДНК и матрицы

При всем поразительном многообразии живых организмов генетическая информация в них записывается единообразно в виде длинных текстов из четырех нуклеотидов (молекулярных букв.). В молекулах ДНК такими буквами выступают аденин A, цитозин C, гуанин G и тимин T, образующие ДНК-алфавит нуклеотидов (в РНК вместо тимина участвует урацил U). ДНК-тексты иногда называют энциклопедией жизни. В двойной спирали ДНК нуклеотиды C и G всегда стоят напротив друг друга и потому называются комплементарной парой. То же самое относится к паре компплементарных нуклеотидов A и T. При этом A и T связаны двумя водородными связями, а C и G - тремя водородными связями (Рис. 2.1). В генетическом кодировании также принимают участие ДНК-алфавиты 16 дуплетов (типа AA, AT, …, GG), 64 триплетов (типа AAA, AAT, …) и т.д.

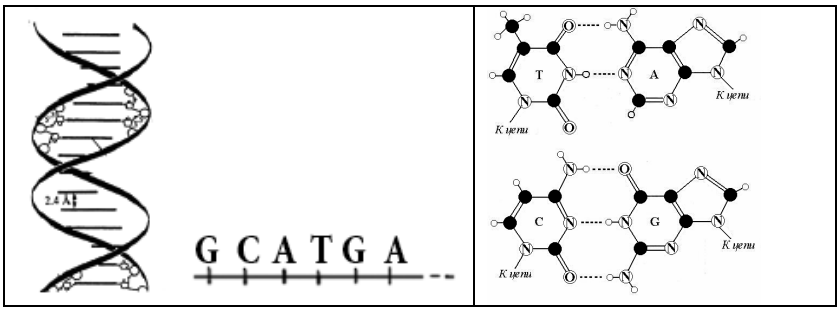


Рис. 2.1. Двойная спираль ДНК и комплементарные пары нуклеооснований A, T, C и G в ней.

Наука не знает, почему алфавит нуклеотидов ДНК состоит всего из этих четырех букв A, T, C, G, но знает, что он является носителем симметрической системы бинарно-оппозиционных признаков (индикаторов):

1) два из этих нуклеотидов являются пуринами (A и G), имея в молекуле два кольца, а два другие (C и T) – пиримидинами, имея одно кольцо. С позиций этих оппозиционных индикаторов C = T = 1, A = G = 0 ;

2) два из этих нуклеотидов являются кето-молекулами (T и G), а два другие (C и A) – амино-молекулами, что дает C = A = 1, T = G = 0.

При этом каждый из 4 нуклеотидов однозначно определяется названной парой бинарных индикаторов.Учитывая это, удобно представить ДНК-алфавиты 4 букв, 16 дуплетов и 64 триплетов в форме квадратных таблиц, столбцы которых нумеруются бинарными индикаторами “пиримидин или пурин” (C=T=0, A=G=1), а строки – бинарными индикаторами “амино или кето” (C=A=0, T=G=1). В таких таблицах все буквы, дуплеты и триплеты автоматически занимают свое строго индивидуальное место (Рис. 2.2).

	0	1
0	C	A
1	T	G

;

	00	01	10	11
00	CC	CA	AC	AA
01	CT	CG	AT	AG
10	TC	TA	GC	GA
11	TT	TG	GT	GG

	000	001	010	011	100	101	110	111
000	CCC	CCA	CAC	CAA	ACC	ACA	AAC	AAA
001	CCT	CCG	CAT	CAG	ACT	ACG	AAT	AAG
010	CTC	CTA	CGC	CGA	ATC	ATA	AGC	AGA
011	CTT	CTG	CGT	CGG	ATT	ATG	AGT	AGG
100	TCC	TCA	TAC	TAA	GCC	GCA	GAC	GAA
101	TCT	TCG	TAT	TAG	GCT	GCG	GAT	GAG
110	TTC	TTA	TGC	TGA	GTC	GTA	GGC	GGA
111	TTT	TTG	TGT	TGG	GTT	GTG	GGT	GGG

Рис. 2.2. Расположение ДНК-алфавитов 4 нуклеотидов, 16 дуплетов и 64 триплетов в таблицах, построенных на учете бинарно-оппозиционных молекулярных признаков в алфавите нуклеотидов C, A, T, G.

Данные 3 таблицы являются не просто таблицами, но они являются начальными членами единого тензорное семейство матриц $[C, A; T, G]^{(n)}$, где символ (n) обозначает тензорную степень и n = 1, 2, 3, 4, … . Вторая и третья тензорная (кронекеровская) степень (2*2)-матрицы [C, A; T, G] автоматически дают эту (4*4)-матрицу алфавита 16 дуплетов и эту (8*8)-матрицу алфавита 64 триплетов на Рис. 2.2. Тензорное произведение матриц играет важную роль в квантовой информатике и квантовой механике многокомпонентных систем: пространство состояния многокомпонентной системы является тензорным произведением пространств состояний ее компонент.

Кратко напомним здесь о пионерской статье [Конопельченко, Румер, 1975], где впервые алфавит 4 нуклеотидов ДНК представлен в виде (2*2)-матрицы, которая без каких-либо пояснений возведена во вторую тензорную (кронекеровскую) степень. Это было сделано для получения (4x4)-матрицы 16 дуплетов в целях анализа симметрий в отношениях вырожденности генетического кода, исходя из общего положения о важности исследования симметрий в природных системах. Данная 4-страничная статья обосновала целесообразность выделения в триплете двух первых букв («корня кодона») и ввела понятие «сильных» и «слабых» корней кодонов для демонстрации упорядоченности генетического кода. В этой статье не затрагивался тот факт, что алфавит 4 нуклеотидов наделен системой бинарно-оппозиционных признаков, учет которых ведет к алгоритмическому построению многочисленного тензорного семейства матриц (Рис. 2.2) без необходимости волевого введения исследователем операции тензорного произведения в анализ генетических структур. Данная статья явилась первой публикацией по матричной генетике. Она оставалась без продолжения и должной оценки в течение более чем четверти века, о чем можно только сожалеть. Вместе с нашими публикациями она свидетельствует о приоритете и ведущем положении отечественной науки в перспективной области алгебраической биологии – матричной генетике. Важная роль и отличительные особенности данной статьи описаны в специальном параграфе книги [Петухов, 2008, с. 134-136].

64 триплета кодируют 20 наследуемых аминокислот, из которых состоят белки живых тел, а также стоп-сигналы белкового синтеза (стоп-кодоны). Как же в этой матрице 64 триплетов (Рис. 2.2) расположены кодируемые ими 20 аминокислот и стоп-сигналы белкового синтеза? Это расположение нельзя предсказать, поскольку аминокислоты и нуклеотиды A, C, G, T совершенно различны по строению. Число вариантов расположения аминокислот с их какими-то повторениями для заполнения всей (8*8)-матрицы необозримо: $>>10^{100}$ (для сравнения в физике время существования Вселенной оценивается в 1017 секунд). Будет ли это расположение аминокислот хаотичным или вдруг окажется закономерно симметричным?

Оказывается, что из океана возможностей природа выбрала алгебраически закономерный вариант повторения и расположения аминокислот и стоп-сигналов в этой матрице 64 триплетов, показанный на Рис. 2.3 для случая генетического кода митохондрий позвоночных (этот диалект генетического кода считается наиболее симметричным и древним [Франк-Каменецкий, 1988]); в теории симметрии принято начинать анализ с симметричных паттернов, переходя затем к изучению случаем нарушения их симметрий). Показанная матрица кодируемых аминокислот и стоп-сигналов состоит из пар соседних строк 1-2, 3-4, 5-6, 7-8, идентичных по составу аминокислот и стоп-кодонов.

CCC Pro	CCA Pro	CAC His	CAA Gln	ACC Thr	ACA Thr	AAC Asn	AAA Lys
CCT Pro	CCG Pro	CAT His	CAG Gln	ACT Thr	ACG Thr	AAT Asn	AAG Lys
CTC Leu	CTA Leu	CGC Arg	CGA Arg	ATC Ile	ATA Met	AGC Ser	AGA Stop
CTT Leu	CTG Leu	CGT Arg	CGG Arg	ATT Ile	ATG Met	AGT Ser	AGG Stop
TCC Ser	TCA Ser	TAC Tyr	TAA Stop	GCC Ala	GCA Ala	GAC Asp	GAA Glu
TCT Ser	TCG Ser	TAT Tyr	TAG Stop	GCT Ala	GCG Ala	GAT Asp	GAG Glu
TTC Phe	TTA Leu	TGC Cys	TGA Trp	GTC Val	GTA Val	GGC Gly	GGA Gly
TTT Phe	TTG Leu	TGT Cys	TGG Trp	GTT Val	GTG Val	GGT Gly	GGG Gly

Рис. 2.3. Симметричное расположение 20 аминокислот и стоп-кодонов в матрице кодирующих 64 триплетов. Показан случай генетического кода митохондрий позвоночных. Используются общепринятые краткие обозначения этих аминокислот и стоп-кодонов (Stop). Одинаковым цветом выделены строки 1-2, 3-4, 5-6 и 7-8, идентичные по составу аминокислот и стоп-кодонов.

Эта закономерная связь кодируемых аминокислот и стоп-кодонов с тензорным семейством ДНК-алфавитов *n*-плетов является одним из свидетельств глубокой связи информационной системы генетического кодирования с формализмами квантовой информатики.

3. Гештальт правила процентного состава композитных ДНК-текстов геномов, связанные с *m*-позиционными тетрагруппировками *n*-плетов.

Будем исследовать каждый геномный ДНК-текст как композитный, многослойный текст, в котором каждый *n*-й слой является последовательностью *n*-плетов. Например, в тексте ACCTGTAACG… первый слой является последовательностью членов алфавита 4 нуклеотидов (A-C-C-T-G-…), второй – последовательностью членов алфавита 16 дуплетов (AC-CT-GT-AA-CG-…), третий – последовательностью членов алфавита 64 триплетов (ACC-TGT-AAC-..) и т.д. Тем самым, текст каждого такого слоя написан в своем собственном алфавите, но все *n*-плетные алфавиты слоев тензорно взаимосвязаны через тензорное семейство матриц на Рис. 2.2. В каждом *n*-м слое будем подсчитывать проценты каждого из 4^n видов *n*-плетов и изучать взаимосвязи между ними. Этот авторский подход к анализу ДНК-текста как набора параллельных текстов обнаруживает важные алгебраические закономерности в генетической информатике высших и низших организмов.

Начнем с ДНК-текста хромосомы человека №1, содержащего около 250 миллионов букв C, G, A, T (исходные данные взяты из GenBank: https://www.ncbi.nlm.nih.gov/nuccore/NC_000001.11). Сначала подсчитаем в его первом слое проценты каждого из 4-х видов нуклеотидов C, G, A, T. Получаем %C ≈ 0.2085 , %G ≈ 0.2089, %A ≈ 0.2910, %T ≈ 0.2917. Подстановка этих процентов в соответствующие ячейки описанной матрицы нуклеотидов из Рис. 2.2 дает числовую матрицу на Рис. 3.1.

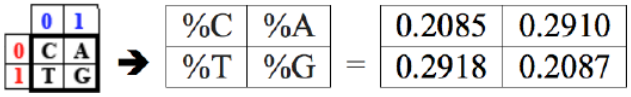


Рис. 3.1. Процентный состав первого – нуклеотидного - слоя ДНК-текста хромосомы человека №1.

Затем в его втором - дуплетном – слое подсчитаем проценты каждого из 16 дуплетов и вставим их в соответствующие ячейки матрицы 16 дуплетов $[C, A; T, G]^{(2)}$ названного тензорного семейства (Рис. 3.2).

%CC	%CA	%AC	%AA	=	0.05409	0.07274	0.05033	0.09504
%CT	%CG	%AT	%AG		0.07134	0.01031	0.07429	0.07137
%TC	%TA	%GC	%GA		0.06008	0.06312	0.04402	0.06008
%TT	%TG	%GT	%GG		0.09568	0.07286	0.05046	0.05419

Рис. 3.2. Процентный состав второго – дуплетного - слоя ДНК-текста хромосомы человека №1.

Рис. 3.3 показывает подсчитанные проценты каждого из 64 триплетов и 256 тетраплетов соответственно в третьем (триплетном) и четвертом (тетраплетном) слоях того же ДНК-текста.

0.0138	0.0188	0.0152	0.0186	0.0118	0.0198	0.0145	0.0369
0.0185	0.0029	0.0179	0.0210	0.0162	0.0025	0.0238	0.0199
0.0176	0.0127	0.0025	0.0023	0.0132	0.0194	0.0144	0.0224
0.0201	0.0209	0.0026	0.0029	0.0239	0.0178	0.0161	0.0185
0.0159	0.0196	0.0110	0.0199	0.0125	0.0146	0.0096	0.0196
0.0223	0.0023	0.0194	0.0128	0.0144	0.0025	0.0133	0.0176
0.0197	0.0198	0.0146	0.0195	0.0096	0.0112	0.0126	0.0160
0.0372	0.0188	0.0199	0.0190	0.0145	0.0153	0.0119	0.0138

.0033	.0055	.0042	.0044	.0040	.0056	.0032	.0070	.0030	.0042	.0040	.0053	.0032	.0059	.0055	.0149
.0041	.0010	.0044	.0058	.0047	.0010	.0040	.0044	.0041	.0005	.0051	.0054	.0047	.0006	.0095	.0071
.0050	.0029	.0008	.0006	.0036	.0039	.0049	.0059	.0037	.0032	.0006	.0006	.0040	.0070	.0037	.0066
.0048	.0058	.0006	.0009	.0057	.0047	.0045	.0058	.0049	.0044	.0007	.0007	.0071	.0057	.0049	.0047
.0052	.0057	.0027	.0038	.0010	.0006	.0003	.0006	.0033	.0045	.0032	.0063	.0042	.0049	.0039	.0078
.0058	.0009	.0035	.0028	.0007	.0003	.0005	.0008	.0049	.0005	.0064	.0035	.0046	.0007	.0049	.0059
.0048	.0036	.0046	.0051	.0005	.0004	.0008	.0006	.0047	.0056	.0033	.0050	.0031	.0037	.0047	.0057
.0073	.0044	.0053	.0058	.0006	.0010	.0005	.0010	.0096	.0040	.0051	.0044	.0046	.0047	.0040	.0041
.0046	.0049	.0042	.0051	.0024	.0045	.0028	.0078	.0030	.0041	.0029	.0038	.0023	.0037	.0029	.0073
.0057	.0006	.0051	.0052	.0037	.0004	.0056	.0036	.0047	.0008	.0033	.0046	.0030	.0005	.0046	.0048
.0057	.0040	.0005	.0006	.0030	.0055	.0026	.0041	.0032	.0027	.0006	.0005	.0025	.0031	.0032	.0057
.0068	.0058	.0006	.0006	.0070	.0039	.0032	.0029	.0037	.0048	.0006	.0008	.0041	.0036	.0036	.0051
.0051	.0063	.0034	.0063	.0041	.0049	.0031	.0061	.0023	.0031	.0017	.0035	.0031	.0042	.0023	.0052
.0077	.0006	.0063	.0038	.0049	.0006	.0045	.0057	.0038	.0004	.0032	.0027	.0042	.0010	.0034	.0052
.0073	.0078	.0038	.0051	.0037	.0046	.0042	.0051	.0030	.0028	.0029	.0042	.0023	.0024	.0030	.0046
.0150	.0072	.0055	.0045	.0059	.0057	.0043	.0054	.0055	.0032	.0040	.0042	.0032	.0039	.0030	.0033

Рис. 3.3. Проценты каждого из 64 триплетов и 256 тетраплетов соответственно в третьем (триплетном) и четвертом (тетраплетном) слоях ДНК-текста хромосомы человека №1.

На первый взгляд, показанные множества процентов в каждом слое этого

ДНК-текста вполне хаотичны. При этом процент *n*-плетов во всех слоях зависит от очередности букв. Например, процент дуплетов CG и GC одинакового буквенного состава отличается в несколько раз: %CG = 0.0103, а %GC = 0.0440 (Рис. 3.2). Отметим также, что тензорные степени матрицы процентов нуклеотидов $[\%C, \%A; \%T, \%G]^{(n)} = [0.2085. 0.2910; 0.2918, 0.2087]^{(n)}$ дают такие формальные проценты *n*-плетов, которые совершенно не совпадают с реальными процентами *n*-плетов в *n*-плетных слоях данного ДНК-текста.

Но неожиданно автором обнаружены тетраблочные закономерности в этих множествах процентов дуплетов, триплетов, тетраплетов, и т.д. в слоях геномных ДНК-текстов. Будем исследовать в каждом слое ДНК-текста суммы процентов *n*-плетов, принадлежащих к определенным группировкам. Введем обозначения для этих сумм:

- сумму процентов всех 4 дуплетов в дуплетом слое, начинающихся с буквы C, обозначим Σ%CN = %CC+%CA+%CT+%CG (N - любая из букв C, A, T, G);
- сумму процентов всех 4 дуплетов, имеющих букву C на второй позиции, обозначим Σ%NC = %CC+%AC+%TC+%GC;
- суммы процентов в триплетном слое всех 16 триплетов, содержащих букву C на первой, второй и третьей позициях обозначим соответственно символами Σ%CNN, Σ%NCN и Σ%NNC;
- суммы процентов всех 64 тетраплетов с буквой C на первой, второй, третьей и четвертой позициях обозначим соответственно Σ%CNNN, Σ%NCNN, Σ%NNCN и Σ%NNNC.

Посчитывая такие суммы процентов, получаем удивительную таблицу их взаимосвязи и симметрии (Рис. 3.4).

%C ≈ **0.2085**	%G ≈ **0.2087**	%A ≈ **0.2910**	%T ≈ **0.2918**
Σ%CN ≈ **0.2085**	Σ%GN ≈ **0.2088**	Σ%AN ≈ **0.2910**	Σ%TN ≈ **0.2917**
Σ%NC ≈ **0.2085**	Σ%NG ≈ **0.2087**	Σ%NA ≈ **0.2910**	Σ%NT ≈ **0.2918**
Σ%CNN ≈ **0.2084**	Σ%GNN ≈ **0.2088**	Σ%ANN ≈ **0.2910**	Σ%TNN ≈ **0.2917**
Σ%NCN ≈ **0.2085**	Σ%NGN ≈ **0.2088**	Σ%NAN ≈ **0.2910**	Σ%NTN ≈ **0.2917**
Σ%NNC ≈ **0.2085**	Σ%NNG ≈ **0.2087**	Σ%NNA ≈ **0.2910**	Σ%NNT ≈ **0.2918**
Σ%CNNN ≈ **0.2085**	Σ%GNNN ≈ **0.2088**	Σ%ANNN ≈ **0.2910**	Σ%TNNN ≈ **0.2917**
Σ%NCNN ≈ **0.2085**	Σ%NGNN ≈ **0.2087**	Σ%NANN ≈ **0.2910**	Σ%NTNN ≈ **0.2918**
Σ%NNCN ≈ **0.2085**	Σ%NNGN ≈ **0.2088**	Σ%NNAN ≈ **0.2910**	Σ%NNTN ≈ **0.2918**
Σ%NNNC ≈ **0.2085**	Σ%NNNG ≈ **0.2087**	Σ%NNNA ≈ **0.2910**	Σ%NNNT ≈ **0.2918**

Рис. 3.4. Взаимосвязь величин сумм процентов в *m*-позиционных тетра-группировках *n*-плетов из разных *n*-плетных слоев ДНК-текста хромосомы человека №1.

В каждом из 4-х столбцов, связанном с одним из нуклеотидов C, G, A или T, одно и то же суммарное число процентов повторяется во всех его строках и оно равно проценту этого нуклеотида %C, %G, %A, %T с точностью до третьего или четвертого знака после запятой (Рис. 3.4). Зная процент нуклеотида в геномной ДНК, можно с высокой точностью предсказать суммы процентов 4 дуплетов, 16 триплетов, 64 тетраплетов с этим нуклеотидом на фиксированных позициях в дуплетном, триплетном и тетраплетном слоях ДНК-текста.

Это относительное постоянство сумм процентов мы называем **гештальт-правилом процентов в геномах** (для указанных тетрагруппировок *n*-плетов):

- Все n слоев длинного ДНК-текста, каждый из которых состоит из последовательности 4^n видов n-плетов, имеют примерно одинаковую сумму процентов всех тех n-плетов, которые содержат определяющий нуклеотид (C, G, A или T) в фиксированной позиции m ($m \leq n$).
  Здесь n = 1, 2, 3, 4,… много меньше длины геномного текста.

В разных геномах проценты нуклеотидов %C, %G, %A, %T, которые определяют величины сумм процентов в тетрагруппировках n-плетов, могут сильно различаться. Но это гештальт-правило взаимозависимости сумм процентов n-плетов в n-плетных слоях геномных ДНК-текстов неизменно повторяется на всех протестированных автором геномах высших и низших организмов.

Это Гештальт-правило сумм процентов можно переформулировать в виде геномного принципа запрета:

- сумма процентов всех тех n-плетов, которые содержат атрибутивный нуклеотид (C, G, A или T) на фиксированной позиции m ($m \leq n$), не могут быть существенно различными в разных n-плетных слоях геномного ДНК-текста (здесь n = 1, 2, 3, 4, ... , но много меньше длины ДНК-текста).

Каждая из 16 строк 4 столбцов (Рис. 3.4), отражающих феноменологическое гештальт правило сумм процентов в геномах, представляет в соответствующем слое геномного ДНК-текста определенную m-позиционную тетрагруппировку n-плетов, то есть n-плетов, в которых атрибутивный нуклеотид стоит на позиции m ($m \leq n$). Принимая во внимание принадлежность каждого n-плета к одному из четырех столбцов, то есть к одной из четырех C-, G-, A- и T-групировок, введем для наглядности следующие символьные обозначения для каждого из n-плетов:

- n-плеты первого столбца (C-столбца) обозначим символом ♠;
- n-плеты второго столбца (G-столбца) - символом ♣;
- n-плеты третьего столбца (A-столбца) - символом ♥;
- n-плеты четвертого столбца (T-столбца) – символом ♦.

Тогда данное гештальт-правило для этих тетрагруппировок n-плетов геномов выражается так:

$$\Sigma\%\spadesuit \approx \text{const},\ \Sigma\%\clubsuit \approx \text{const},\ \Sigma\%\heartsuit \approx \text{const},\ \Sigma\%\diamondsuit \approx \text{const} \qquad (3.1)$$

При этом в одиночной нити ДНК выполняются примерные равенства $\Sigma\%\spadesuit \approx \Sigma\%\clubsuit$ и $\Sigma\%\heartsuit \approx \Sigma\%\diamondsuit$ (в двойной спирали ДНК данные два равенства выполняются точно). В следующих разделах статьи эти карточные символы будут использованы для наглядной демонстрации алгебраически закономерного расположения членов данных тетрагруппировок в матрицах тензорного семейства $[C, A; T, G]^{(n)}$.

Если в ДНК-текстах геномов брать только каждую κ-ю букву (κ = 2, 3,…, 10,…, 50,…), то образуются новые укороченные ДНК-тексты с новыми очередностями букв A, T, C, G (Рис. 3.5). Оказывается, что для каждого такого нового текста, называемого эпи-цепью ДНК порядка k, проценты отдельных n-плетов различны, но суммы процентов в тетрагруппировках исходной ДНК и ее эпи-цепей аналогичны [Petoukhov, 2021a]. Это свидетельствует о фрактало-подобном строении длинных ДНК-текстов.

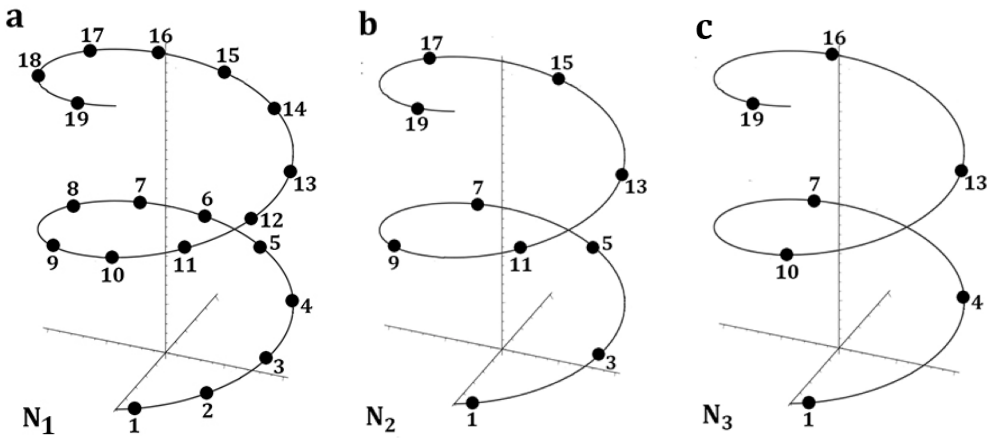


Рис. 3.5. Схематическое изображение одиночной нити ДНК и ее начальных эпи-цепей с нумерованными нуклеотидами, обозначенными черными кружками. **a**, последовательность N_1 нумерованных нуклеотидов в нити ДНК; **b**, эпи-цепь второго порядка N_2 , содержащая нуклеотиды с номерами 1-3-5-7-…; **c**, эпи-цепь третьего порядка N_3 нуклеотидов с номерами 1-4-7-10-… .

Например, для ДНК-текстов хромосомы человека №1 и ее эпи-цепей при разрежении k = 2, 3, 4 имеем в их последовательностях дуплетов различные проценты 16 дуплетов (Рис. 3.6). Но суммы процентов в тетрагруппировках дуплетов во всех этих разных текстах равны соответственно %C≈0.2085, %G≈0.2087, %A≈0.2910, %T≈0.2918 с убедительной точностью независимо от различия процентов отдельных дуплетов.

Исходная ДНК первой хромосомы человека:

%CC	%CA	%AC	%AA
%CT	%CG	%AT	%AG
%TC	%TA	%GC	%GA
%TT	%TG	%GT	%GG

≈

0.05409	0.07274	0.05033	0.09504
0.07134	0.01031	0.07429	0.07137
0.06008	0.06312	0.04402	0.06008
0.09568	0.07286	0.05046	0.05419

Эпи-цепь при k = 2

0.0492	0.0524	0.0538	0.0985
0.0591	0.0478	0.0800	0.0588
0.0612	0.0788	0.0443	0.0613
0.0988	0.0530	0.0539	0.0492

Эпи-цепь при k = 3

0.0487	0.0562	0.0579	0.0912
0.0571	0.0464	0.0849	0.0570
0.0608	0.0826	0.0411	0.061
0.0918	0.0565	0.0578	0.0489

Эпи-цепь при k = 4

0.0487	0.0558	0.0574	0.0940
0.0566	0.0475	0.0831	0.0565
0.0579	0.0835	0.0445	0.0579
0.0943	0.0561	0.0577	0.0487

Рис. 3.6. Проценты 16 дуплетов в ДНК хромосомы человека №1 и в ее эпи-цепях при разряжении k = 2, 3, 4.

Например, %CG = 0.0103 в исходной полной нити ДНК; в ее эпи-цепи второго порядка %CG = 0.0478; в эпи-цепи третьего порядка %CG = 0.0464; в эпи-цепи четвертого порядка %CG = 0.0475. Но в каждой из четырех представленных процентных матриц реализуются равенства сумм процентов, которые удобно представить как равенство координат 4-мерного вектора:

$$[\%A, \%T, \%C, \%G] \approx [\Sigma\%AN, \Sigma\%TN, \Sigma\%CN, \Sigma\%GN] \approx$$
$$[\Sigma\%NA, \Sigma\%NT, \Sigma\%NC, \Sigma\%NG] \approx [0.2910, 0.2918, 0.2085, 0.2087] \quad (3.2)$$

Для сравнительного анализа ситуации с процентами вернемся к тензорному семейству матриц $[C, A; T, G]^{(n)}$ (Рис. 2.2) и рассмотрим формальные (или модельные) проценты *n*-плетов в тензорном семействе процентных матриц $[\%C, \%A; \%T, \%G]^{(n)} \approx [0{,}2085, 0{,}2910; 0{,}2917, 0{,}2089]^{(n)}$. Рис. 3.7 и 3.8 показывают полученные матрицы для дуплетов и триплетов, величины процентов в которых существенно отличаются от реальных процентов *n*-плетов, показанных выше на рис. 3.2 и 3.3.

Но неожиданно гештальт-правило постоянства сумм процентов выполняется не только для реальных процентов, но также для этих модельных процентов *n*-плетов:

$$[\%A, \%T, \%C, \%G] \approx [\Sigma\%AN, \Sigma\%TN, \Sigma\%CN, \Sigma\%GN] \approx$$
$$[\Sigma\%NA, \Sigma\%NT, \Sigma\%NC, \Sigma\%NG] \approx$$
$$[\Sigma\%ANN, \Sigma\%TNN, \Sigma\%CNN, \Sigma\%GNN] \approx$$
$$[\Sigma\%NAN, \Sigma\%NTN, \Sigma\%NCN, \Sigma\%NGN] \approx$$
$$[\Sigma\%NNA, \Sigma\%NNT, \Sigma\%NNC, \Sigma\%NNG] \approx$$
$$[0.2910, 0.2918, 0.2085, 0.2087]. \quad (3.3)$$

%CC	%CA	%AC	%AA
%CT	%CG	%AT	%AG
%TC	%TA	%GC	%GA
%TT	%TG	%GT	%GG

=

0.0435	0.0607	0.0607	0.0847
0.0608	0.0435	0.0849	0.0607
0.0608	0.0849	0.0435	0.0607
0.0851	0.0609	0.0609	0.0436

Рис. 3.7. Формальные (или модельные) проценты 16 дуплетов в матрице, являющейся второй тензорной степенью матрицы [0.2085, 0.2910; 0.2917, 0.2089] процентов нуклеотидов в ДНК-тексте хромосомы человека №1. Все величины округлены до четвертого знака после запятой.

%CCC	%CCA	%CAC	%CAA	%ACC	%ACA	%AAC	%AAA
%CCT	%CCG	%CAT	%CAG	%ACT	%ACG	%AAT	%AAG
%CTC	%CTA	%CGC	%CGA	%ATC	%ATA	%AGC	%AGA
%CTT	%CTG	%CGT	%CGG	%ATT	%ATG	%AGT	%AGG
%TCC	%TCA	%TAC	%TAA	%GCC	%GCA	%GAC	%GAA
%TCT	%TCG	%TAT	%TAG	%GCT	%GCG	%GAT	%GAG
%TTC	%TTA	%TGC	%TGA	%GTC	%GTA	%GGC	%GGA
%TTT	%TTG	%TGT	%TGG	%GTT	%GTG	%GGT	%GGG

=

.00906382	.01265035	.01265035	.01765605	.01265035	.01765605	.01765605	.02464250
.01268315	.00907448	.01770183	.01266522	.01770183	.01266522	.02470639	.01767681
.01268315	.01770183	.00907448	.01266522	.01770183	.02470639	.01266522	.01767681
.01774773	.01269806	.01269806	.00908514	.02477045	.01772264	.01772264	.01268011
.01268315	.01770183	.01770183	.02470639	.00907448	.01266522	.01266522	.01767681
.01774773	.01269806	.02477045	.01772264	.01269806	.00908514	.01772264	.01268011
.01774773	.02477045	.01269806	.01772264	.01269806	.01772264	.00908514	.01268011
.02483467	.01776859	.01776859	.01271298	.01776859	.01271298	.01271298	.00909582

Рис. 3.8. Формальные (или модельные) проценты 64 триплетов в матрице, являющейся третьей тензорной степенью матрицы [0.2085, 0.2910; 0.2917, 0.2089] процентов нуклеотидов в ДНК-тексте хромосомы человека №1.

Следует особенно отметить, что в случае тензорного семейства процентных матриц $[\%C, \%A; \%T, \%G]^{(n)}$ Гештальт правило выполняется с абсолютной точностью. Например, если %A = 0.291001313, %T = 0.291755765, %C = 0.208498924, %G = 0.208743998, то мы имеем точные равенства рассматриваемых сумм процентов:

$$
\begin{aligned}
&[\%A, \%T, \%C, \%G] = [\Sigma\%AN, \Sigma\%TN, \Sigma\%CN, \Sigma\%GN] = \\
&[\Sigma\%NA, \Sigma\%NT, \Sigma\%NC, \Sigma\%NG] = \\
&[\Sigma\%ANN, \Sigma\%TNN, \Sigma\%CNN, \Sigma\%GNN] = \\
&[\Sigma\%NAN, \Sigma\%NTN, \Sigma\%NCN, \Sigma\%NGN] = \\
&[\Sigma\%NNA, \Sigma\%NNT, \Sigma\%NNC, \Sigma\%NNG] = \\
&[\Sigma\%ANNN, \Sigma\%TNNN, \Sigma\%CNNN, \Sigma\%GNNN] = \\
&[\Sigma\%NANN, \Sigma\%NTNN, \Sigma\%NCNN, \Sigma\%NGNN] = \\
&[\Sigma\%NNAN, \Sigma\%NNTN, \Sigma\%NNCN, \Sigma\%NNGN] = \\
&[\Sigma\%NNNA, \Sigma\%NNNT, \Sigma\%NNNC, \Sigma\%NNNG] = \\
&[0.291001313, 0.291755765, 0.208498924, 0.208743998]. \qquad (3.4)
\end{aligned}
$$

Это свидетельствует о том, что фундаментальные генетические явления, отраженные в сформулированном правиле гештальта, связаны с алгебраической операцией тензорного произведения, которая так важна в квантовой механике и квантовой информатике. Соответственно, тензорное семейство процентных матриц $[\%C, \%A; \%T, \%G]^{(n)}$ и их процентное содержание можно рассматривать в каждом конкретном случае как некий эталон сравнения при анализе длинных ДНК-текстов. Ниже мы покажем, что разница между тетрагруппировками реальных процентов и эталонных процентов связана с унитарными операторами. Процентные значения *n*-плетов в тензорном семействе матриц $[\%C, \%A; \%T, \%G]^{(n)}$ условно назовем эталонными процентами n-плетов в анализируемом длинном ДНК-тексте.

Кратко упомянем еще о следующем. Выше были представлены правила многослойных ДНК-текстов геномов, связанные с постоянством сумм процентов *n*-плетов в различных *m*-позиционных тетрагруппировках, то есть группировках, в которых роль атрибутивного элемента группировки выполнял один из 4 нуклеотидов C, G, A, T. Оказывается, что аналогичные геномные гештальт правила выполняются для тетрагруппировок, в которых роль атрибутивного элемента группировки *n*-плетов выполняет один из 16 дуплетов, или один из 64 триплетов, и т.д. (*mm*-, *mmm*-тетрагруппировки). В этом случае мы имеем 16 тетрагруппировок типа CA-тетрагруппировка, 64 тетра-группировки типа CAT-тетрагруппировка, и т.д. (см. подробности в препринте [Petoukhov, 2021a].

Например, для хромосомы человека №1 имеем постоянства сумм процентов для всех *16 mm*-тетрагруппировок *n*-плетов, показанные на Рис. 3.8.

%AA ≈ Σ%AAN ≈ Σ%NAA ≈ Σ%AANN ≈ Σ%NAAN ≈ Σ%NNAA ≈ 0.095
%AT ≈ Σ%ATN ≈ Σ%NAT ≈ Σ%ATNN ≈ Σ%NATN ≈ Σ%NNAT ≈ 0.074
%AC ≈ Σ%ACN ≈ Σ%NAC ≈ Σ%ACNN ≈ Σ%NACN ≈ Σ%NNAC ≈ 0.050
%AG ≈ Σ%AGN ≈ Σ%NAG ≈ Σ%AGNN ≈ Σ%NAGN ≈ Σ%NNAG ≈ 0.071
%TA ≈ Σ%TAN ≈ Σ%NTA ≈ Σ%TANN ≈ Σ%NTAN ≈ Σ%NNTA ≈ 0.063
%TT ≈ Σ%TTN ≈ Σ%NTT ≈ Σ%TTNN ≈ Σ%NTTN ≈ Σ%NNTT ≈ 0.096
%TC ≈ Σ%TCN ≈ Σ%NTC ≈ Σ%TCNN ≈ Σ%NTCN ≈ Σ%NNTC ≈ 0.060
%TG ≈ Σ%TGN ≈ Σ%NTG ≈ Σ%TGNN ≈ Σ%NTGN ≈ Σ%NNTG ≈ 0.073
%CA ≈ Σ%CAN ≈ Σ%NCA ≈ Σ%CANN ≈ Σ%NCAN ≈ Σ%NNCA ≈ 0.073
%CT ≈ Σ%CTN ≈ Σ%NCT ≈ Σ%CTNN ≈ Σ%NCTN ≈ Σ%NNCT ≈ 0.071
%CC ≈ Σ%CCN ≈ Σ%NCC ≈ Σ%CCNN ≈ Σ%NCCN ≈ Σ%NNCC ≈ 0.054
%CG ≈ Σ%CGN ≈ Σ%NCG ≈ Σ%CGNN ≈ Σ%NCGN ≈ Σ%NNCG ≈ 0.010
%GA ≈ Σ%GAN ≈ Σ%NGA ≈ Σ%GANN ≈ Σ%NGAN ≈ Σ%NNGA ≈ 0.060
%GT ≈ Σ%GTN ≈ Σ%NGT ≈ Σ%GTNN ≈ Σ%NGTN ≈ Σ%NNGT ≈ 0.050
%GC ≈ Σ%GCN ≈ Σ%NGC ≈ Σ%GCNN ≈ Σ%NGCN ≈ Σ%NNGC ≈ 0.044
%GG ≈ Σ%GGN ≈ Σ%NGG ≈ Σ%GGNN ≈ Σ%NGGN ≈ Σ%NNGG ≈ 0.054

Рис. 3.8. Постоянство процентных сумм для *mm*-позиционных тетрагруппировок, относящихся к 64 триплетам и 256 тетраплетам в соответствующих триплетных и тетраплетных слоях ДНК-текста хромосомы человека №1. Каждая из этих тетрагруппировок определяется одним из 16 дублетов в качестве атрибутивного позиционного элемента, расположенного в начале, в середине или в конце *n*-плетов. Численные значения процентных сумм рассчитаны на основе данных о процентах отдельных *n*-плетов на Рис. 3.2 и 3.3.

Рис. 3.9 показывает другой пример: постоянство сумм процентов для всех 64 *mmm*-тетрагруппировок *n*-плетов той же ДНК.

%CCC ≈ Σ%CCCN ≈ Σ%NCCC ≈ 0.014	%GAA ≈ Σ%GAAN ≈ Σ%NGAA ≈ 0.020
%CCG ≈ Σ%CCGN ≈ Σ%NCCG ≈ 0.003	%GAT ≈ Σ%GATN ≈ Σ%NGAT ≈ 0.013
%CGC ≈ Σ%CGCN ≈ Σ%NCGC ≈ 0.003	%GTA ≈ Σ%GTAN ≈ Σ%NGTA ≈ 0.011
%CGG ≈ Σ%CGGN ≈ Σ%NCGG ≈ 0.003	%GTT ≈ Σ%GTTN ≈ Σ%NGTT ≈ 0.014
%GCC ≈ Σ%GCCN ≈ Σ%NGCC ≈ 0.013	%ACC ≈ Σ%ACCN ≈ Σ%NACC ≈ 0.012
%GCG ≈ Σ%GCGN ≈ Σ%NGCG ≈ 0.003	%ACG ≈ Σ%ACGN ≈ Σ%NACG ≈ 0.003
%GGC ≈ Σ%GGCN ≈ Σ%NGGC ≈ 0.013	%AGC ≈ Σ%AGCN ≈ Σ%NAGC ≈ 0.014
%GGG ≈ Σ%GGGN ≈ Σ%NGGG ≈ 0.014	%AGG ≈ Σ%AGGN ≈ Σ%NAGG ≈ 0.018
%CCA ≈ Σ%CCAN ≈ Σ%NCCA ≈ 0.019	%TCC ≈ Σ%TCCN ≈ Σ%NTCC ≈ 0.016
%CCT ≈ Σ%CCTN ≈ Σ%NCCT ≈ 0.019	%TCG ≈ Σ%TCGN ≈ Σ%NTCG ≈ 0.002
%CGA ≈ Σ%CGAN ≈ Σ%NCGA ≈ 0.002	%TGC ≈ Σ%TGCN ≈ Σ%NTGC ≈ 0.015
%CGT ≈ Σ%CGTN ≈ Σ%NCGT ≈ 0.003	%TGG ≈ Σ%TGGN ≈ Σ%NTGG ≈ 0.019
%GCA ≈ Σ%GCAN ≈ Σ%NGCA ≈ 0.015	%ACA ≈ Σ%ACAN ≈ Σ%NACA ≈ 0.020
%GCT ≈ Σ%GCTN ≈ Σ%NGCT ≈ 0.014	%ACT ≈ Σ%ACTN ≈ Σ%NACT ≈ 0.016
%GGA ≈ Σ%GGAN ≈ Σ%NGGA ≈ 0.016	%AGA ≈ Σ%AGAN ≈ Σ%NAGA ≈ 0.022
%GGT ≈ Σ%GGTN ≈ Σ%NGGT ≈ 0.012	%AGT ≈ Σ%AGTN ≈ Σ%NAGT ≈ 0.016
%CAC ≈ Σ%CACN ≈ Σ%NCAC ≈ 0.015	%TCA ≈ Σ%TCAN ≈ Σ%NTCA ≈ 0.020
%CAG ≈ Σ%CAGN ≈ Σ%NCAG ≈ 0.021	%TCT ≈ Σ%TCTN ≈ Σ%NTCT ≈ 0.022
%CTC ≈ Σ%CTCN ≈ Σ%NCTC ≈ 0.018	%TGA ≈ Σ%TGAN ≈ Σ%NTGA ≈ 0.019
%CTG ≈ Σ%CTGN ≈ Σ%NCTG ≈ 0.021	%TGT ≈ Σ%TGTN ≈ Σ%NTGT ≈ 0.020
%GAC ≈ Σ%GACN ≈ Σ%NGAC ≈ 0.010	%AAC ≈ Σ%AACN ≈ Σ%NAAC ≈ 0.014
%GAG ≈ Σ%GAGN ≈ Σ%NGAG ≈ 0.018	%AAG ≈ Σ%AAGN ≈ Σ%NAAG ≈ 0.020
%GTC ≈ Σ%GTCN ≈ Σ%NGTC ≈ 0.010	%ATC ≈ Σ%ATCN ≈ Σ%NATC ≈ 0.013
%GTG ≈ Σ%GTGN ≈ Σ%NGTG ≈ 0.015	%ATG ≈ Σ%ATGN ≈ Σ%NATG ≈ 0.018
%CAA ≈ Σ%CAAN ≈ Σ%NCAA ≈ 0.019	%TAC ≈ Σ%TACN ≈ Σ%NTAC ≈ 0.011
%CAT ≈ Σ%CATN ≈ Σ%NCAT ≈ 0.018	%TAG ≈ Σ%TAGN ≈ Σ%NTAG ≈ 0.013
%CTA ≈ Σ%CTAN ≈ Σ%NCTA ≈ 0.013	%TTC ≈ Σ%TTCN ≈ Σ%NTTC ≈ 0.020
%CTT ≈ Σ%CTTN ≈ Σ%NCTT ≈ 0.020	%TTG ≈ Σ%TTGN ≈ Σ%NTTG ≈ 0.019
%GAA ≈ Σ%GAAN ≈ Σ%NGAA ≈ 0.020	%AAA ≈ Σ%AAAN ≈ Σ%NAAA ≈ 0.037
%GAT ≈ Σ%GATN ≈ Σ%NGAT ≈ 0.013	%AAT ≈ Σ%AATN ≈ Σ%NAAT ≈ 0.024
%GTA ≈ Σ%GTAN ≈ Σ%NGTA ≈ 0.011	%ATA ≈ Σ%ATAN ≈ Σ%NATA ≈ 0.019
%GTT ≈ Σ%GTTN ≈ Σ%NGTT ≈ 0.014	%ATT ≈ Σ%ATTN ≈ Σ%NATT ≈ 0.024
	%TAA ≈ Σ%TAAN ≈ Σ%NTAA ≈ 0.020
	%TAT ≈ Σ%TATN ≈ Σ%NTAT ≈ 0.019
	%TTA ≈ Σ%TTAN ≈ Σ%NTTA ≈ 0.020
	%TTT ≈ Σ%TTTN ≈ Σ%NTTT ≈ 0.037

Рис. 3.9. Постоянство процентных сумм для *mmm*-позиционных тетрагруппировок, относящихся к 256 тетраплетам в тетраплетном слое ДНК-текста хромосомы человека №1. Каждая из этих тетрагруппировок определяется одним из 64 триплетов в качестве атрибутивного позиционного элемента, расположенного в начале или конце тетраплетов. Численные значения процентных сумм рассчитаны на основе данных о процентах отдельных триплетов и тетраплетов на Рис. 3.3.

Представленные данные о закономерном постоянстве сумм процентов *n*-плетов в *mm*- и *mmm*-тетрагруппировках являются частными примерами весьма общего принципа эстафетной (или волновой) гештальт связи вероятностей содержания *n*-плетов в соседних слоях длинных ДНК. Этот принцип заключается в количественном соответствии между каждым видом *n*-плетов P_n в *n*-м слое ДНК и теми двумя четверками (*n*+1)-плетов очередного (*n*+1)-го слоя, для которых он является атрибутивным элементом (то есть четверкой P_nC, P_nG, P_nA, P_nT или четверкой CP_n, GP_n, AP_n, TP_n). Данное соответствие заключается в том, что сумма процентов каждой из двух названных четверок (*n*+1)-плетов в (*n*+1)-м слое примерно равна проценту *n*-плета $P_к$ в предыдущем *n*-м слое ДНК:

$$\%P_n \approx \%P_nC+\%P_nG+\%P_nA+\%P_nT \approx \%CP_n+\%GP_n+\%AP_n+\%TP_n \qquad (3.5)$$

Выполнение этого гештальт принципа (3.5) эстафеты вероятностей для взаимосвязи первого, второго, третьего и четвертого слоев геномных ДНК подтверждается приведенными данными на Рис. 3.4, 3.8 и 3.9. Но этот же принцип гештальт взаимосвязи слоев по их процентному составу *n*-плетов выдерживается для слоев с более высокими номерами *n*, хотя с ростом значений *n* точность выполнения названного равенства уменьшается. Например, обратимся к пятому и шестому слоям ДНК-текста первой хромосомы человека

№1, которые представляют собой соответственно последовательности 49791284 пентаплетов и 41492737 шестиплетов, и сопоставим процент пентаплета ACGTT в пятом слое с суммой процентов шестиплетов ACGTTC, ACGTTG, ACGTTA и ACGTTT в шестом слое. Компьютерный анализ этого ДНК-текста показывает следующее. В его пятом слое содержится 9130 пентаплетов ACGTT, то есть процент этого вида пентаплетов равен 9130 / 49791284 = 0,00018337. В его шестом слое содержатся количества 1679, 1753, 1474 и 2757 шестиплетов ACGTTC, ACGTTG, ACGTTA и ACGTTT соответственно; сумма процентов этой четверки шестиплетов равна (1679+1753+1474+2757) / 41492737 = 0.00018468, то есть примерно равна величине 0,00018337 процента пентаплета ACGTT в предыдущем пятом слое.

Данная эстафетная гештальт связь (3.5) в геномах каждого *n*-плета с четверкой (*n*+1)-плетов очередного (*n*+1)-го слоя напоминает общебиологический процесс учетверения при мейозе, в ходе которого из одной половой клетки образуется четыре аналогичные половые клетки. Эта же эстафетная связь (3.5) отдаленно напоминает принцип Гюйгенса эстафетного характера: каждая точка волнового фронта является вторичным, то есть новым источником сферических волн (его дополнение привело к принципу Гюйгенса-Френеля, который служит основным постулатом волновой теории в физике, описывающим и объясняющим механизм распространения волн).

4. О гештальт-генетике и наследуемых феноменах гештальт-биологии.

Описанная выше относительная независимость сумм процентов *n*-плетов в их блочных тетрагруппировках в слоях геномных ДНК-текстов от значений отдельных слагаемых является аналогом гештальт-образов в психологии. Гештальт-психология изучает свойства мозга формировать целостные образы, относительно независимые от их частных компонент.

Так, музыкальная мелодия распознается нами, даже когда она играется на разных инструментах и в разных частотных диапазонах с измененным частотным составом ее компонентов. То же самое относится к зрительному восприятию картин. Это - наследуемое фундаментальное свойство психики: искать в разрозненном целое. Благодаря способности мыслить гештальтами, вы можете понять предложение, даже если в каждом слове поменять порядок букв и оставить на месте только начальные и конечные. Так, каждый без труда поймет следующую фразу, сильно «мутированную» локальными перестановками: Нпармиер, вы лгкео мжотее почрстеь эту фазру.

Ряд исследователей связывает гештальт-образы восприятия с квантовой запутанностью. Гештальт-психология существует сегодня в синтезе с когнитивной психологией.

Автор полагает, что аналогии между гештальт-феноменами восприятия и феноменами в геномном гештальт-правиле позволяют развивать гештальт-генетику. Последняя изучает целостные генетические паттерны, относительно независимые от частных компонент. Она соприкасается с учением создателя аналитической психологии К.Юнга и его сподвижника Нобелевского лауреата по физике В.Паули об архетипах бессознательного; они связывали эти архетипы со схемами Инь-Ян древнекитайской «Книги перемен» (или «Ицзин») и ее таблицей 64 гексаграмм, имеющими глубокие структурные аналогии с алфавитами ДНК. Эти аналогии были отмечены, в том числе, видными генетиками: Г.С. Стентом в 1969 году и Нобелевским лауреатом по генетике

Ф. Жакобом в 1977 году (детали многих таких аналогий описаны в книгах [Петухов, 2008; Petoukhov, He, 2010]).

По нашему мнению, истоки генетически наследуемой способности мозга работать с гештальт-образами следует искать в гештальт-генетике. В частности, гештальт-генетика способна предложить новые подходы к пониманию помехоустойчивости генетической информации при мутациях ДНК-текстов.

В дополнение к феноменам гештальт-психологии в живых организмах существует множество генетически унаследованных физиологических феноменов, в которых одна и та же целостная картина реализуется в условиях широкого разнообразия составляющих элементов и которые можно отнести к гештальт-биологии (это новое название предлагается автором как объединение генетически унаследованных гештальт-подобных феноменов разных типов). Например, гештальт-биология включает некоторые генетически унаследованные явления морфогенеза (законы филлотаксиса, спирализация биологических структур на различных уровнях и ветвях биологической эволюции), а также некоторые функциональные явления такие, как гомеостаз на разных этапах онтогенеза и обработка сенсорной информации от разных органов чувств в соответствии с основным психофизическим законом Вебера-Фехнера. В частности, гештальт-генетика включает в сферу своего охвата наблюдение классика эмбриологии К. Бэра о том, что куриные эмбрионы сильно отличаются друг от друга, в то время как полученные из них взрослые организмы удивительно похожи. Автор считает, что в основе гештальт-биологии лежат различные генетически унаследованные феномены гештальт-генетики.

Также следует напомнить, что молекулярный состав живого тела постоянно меняется, сохраняя при этом форму тела. Белки нашего тела участвуют в непрерывных циклах «жизнь-смерть» их сборки и разборки на аминокислоты. Например, период полураспада гормона инсулина составляет 6-9 минут, и т. д. Другими словами, генетически унаследованные части нашего тела постоянно умирают и возрождаются. Принимая во внимание такие явления, известный физиолог А.Г. Гурвич утверждал: «*основной проблемой биологии является поддержание формы при постоянном обновлении субстрата*» [Гурвич, 1977]. На наш взгляд, описанные явления гештальт-генетики с ее гештальт-правилами процентных соотношений в геномных многослойных ДНК-текстах напрямую связаны с этой фундаментальной проблемой биологии.

Гештальт-генетика определяет новые подходы к пониманию онтогенеза. Используя терминологию гештальт-психологии и гештальт-терапии, автор предлагает трактовать онтогенетические процессы как пошаговые процессы «закрытия» определенных генетических гештальтов. Следует отметить, что все физиологические системы вынуждены нести структурную печать генетического кода, поскольку они должны быть генетически закодированы для передачи потомкам и выживания.

Гештальт-генетика касается также феномена биомеханики движений, описанного классиком биомеханики Н.А.Бернштейном: общая целевая задача движения выполняется точно независимо от неточностей составляющих его двигательных подзадач [Бернштейн, 1966]. Например, при повторении точного удара молотком по гвоздю человек каждый раз использует разные траектории, скорости и ускорения частей тела с изменением как углов сгибаний в суставах,

так и активности многих мышц каждого сустава с множеством мотонейронов каждой мышцы (Рис. 4.1).

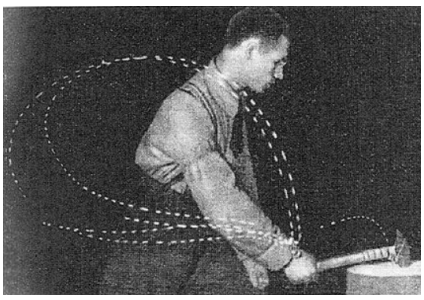


Рис. 4.1. Исследования биомеханики движений по Н.А.Бернштейну (с сайта https://ru.wikipedia.org/wiki/Бернштейн,_Николай_Александрович).

Этот вопрос о том, как центральная нервная система способна адекватно контролировать множество степеней свободы костно-мышечной системы был впервые поставлен Бернштейном и теперь известен как «проблема Бернштейна». Проблема степеней свободы в моторном контроле заключается в том, что у людей и животных есть несколько способов выполнить движение для достижения той же цели, используя избыточные нейрофизиологические степени свободы. То, как нервная система «выбирает» подмножество этих почти бесконечных степеней свободы, является большой трудностью в понимании моторного контроля и моторного обучения. Другими словами, при нормальных обстоятельствах не существует простого однозначного соответствия между моторной проблемой (или задачей) и моторным решением проблемы.

Гештальт психология рассматривает также хорошо известные явления постоянства восприятия, отражающие способность мозга стабильно распознавать формы предметов и других структур внешнего мира в условиях значительного изменения условий их предъявления органам чувств. Например, изменение цветовой освещенности комнаты не мешает человеку распознавать формы и цвета находящихся в ней предметов, хотя на сетчатку глаз попадают другие световые частоты от предметов. Подобных свойств можно ожидать от информационных явлений в гештальт-генетике, которые предвосхищают феномены гештальт-психологии.

Согласно закону Менделя независимого наследования признаков, информация с уровня молекул ДНК диктует макроструктуры живых организмов через множество независимых каналов, несмотря на сильные шумы и помехи. Например, цвета волос, глаз и кожи наследуются независимо друг от друга. Соответственно, каждый организм - это машина многоканального помехоустойчивого кодирования. Гештальт-генетика с ее геномными гештальт-правилами может помочь в понимании этой многоканальной помехоустойчивости.

5. Геномные ДНК-тексты и квантовая информатика.

Ряд авторов давно высказывает предположение о квантово-информационной природе живых тел. Например, Р. Пенроуз, рассматривая тело как квантовый компьютер, обращается к тубулиновым белкам, которые могут переключаться из одного состояния в другое по аналогии с триггерами [Пенроуз, 2005]. Анализ живых тел на более глубоком - геномном - уровне приводит к более

убедительным и плодотворным свидетельствам об организмах как квантовых информационных объектах (или как квантово-подобных объектах, математическое моделирование которых может строиться на формализмах квантовой информатики).

В квантовой механике и квантовой информатике при анализе вероятностей событий традиционно учитываются амплитуды этих вероятностей, равные квадратному корню из их значений. Правило геномного гештальта для процентного соотношения *n*-плетов (в C-, G-, A- и T-группировках) было выявлено при анализе сумм вероятностей *n*-плетов из определенных группировок, например, всех дублетов, начинающихся с буквы C: Σ%CN = %CC + %CA + %CT + %CG. Каждая из этих вероятностей %CC, %CA, %CT и %CG имеет свою собственную амплитуду в виде квадратного корня из нее. Сумму процентов (то есть вероятностей) *n*-плетов в каждой из этих четырех группировок можно интерпретировать как квадрат длины соответствующего вектора, компоненты которого равны амплитудам этих вероятностей (то есть квадратным корням из данных вероятностей). Например, величина Σ%CN суммы процентов 4 дуплетов %CC+%CG+%CA+%CT является квадратом длины 4-мерного вектора $[\sqrt{\%CC}\ ,\ \sqrt{\%CG}\ ,\ \sqrt{\%CA}\ ,\ \sqrt{\%CT}\ ]$. Аналогично, величина Σ%CNN суммы процентов 16 триплетов с буквой C на первой позиции равна величине квадрата длины 8-мерного вектора $[\sqrt{\%CCC}, \sqrt{\%CCG}, \sqrt{\%CCA}, \sqrt{\%CCT}, \sqrt{\%CGC}, \sqrt{\%CGG}, \sqrt{\%CGA}, \sqrt{\%CGT}, \sqrt{\%CAC}, \sqrt{\%CAG}, \sqrt{\%CAA}, \sqrt{\%CAT}, \sqrt{\%CTC}, \sqrt{\%CTG}, \sqrt{\%CTA}, \sqrt{\%CTG}]$, и так далее.

С этой точки зрения постоянство величин сумм процентов в столбцах таблицы на Рис. 3.4 означает постоянство длины тех 2^n-мерных векторов, координаты которых являются амплитудами вероятностей соответствующих *n*-плетов. Такой метрический подход позволяет разрабатывать новые методы сравнительного векторного анализа в генетике, которые сейчас изучаются в нашей лаборатории.

Сравним все *m*-позиционные тетрагруппировки в случае реальных процентов *n*-плетов (аналогично показанным на Рис. 3.2 и 3.3) и в случае эталонных процентов (аналогичных показанным на Рис. 3.7 и 3.8). Во всех сравниваемых тетрагруппировках соответствующие 2^n-мерные векторы амплитуд вероятностей имеют разные значения координат. Но длины этих векторов во всех сравниваемых случаях равны между собой в силу правила гештальта, которое говорит о равенстве сумм процентов *n*-плетов в каждой группировке одному и тому же значению, то есть проценту соответствующего нуклеотида в анализируемом геномном ДНК-тексте. Соответственно, эти векторы одинаковой длины могут быть преобразованы друг в друга с помощью унитарных преобразований, которые не изменяют длину векторов и являются либо поворотами, либо зеркальными отражениями.

Таким образом, алгебраическая гештальт-генетика оказывается связанной с унитарными операторами, которые являются ключевыми для квантовой информатики: как известно, все вычисления в квантовых компьютерах и алгоритмах квантового поиска основаны на унитарных операторах в качестве квантовых вентилей. Любая унитарная матрица может служить квантовым вентилем. В квантовой механике эволюция замкнутой квантовой системы описывается унитарными преобразованиями.

Основоположник квантовой информатики Ю.И.Манин ввел в своей книге [Манин, 1980, с.15] понятие квантового компьютера именно при анализе особенностй скоростной обработки «*генетическими автоматами*» информации

ДНК в хромосомах, пророчески указывая при этом на важную роль унитарных вращений и тензорных произведений: «*Квантовый автомат должен быть абстрактным: его математическая модель должна использовать лишь самые общие квантовые принципы, не предрешая физических реализаций. Тогда модель эволюции есть унитарное вращение в конечномерном гольбертовом пространстве, а модель виртуального разделения на подсистемы отвечает разложению пространства в тензорное произведение. Где-то в этой картине должно найти место взаимодействие, описываемое по традиции эрмитовыми операторами и вероятностями*». Тем самым, само рождение квантовой информатики, столь перспективной для задач искусственного интеллекта, произошло благодаря стремлению понять особенности генетической информатики! Очевидно, что излагаемые в данной статье универсальные правила вероятностей в геномных ДНК как многослойных текстах, написанных на языках тензорно взаимосвязанных алфавитов *n*-плетов ДНК, вместе с привязанными к этим правилам унитарными вращениями, согласуются с этим предвидением Ю.И.Манина.

Поскольку в данной статье речь идет о многих слоях текстов геномной ДНК, можно предположить, что, в частности, в многослойных текстах ДНК работает целый набор алгоритмов квантового поиска, каждый из которых индивидуально ориентирован на определенный текстовый слой. В статьях [Patel, 2001a-c] предполагается, что генетический код связан с квантовым алгоритмом Гровера. Статья [Petoukhov, 2018] представляет другие свидетельства связи структуриро-ванной системы генетического кодирования с унитарными матрицами.

Попутно можно отметить, что вся генетически унаследованная кинематическая схема движений нашего тела с его частями построена на поворотах в суставах и зеркальных отражениях, то есть на унитарных трансформациях, имеющих физиологическое значение.

Рис. 5.1. Кинематическая схема движений тела построена на унитарных преобразованиях вращения в суставах и зеркального отражения.

Автор обращает особое внимание на унитарные матрицы в развитии гештальт-генетики, поскольку унитарные преобразования вращения могут рассматриваться в качестве основы для моделирования известного морфогенетического гештальт-феномена спирализации в биологии: систематического существования спиральных морфологических конфигураций на разных уровнях и ветвях биологической организации независимо от их генетически унаследованного биоматериала. Гете даже называл спирали «символами жизни» из-за множества реализаций унаследованных спиральных структур и процессов в живых организмах на различных линиях и уровнях биологической эволюции. В организме человека спиральные структуры, генетически передаваемые от поколения к поколению, представлены в мышцах,

сердце, кровеносных сосудах, костях, нервах, органе слуха (улитка), клеточной организации зародыша (зигота) и др. Структура сухожилий и связок состоит из спиралей, которые, в свою очередь, состоят из коллагена, имеющего тройную спиральную структуру. Спиральные движения (нутация) наблюдаются во время роста корней и побегов, усики растений заворачиваются по спирали, ткань стволов деревьев растет по спирали и т.д. [Петухов, Свирин, Хазина, 2015]. Из-за спиральной биоконфигурации течение всех жидкостей в организме (кровь, лимфа, и моча) имеет спиралевидный характер. Название книги о биоспиралях - «Линии жизни» [Cook, 1914] - отражает их важность для живого вещества. Структурные связи молекулярно-генетических систем с унитарными матрицами также описаны в предыдущих публикациях автора [Петухов, 2008; Petoukhov, He, 2010].

В данной статье отметим дополнительную структурную связь унитарных матриц с нуклеотидным строением двойной спирали ДНК. В этой двойной спирали процентные доли содержания комплемантарных нуклеотидов C и G в точности равны (%C = %G), как и процентные доли комплементарных A и T (%A = %T). Соответственно для двойной спирали ДНК матрица процентов нуклеотидов [%C, %A; %T, %G] является бисимметрической и дважды стохастической (при ее умножении на 2): сумма процентов в каждой ее строке и каждом столбце равна единице. Но дважды стохастические матрицы связаны с унитарными матрицами согласно известной теореме [Прасолов, 2008, с. 350]:

Если матрица $A = \|a_{ij}\|_1^n$ *унитарная, то матрица* $B = \|b_{ij}\|_1^n$, *где* $b_{ij} = |a_{ij}|^2$, *дважды стохастическая.*

Исходя из этой теоремы, Рис. 5.2 показывает взаимосвязь четырех унитарных (2*2)-матриц и дважды стохастической (2*2)-матрицы процентов нуклеотидов в двойной спирали ДНК. Две из показанных четырех унитарных матриц представляют преобразования зеркального отражения. Две другие унитарные матрицы представляются преобразования поворота векторов по часовой стрелке и против нее; одновременно они являются матричными представлениями комплексного числа $z = \sqrt{2\%C} + i*\sqrt{2\%A}$, где $i^2 = -1$.

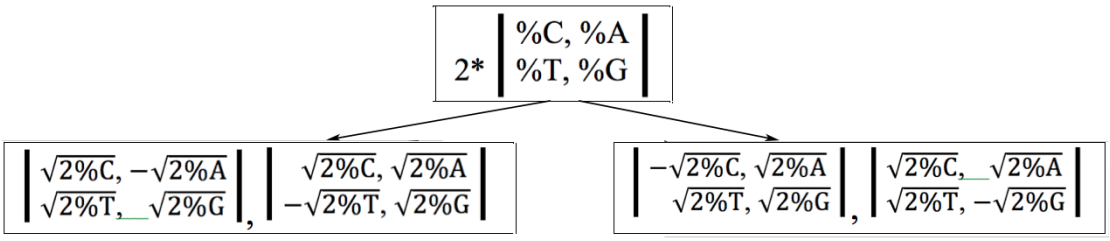


Рис. 5.2. Взаимосвязь дважды стохастической матрицы процентов нуклеотидов в двойной спирали ДНК (верхний ряд) и четырех унитарных матриц (нижний ряд). Две унитарные матрицы (слева) представляют преобразования вращения, а две другие унитарные матрицы (справа) – преобразования зеркального отражения.

Описанная связь матриц процентного содержания нуклеотидов в двойной спирали ДНК с комплексными числами, которые представлены отмеченными унитарными матрицами, полезна для алгебраического и квантово-

механического моделирования генетических явлений, поскольку квантовая механика тесно связана с комплексными числами.

Взаимосвязи (2*2)-матриц, показанные на Рис. 5.2, можно обобщить для (2^n*2^n)-матриц, образованных тензорными степенями исходной унитарной (2*2)-матрицы и дважды стохастической (2*2)-матрицы (как известно, тензорные степени унитарных матриц генерируют новые унитарные матрицы увеличенных порядков). Тензорные степени (2*2)-матриц, представляющих комплексные числа, порождают (2^n*2^n)-матрицы, которые представляют алгебраические расширения комплексных чисел, относящиеся к соответствующим системам многомерных чисел.

Материалы данной статьи дополняют опубликованные ранее работы автора по теме квантовой биологии и формализмов квантовой информатики в биологии [Петухов, 2008; 2020a-c, 2021a,b; Petoukhov, He, 2010; Petoukhov, Petukhova, Svirin, 2019]. Выявленная связь генетики с квантовой информатикой открывает возможность внедрения богатых формализмов квантовой механики и квантовой информатики в алгебраическую биологию для взаимного обогащения этих наук и включения биологии в область развитого математического естествознания. В организме человека около 100 триллионов клеток, образующих единую систему. Формализмы квантовой информатики и гештальт-генетики могут помочь в понимании таких явлений когерентности.

Следует упомянуть предположения многих авторов о том, что формализмы квантовой информатика могут быть эффективно использованы для глубокого понимания и моделирования биологических тел (см., например, [Abbot, Davies, Pati, 2008; Fimmel, Petoukhov, 2020a,b; Igamberdiev, 1993; Matsuno, 1999; Matsuno, Paton, 2000]).

6. Матричные алгебры и матрицы *m*-позиционных тетрагруппировок.

Вернемся к четырем столбцам (Рис. 3.4) сумм процентов *n*-плетов в их *m*-позиционных тетрагруппировках из разных *n*-плетных слоев ДНК-текста хромосомы человека №1. Выше в разделе №3 все *n*-плеты, фигурирующие в суммах одного из 4 столбцов, были обозначены соответствующим карточным символом:

- *n*-плеты первого столбца (C-столбца) - символом ♠ ;
- *n*-плеты второго столбца (G-столбца) - символом ♣ ;
- *n*-плеты третьего столбца (A-столбца) - символом ♥ ;
- *n*-плеты четвертого столбца (T-столбца) – символом ♦ .

Каково расположение членов этих 4 группировок в исходном тензорном семействе матриц $[C, A; T, G]^{(n)}$ ДНК-алфавитов *n*-плетов (Рис. 2.2), построенных на основе системы бинарных оппозиций в алфавите 4 нуклеотидов? Рис. 6.1-6.4 показывают их матричное расположение для случаев $n = 1, 2, 3, 4$, используя введенные символы группировок *n*-плетов.

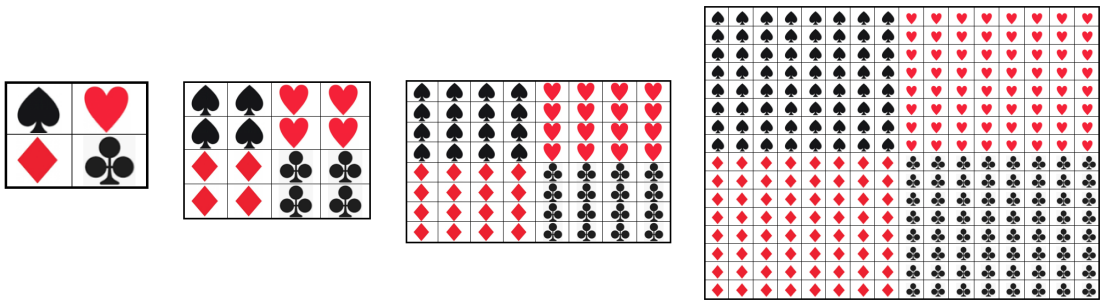


Fig. 6.1. Расположение *n*-плетов, начинающихся с атрибутивного нуклеотида C, G, A или T, в матрицах тензорного семейства $[C, A; T, G]^{(n)}$ при $n = 1, 2, 3, 4$ (объяснение в тексте).

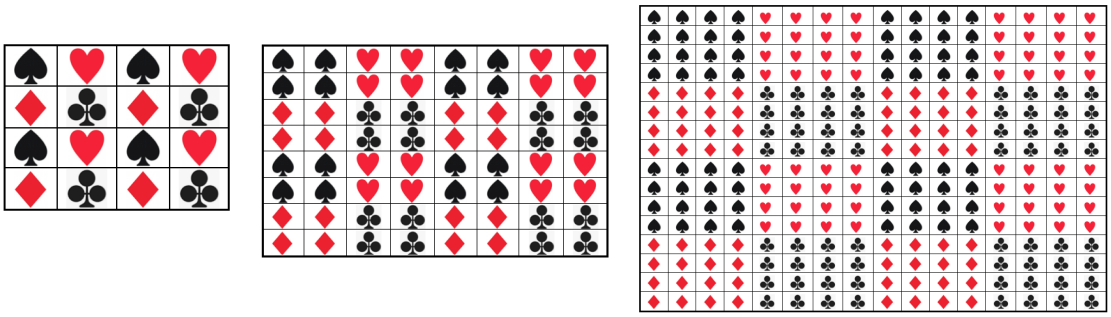


Рис. 6.2. Расположение 16 дуплетов, 64 триплетов и 256 тетраплетов, имеющих атрибутивный нуклеотид C, G, A или T на второй позиции, в матрицах тензорного семейства $[C, A; T, G]^{(n)}$ при $n = 2, 3, 4$.

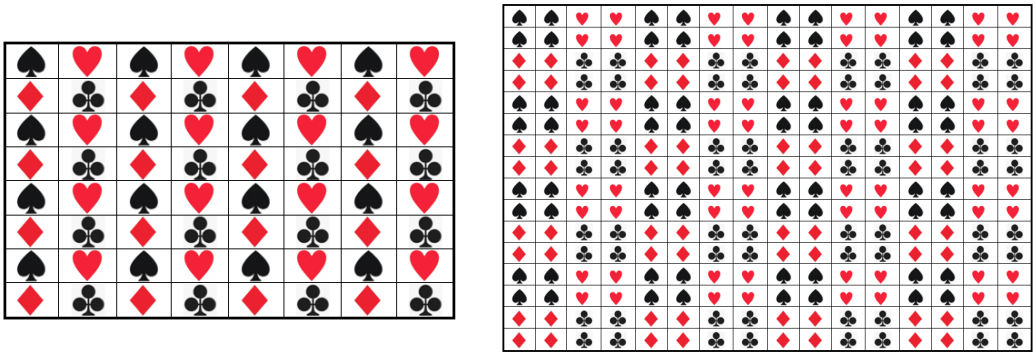


Рис. 6.3. Расположение 64 триплетов и 256 тетраплетов, которые имеют атрибутивный нуклеотид C, G, A или T на третьей позиции, в матрицах тензорного семейства $[C, A; T, G]^{(n)}$ при $n = 3, 4$.

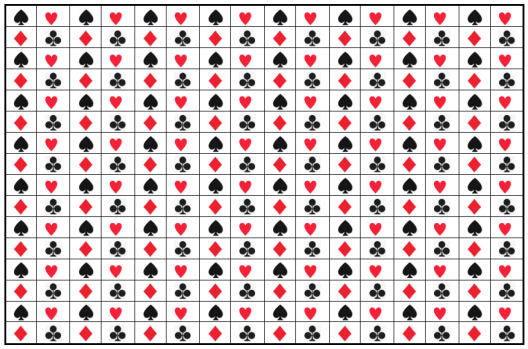


Рис. 6.4. Расположение в матрице [C, A; T, G]$^{(4)}$ тетрагруппировки 256 тетраплетов, которые имеют атрибутивный нуклеотид C, G, A или T на четвертой позиции.

Мозаика всех этих матриц имеет меандровый характер: чередующиеся в них связки черных и красных символов имеют одинаковую длину. Это напоминает функции Радемахера-Уолша, широко используемые в теории помехоустойчивого кодирования информации и состоящие из меандровых последовательностей элементов +1 и -1. Заменяя в меандровых геноматрицах черные символы на «+1», а красные – на «-1», получаем числовые представления этих генетических (2^n*2^n)-матриц, которые оказываются 2^n-мерными гиперболическими числами с единичными координатами. Поясним это на примере одной из показанных мозаичных (4*4)-геноматриц, взятой с Рис. 6.1.

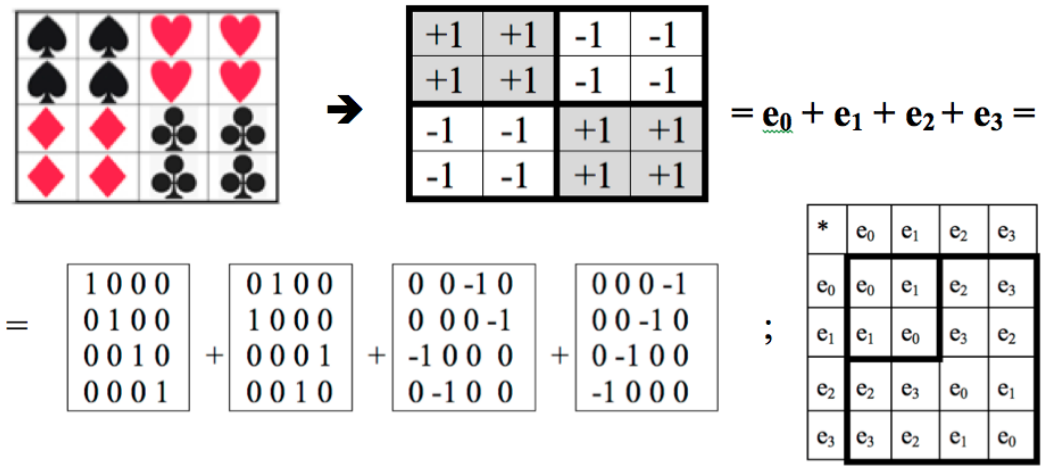


Рис. 6.5. Декомпозиция матрицы, взятой с Рис. 6.1, дает четыре разреженных матрицы, определяющих таблицу умножения алгебры 4-мерных гиперболических чисел (внизу справа). В таблице этой алгебры жирной рамкой указана ее подалгебра.

Показанная на Рис. 6.5 декомпозиция числовой матрицы дает набор из четырех разреженных матриц e_0, e_1, e_2, e_3, замкнутый относительно умножения и дающий таблицу умножения, известную в алгебре 4-мерных гиперболических чисел как таблица умножения ее базисных единиц. Каждая из этих разреженных матриц является матрицей перестановок, применяемой для анализа структурированных текстов ДНК.

Примеры числовых аналогов других матриц, указанных на Рис. 6.1-6.4, и их декомпозиций на наборы разреженных матриц, также замкнутых относительно умножения и определяющих таблицы умножения алгебр соответствующих 2^n-мерных гиперболических чисел, полезных для моделирования генетических структур, имеются в работах [Petoukhov, 2021a,b].

Частным случаем 2-мерных гиперболических чисел являются гиперболические повороты, известные в специальной теории относительности под именем преобразований Лоренца. Они имеют отношение также к генетически наследуемому основному психофизическому закону Вебера-Фехнера. Покажем связь гиперболических поворотов с этим логарифмическим законом (натуральные логарифмы в начале своей истории недаром назывались гиперболическими логарифмами).

Восприятие мира - зрительное, слуховое, обонятельное, тактильное, вкусовое и пр. - по наследуемым сенсорным каналам подчиняется основному психофизическому закону Вебера-Фехнера, несмотря на различия органов чувств (глаз, ушей, носа,…) по внешнему виду и составу (Рис. 6.6). Данный закон гласит, что интенсивность восприятия пропорциональна логарифму интенсивности стимула:

$$p = k*\ln(x/x_0) = k*\{\ln(x) - \ln(x_0)\}, \qquad (6.1)$$

где p – интенсивность восприятия, x – интенсивность сти-мула, x_0 – пороговое значение стимула, k – коэффициент, ln – символ натурального логарифма.

В силу этого закона, например, сила звука в технике измеряется в логарифмической шкале децибел.

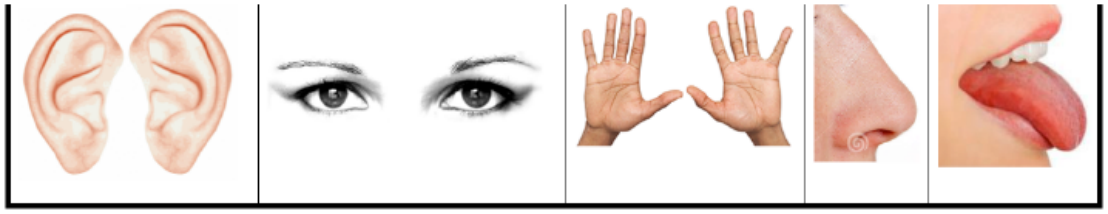


Рис. 6.6. Различные органы чувств, сильно различающиеся по форме и составу, объединены тем, что их сенсорика подчиняется закону Вебера-Фехнера.

Известно, что для любого вещественного числа $a>1$ натуральный логарифм определяется как площадь под гиперболой $y = 1/x$ от 1 to a (Рис. 6.7). Соответственно две точки $(x, 1/x)$ и $(x_0, 1/x_0)$ этой гиперболы задают величину $\ln(x) - \ln(x_0)$ из значения интенсивности восприятия p в законе Вебера-Фехнера. Изменение величины интенсивности стимула x_1 в новое значение x_2 соответствует гиперболическому повороту, который в этой гиперболе переводит точку $(x_1, 1/x_1)$ в точку $(x_2, 1/x_2)$ и характеризует величину изменения интенсивности восприятия в данном законе: $\Delta p = k*\ln(x_2/x_1)$.

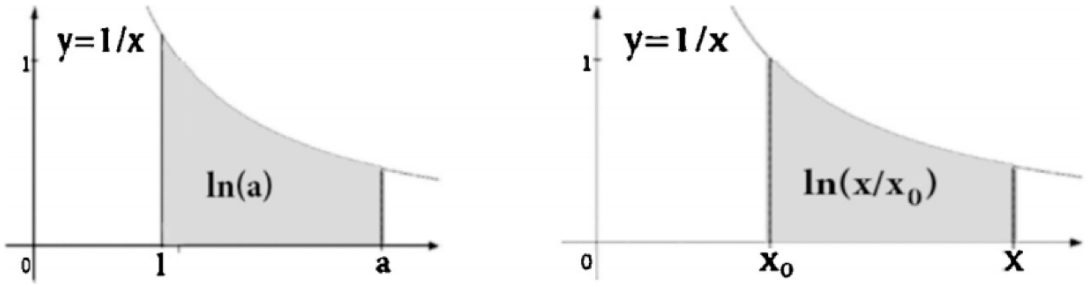


Рис. 6.7. Натуральный логарифм как площадь под гиперболой y = 1/x.
Слева: ln(a) равен площади под гиперболой от 1 до a.
Справа: $\ln(x/x_0)$ равен площади под гиперболой от x_0 до x.

Генетически наследуемый закон Вебера-Фехнера является законом далеко не только нервной системы. Он также выполняется для низших организмов, не имеющих нервных клеток: "*этот закон применим для хемотропических, гелиотропических и геотропических движений бактерий и грибов и антерозоидов папоротников, мхов и явнобрачных…. Закон Вебера-Фехнера, следовательно, - не закон нервной системы и центров, но закон протоплазмы вообще и ее способности отвечать на раздражения*" [Шульц, 1916].

Закон Вебера-Фехнера является одной из частей темы алгебраической гештальт-генетики и ее влияния на физиологию.

7. Наклонные проекторы и проекторные тетрагруппировки *n*-плетов в слоях ДНК-текстов геномов.

Продолжим изучение свойств процентного содержания *n*-плетов в многослойном представлении длинных ДНК-текстов. До сих пор мы рассматривали правило гештальта, относящееся к *m*-позиционным тетрагруппировкам *n*-плетов, которые связаны с алгеброй гиперболических чисел (см. предыдущий раздел). По этой причине мы дополнительно называем такие *m*-позиционные тетрагруппировки гиперболическими тетрагруппировками, а соответствующее генетическое гештальт правило - гиперболическим гештальт правилом. Этот раздел посвящен другому типу тетрагруппировок, называемых проекторными тетрагруппировками *n*-плетов, поскольку они связаны с наклонными проекторами. Тетрагруппировки проекторов сопряжены с соответствующим проекторным гештальт-правилом, описанным ниже.

Наше зрительное восприятие и «метод проекции» в рисовании основаны на проекциях внешних объектов на сетчатку и плоскость рисунка (рис. 7.1). В математике такие операции проекции выражаются с помощью квадратных матриц, которые называются «проекторными операторами» (или сокращенно «проекторами»).

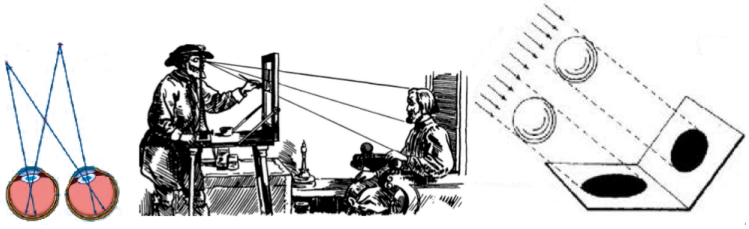


Рис. 7.1. Примеры проецирования. Слева: ход световых лучей к сетчатке глаз.
В середине: метод проецирования при рисовании.
Справа: ортогональное и наклонное проецирование на плоскость.

Понятие проекторных (или проекционных) операторов - одно из важных понятий во многих небиологических науках: физике, включая квантовую механику; математике; информатике; химия; математическая логика и т. д. Матрица P на Рис. 7.2 представляет собой пример проектора, проецирующего вектор [x, y, z] на плоскость [x, y, 0].

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad P \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$$

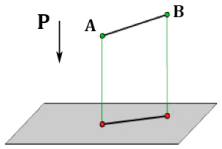


Рис. 7.2. Пример матричного оператора ортогонального проецирования.

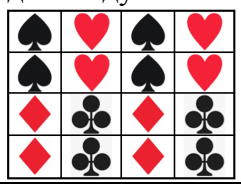
Проекторы математически выражаются квадратными матрицами (http://mathworld.wolfram.com/ProjectionMatrix.html). Необходимым и достаточным условием того, что матрица P является оператором проектирования, является выполнение равенства $P^2 = P$. Все множество проекторов разделено на два класса:

- Ортогональные проекторы, которые выражаются симметричными матрицами. Они наиболее известны и имеют множество приложений;
- Наклонные проекторы, которые выражаются несимметричными матрицами. Они значительно менее известны, но именно они будут в центре внимания в данном разделе в связи с геномными гештальт феноменами.

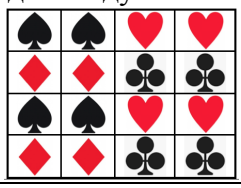
Излагаемые авторские результаты свидетельствуют о том, что значение проекторов для биоинформатики не ограничивается их ролью в зрительном восприятии (Рис. 7.1), но что вся система генетической информатики активно использует их. Опишем эти проекторные тетрагруппировки *n*-плетов, которые существуют в разных *n*-плетных слоях геномных ДНК-текстов; в каждой из их четырех группировок сохраняется величина сумм процентов соответствующих *n*-плетов как в разных слоях, так и в разных проекторных тетрагруппировках *n*-го слоя (n = 2, 3, 4,…, но намного меньше длины рассматриваемого текста геномной ДНК).

Вернемся к процентным матрицам 16 дублетов и 64 триплетов в хромосоме №1 человека (Рис. 3.2 и 3.3). Для облегчения визуального представления этих тетрагруппировок проекторов *n*-плетов, процентные матрицы на Рис. 3.2 и 3.3 преобразуются в соответствующие символические матрицы на Рис. 7.3 с 4 типами символов игральных карт в ячейках: пики (♠), червы (♥), бубны (♦) и трефы (♣).

Тетрагруппировка №1 для 16 дуплетов:



Тетрагруппировка №2 для 16 дуплетов:



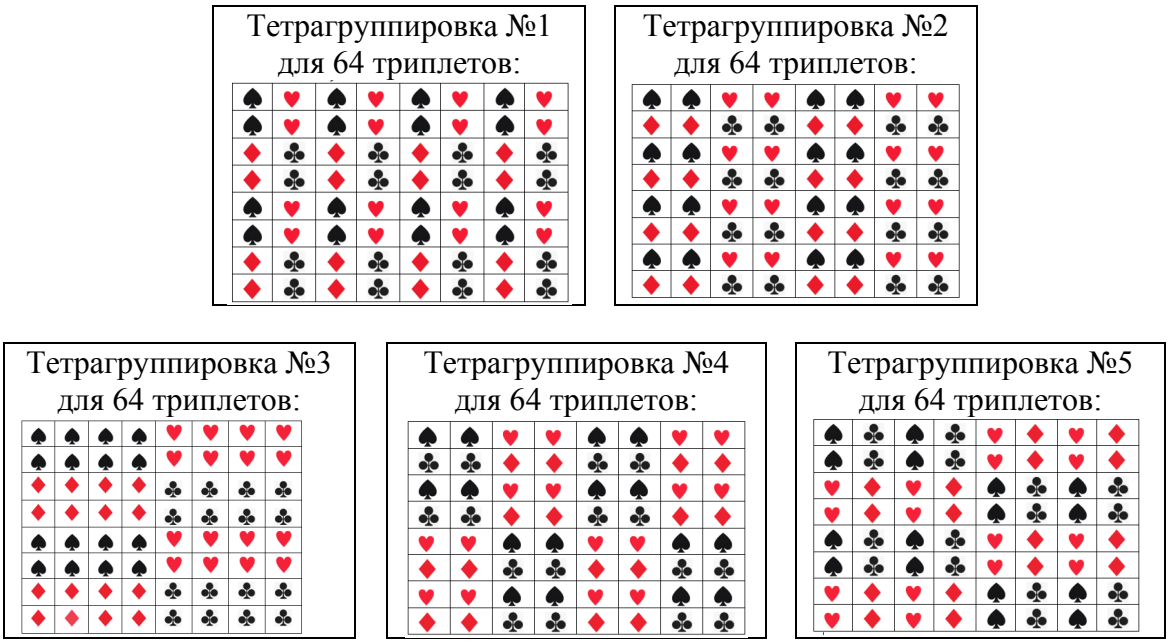


Fig. 7.3. Расположение *n*-плетов из разных проекторных тетрагруппировок в матрицах 16 дублетов и 64 триплетов из Рис. 2.2. Соответствующие ячейки каждой из четырех группировок обозначаются одним из символов игральных карт.

Удивительным результатом исследования разных геномов является обнаружение константности сумм процентов Σ%♥, Σ%♦, Σ%♠, and Σ%♣ *n*-плетов, которые обозначены одинаковым символом, во всех матрицах с различными мозаиками тетрагруппировок на Рис. 7.3:

$$\Sigma\%\heartsuit \approx \text{constant}, \ \Sigma\%\diamondsuit \approx \text{constant}, \ \Sigma\%\spadesuit \approx \text{constant}, \ \Sigma\%\clubsuit \approx \text{constant}. \qquad (7.1)$$

Например, для ДНК хромосомы человека №1, процентная (4*4)-матрица дуплетов которой показана на Рис. 3.2, имеем следующие процентные суммы (7.2) в случае проекторной тетрагруппировки №1 для 16 дуплетов из Рис. 7.3:

$$\begin{aligned}
\Sigma\%\heartsuit &= \%CA + \%CG + \%AA + \%AG \approx \\
&\approx 0.07274 + 0.01031 + 0.09504 + 0.07137 \approx \mathbf{0.250}, \\
\Sigma\%\diamondsuit &= \%TC + \%TT + \%GC + \%GT \approx \\
&\approx 0.06008 + 0.09568 + 0.04402 + 0.05046 \approx \mathbf{0.250}, \\
\Sigma\%\spadesuit &= \%CC + \%CT + \%AC + \%AT \approx \\
&\approx 0.05409 + 0.07134 + 0.05033 + 0.07429 \approx \mathbf{0.250}, \\
\Sigma\%\clubsuit &= \%TA + \%TG + \%GA + \%GG \approx \\
&\approx 0.06312 + 0.07286 + 0.06008 + 0.05419 \approx \mathbf{0.250.}
\end{aligned} \qquad (7.2)$$

Тетрагруппировка №2 для 16 дуплетов из Рис. 7.3 имеет другую комбинацию 4 дуплетов и процентов в каждой группировке, но те же самые величины процентных сумм (7.3):

$$\begin{aligned}
\Sigma\%\heartsuit &= \%AC + \%AA + \%GC + \%GA \approx \\
&\approx 0.05033 + 0.09504 + 0.04402 + 0.06008 \approx \mathbf{0.250},
\end{aligned}$$

$\Sigma\%\blacklozenge$ = %CT + %CG + %TT + %TG ≈
≈ 0.07134 + 0.01031 + 0.09568 + 0.07286 ≈ **0.250**,
$\Sigma\%\spadesuit$ = %CC + %CA + %TC + %TA ≈
≈ 0.05409 + 0.07274 + 0.06008 + 0.06312 ≈ **0.250**,
$\Sigma\%\clubsuit$ = %AT + %AG + %GT + %GG ≈
≈ 0.07429 + 0.07137 + 0.05046 + 0.05419 ≈ **0.250** (7.3)

Аналогичное правило константности величин процентных сумм в проекторных тетрагруппировках выполняется для всех пяти (8*8)-матриц (Рис. 7.3), мозаики которы представляю различные проекторные тетрагруппировки для ДНК-алфавита 64 триплетов. Каждый читатель может сам проверить выполнение этого правила константности посредством суммирования соответствующих процентов. Автор проиллюстрирует здесь эту константность только для величин процентных сумм $\Sigma\%\spadesuit$ 16 триплетов, обозначенных символом ♠, используя мозаичные матрицы на Рис. 7.3 и процентные матрицы триплетов на Рис. 3.3.

Тетрагруппировке №1 для 64 триплетов на Рис. 7.3 соответствует сумма:

$\Sigma\%\spadesuit$ = %CCC+%CCT+%CAC+%CAT+%ACC+%ACT+%AAC+%AAT+
+%TTC+TCT+%TAC+%TAT+%GCC+%GCT+%GAC+%GAT ≈
≈ 0.0138+0.0185+0.0152+0.0179+0.0118+0.0162+0.0145+0.0238+0.0159
+0.0223+0.0110+0.0194+0.0125+0.0144+0.0096+0.0133 ≈ **0.250**. (7.4)

Тетрагруппировке №2 для 64 триплетов на Рис. 7.3 соответствует сумма:

$\Sigma\%\spadesuit$ = %CCC+%CCA+%ACC+%ACA+%CTC+%CTA+%ATC+%ATA+
+%TCC+%TCA+%GCC+%GCA+%TTC+%TTA+%GTC+%GTA ≈
≈ 0.0138+0.0188+0.0118+0.0198+0.0176+0.0127+0.0132+0.0194+0.0159
+0.0196+0.0125+0.0146+0.0197+0.0198+0.0096+0.0112 ≈ **0.250**. (7.5)

Тетрагруппировке №3 для 64 триплетов на Рис. 7.3 соответствует сумма:

$\Sigma\%\spadesuit$ = %CCC+%CCA+%CAC+%CAA+%CCT+%CCG+%CAT+%CAG+%TCC+
+%TCA+%TAC+%TAA+%TCT+%TCG+%TAT+%TAG ≈
≈ 0.0138+0.0188+0.0152+0.0186+0.0185+0.0029+0.0179+0.0210+0.0159+
+0.0196+0.0110+0.0199+0.0223+0.0023+0.0194+0.0128 ≈ **0.250**. (7.6)

Тетрагруппировке №4 для 64 триплетов на Рис. 7.3 соответствует сумма:

$\Sigma\%\spadesuit$ = %CCC+%CCA+%ACC+%ACA+%CTC+%CTA+%ATC+%ATA+%TAC
+%TAA+%GAC+%GAA+%TTA+%TGA+%GGC+%GGA ≈
≈ 0.0138+0.0188+0.0118+0.0198+0.0176+0.0127+0.0132+0.0194+
+0.0110+0.0096+0.0196+0.0146+0.0195+0.0126+0.0160 ≈ **0.250**. (7.7)

Тетрагруппировке №5 для 64 триплетов на Рис. 7.3 соответствует сумма:

$\Sigma\%\spadesuit$ = %CCC+%CCT+%CAC+%CAT+%ATC+%ATT+%AGC+%AGT+%TCC
+%TCT+%TAC+%TAT+%GTC+%GTT+%GGC+%GGT ≈
≈ 0.0138+0.0185+0.0152+0.0179+0.0132+0.0239+0.0144+0.0161+0.0159+

+0.0223+0.0110+0.0194+0.0096+0.0145+0.0126+0.0119 ≈ **0.250**. (7.8)

В показанном частном числовом примере хромосомы человека №1, выражения (7.2-7.8) показывают, что Σ%♥ ≈ Σ%♦ ≈ Σ%♠ ≈ Σ%♣ ≈ 0.250. Но в общем случае для геномных ДНК-текстов выполняются равенства (7.9):

$$\Sigma\%\heartsuit \approx \Sigma\%\diamondsuit, \qquad \Sigma\%\spadesuit \approx \Sigma\%\clubsuit \tag{7.9}$$

Например, в случае бактериального генома *Bradyrhizobium japonicum* (см. Приложение I) выполняются следующие равенства для проекторных тетрагруппировок *n*-плетов во всех 7 мозаичных матрицах дуплетов и триплетов на Рис. 7.3:

$$\Sigma\%\heartsuit \approx \Sigma\%\diamondsuit \approx 0.277, \qquad \Sigma\%\spadesuit \approx \Sigma\%\clubsuit \approx 0.223 \tag{7.10}$$

Автор рассматривает этот геномный феномен постоянства величин суммарных процентов (7.10) в проекторных тетрагруппировках, имеющих различный процент отдельных *n*-плетов в различных слоях длинных ДНК-текстов, как основу генетического гештальт-правила вероятностей для проекторных тетрагруппировок и как основу соответствующего предполагаемого принципа запрета в генетике.

Следует отметить существенное различие гештальт феноменов в случае *m*-позиционных тетрагруппировок, которые связаны с 4 столбцами на Рис. 3.4, и в случае проекторных тетрагруппировок, которые связаны с выражение (7.9). В *m*-позиционных тетрагруппировках величины сумм процентов *n*-плетов в каждой из 4 групп равны проценту соответствующих нуклеотидов A, T, C или G. В проекторных тетрагруппировках величины сумм процентов *n*-плетов в каждой из 4 группировок не имеют видимой связи с процентным содержанием отдельных нуклеотидов.

Следует отметить, что в данной статье представлены лишь некоторые примеры проекторных тетрагруппировок *n*-плетов. В наборе из 256 тетраплетов существует большее количество проекторных тетрагруппировок тетраплетов. Кажется естественным предположить, что тетрагруппировки проекторов существуют также в геномных наборах из 4^n видов *n*-плетов при $n = 5, 6, 7, \ldots$ но автор не проверял это предположение на конкретных геномах.

Автор рассматривает этот генетический феномен постоянства величины процентных сумм (7.10) в различных проекторных тетрагруппировках (имеющих различные проценты отдельных *n*-плетов в описываемых слоях длинных ДНК-текстов), как основу соответствующего **генетического гештальт правила для проекторных тетрагруппировок**.

Аналогичные процентные суммы для этих 7 тетрагруппировок (Рис. 7.3) имеют место также в случае модельных процентов 16 дуплетов и 64 триплетов в тех матрицах (Рис. 3.7, 3.8), которые являются второй и третьей тензорными степенями процентной матрицы нуклеотидов и которые имеют свои собственные проценты отдельных дуплетов и триплетов, отличающиеся от реально существующих:

$$\mathbf{\Sigma\%}\spadesuit \approx 0.250; \quad \mathbf{\Sigma\%}\clubsuit \approx 0.250; \quad \mathbf{\Sigma\%}\heartsuit \approx 0.250; \quad \mathbf{\Sigma\%}\diamondsuit \approx 0.250. \tag{7.11}$$

Рис. 7.4 показывает, что некоторые примеры (4*4)-матриц с проекторными тетрагруппировками 16 дуплетов и (8*8)-матриц с проекторными тетрагруппировками 64 триплетов (представленные на Рис. 7.3) взаимно связаны через тензорное произведение этих (4*4)-матриц на (2*2)-матрицу [1, 1; 1, 1].

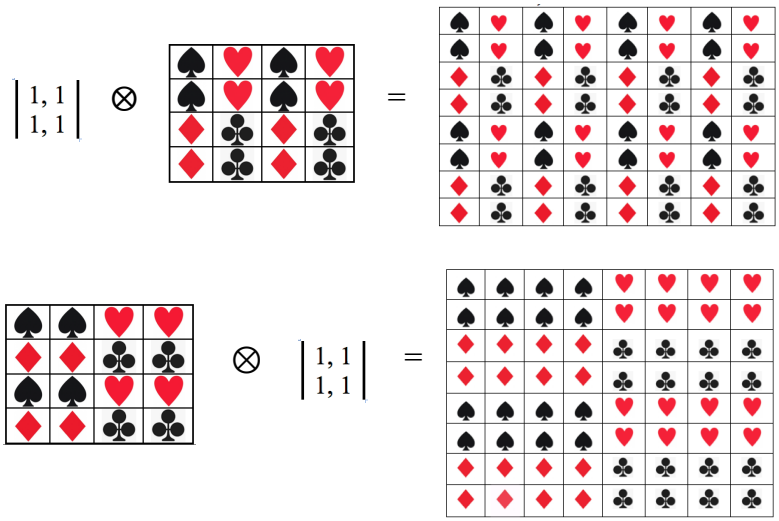
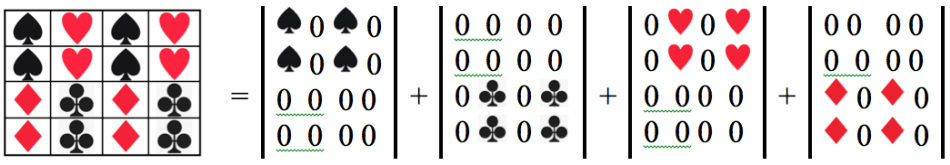


Рис. 7.4. Тензорная взаимосвязь между примерами проекторных матриц четвертого и восьмого порядков из Рис. 7.3. Символ ⊗ означает тензорное произведение.

Следует отметить, что список связей наклонных проекторных операторов с процентным составом *n*-плетов в многослойной презентации геномных ДНК-текстов далеко не ограничивается тетрагруппировками, отмеченными выше в этой статье.

Декомпозиция каждой матрицы проекторных тетрагруппировок по ее карточным мастям дает набор разреженных матриц по аналогии с примером на Рис. 7.5



$$e_0+e_1+e_2+e_3=\begin{vmatrix}1&0&1&0\\1&0&1&0\\0&0&0&0\\0&0&0&0\end{vmatrix}+\begin{vmatrix}0&0&0&0\\0&0&0&0\\0&1&0&1\\0&1&0&1\end{vmatrix}+\begin{vmatrix}0&1&0&1\\0&1&0&1\\0&0&0&0\\0&0&0&0\end{vmatrix}+\begin{vmatrix}0&0&0&0\\0&0&0&0\\1&0&1&0\\1&0&1&0\end{vmatrix};$$

*	e_0	e_1	e_2	e_3
e_0	e_0	e_2	e_2	e_0
e_1	e_3	e_1	e_1	e_3
e_2	e_0	e_2	e_2	e_0
e_3	e_3	e_1	e_1	e_3

Рис. 7.5. Декомпозиция проекторной матрицы с тетрагруппировкой №1 для 16 дуплетов (из Рис. 7.3) дает четыре разреженных матрицы (верхний ряд), которые аналогичны набору матричных операторов наклонных

проекторов e_0, e_1, e_2, e_3, каждый из которых удовлетворяет критерию проекторов: $e_i^2 = e_i$ (нижний ряд слева). Внизу справа показана таблица умножения этих наклонных проекторов.

Этот набор наклонных проекторов e_0, e_1, e_2, e_3 замнут относительно умножения и определяет показанную таблицу умножения некой алгебры 4-мерных чисел (Рис. 7.5). Декомпозиция по мастям других (2^n*2^n)-матриц с проекторными тетрагруппировками дает наборы матриц наклонных проекторов, замкнутые относительно умножения и определяющие таблицы умножения алгебр 2^n-мерных числовых систем. Подробности этого представлены в препринте [Petoukhov, 2021a].

Данный тип наклонных проекторов, возникший при анализе композитных ДНК-текстов и условно называемый генетическим, обладает важными модельными свойствами. Продемонстрируем эти свойства на примере воздействия таких проекторов на 4-мерный вектор состояния системы $\bar{A} = [a_0(t), a_1(t), a_2(t), a_3(t)]$, координаты которого представляют, например, исходно не связанные между собой циклические или волновые функции времени. Подействуем на этот вектор, например, наклонным проектором $e_2 = [0\ 1\ 0\ 1;\ 0\ 1\ 0\ 1\ ;\ 0\ 0\ 0\ 0;\ 0\ 0\ 0\ 0]$ из Рис. 7.5:

$$[a_0(t), a_1(t), a_2(t), a_3(t)]*e_2 = [0,\ a_0(t)+a_1(t),\ 0,\ a_0(t)+a_1(t)] \qquad (7.12)$$

В результате из вектора состояния со случайным набором координат возник симметризованный вектор состояния (7.12), ненулевые координаты которого синхронно изменяются во времени и вообще идентичны. Образно говоря, эти генетические наклонные проекторы создают порядок из хаоса.

Аналогичные результаты дают другие наклонные проекторы данного «генетического» типа при действии на 2^n-мерные вектора состояния (см. препринт [Petoukhov, 2021a]). Такие проекторы позволяют моделировать многие генетически наследуемые согласования в биологических феноменах. Ведь живой организм представляет собой огромный хор согласованных циклических процессов. Согласно положениям хрономедицины, все наши заболевания являются следствием нарушений в этой согласованности. Насколько автор может судить, такой тип наклонных проекторов ранее не применялся в науке и является новым инструментом анализа и моделирования. Здесь вспоминается известное утверждение Ж.Фурье: «*Пристальное, глубокое изучение природы есть источник самых плодотворных открытий математики*".

8. Гиперболические правила количеств *n*-плетов в геномах и гармоническая прогрессия.

При описанном композитном, многослойном представлении длинного ДНК-текста каждый его *n*-й слой, имеющий вид цепи *n*-плетов как его звеньев, содержит в *n* раз меньше звеньев, чем количество звеньев в первом однобуквенном слое (поскольку в *n*-м слое размер звеньев больше в *n* раз по сравнению с размером звеньев в первом слое). В силу описанного в параграфе 3 феноменологического равенства в геномных ДНК-текстах сумм процентов *n*-плетов, принадлежащих к A-, T-, C- и G-группировкам, можно ожидать, что суммарное количество *n*-плетов, например, с буквой A на их первой позиции будет

зависеть от номера n слоя по гиперболической зависимости $1/n$. Покажем результаты исследования этого ожидания.

Подсчитаем в каждом из первых 20 слоев ДНК-текста хромосомы человека №1 количества их n-плетов, начинающихся с A. Получаем серию из 20 многомиллионных чисел, начало и конец которой показаны на Рис. 8.1.

n	1(мономеры)	2(дуплеты)	3(триплеты)	4(тетраплеты)	...	20 (20-плеты)
$S_A(n)$	67070277	33537501	22360413	16768845	...	3354107

Рис. 8.1. Фрагмент последовательности количеств n-плетов в первых 20 слоях ДНК-текста хромосомы человека №1.

Эта полученная последовательность 20 чисел практически совпадает с гиперболической последовательностью $S(n) = 67070277/n$, числителем которой служит количество нуклеотидов A в этой ДНК. Отклонения реальных количеств от модельной последовательности $S(n)$ составляют сотые доли процента и лежат в диапазоне $\pm 0.03\%$. Сходные результаты получается в этом же ДНК-тексте при его многослойном представлении с подсчетом суммарного количества n-плетов, начинающихся с любой из других трех букв T, C или G (в числителях этих гиперболических зависимостей стоят количества соответствующих нуклеотидов). Рис. 8.2 показывает графическое представление данных результатов.

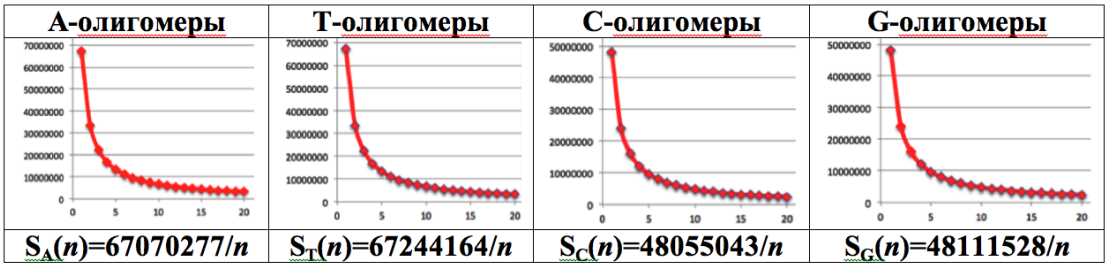


Рис. 8.2. Гиперболические последовательности количеств n-плетов из A-, T-, C- и G-группировок в первых 20 слоях ДНК-текста хромосомы человека №1. По оси абцисс отложены значения n, а по оси ординат соответствующие количества n-плетов (олигомеров длины n).

Аналогичные гиперболические взаимосвязи n-плетных сумм в каждой из A-, T-, C- и G-группировок n-плетов 20 первых слоев геномных ДНК-текстов получены для ДНК-текстов: 1) всех 24 хромосом человека, сильно различающихся своей длиной, а также количеством и видом генов в них (Рис. 8.3); 2) всех хромосом дрозофилы, мыши, червя, многих растений; 3) 19 геномов бактерий и архей; 4) многих экстремофилов, живущих в экстремальных условиях, например, радиации с уровнем в 1000 раз превышающим смертельный для человека. Подробные численные данные анализа этим авторским методом геномных ДНК-текстов эукариотов и прокариотов представлены в препринтах [Petoukhov, 2019-2021]).

Рис. 8.3. 24 хромосомы человека (https://www.ncbi.nlm.nih.gov/genome/51).

Найденные гештальт-правила процентов в геномах и соответствующие гиперболические взаимосвязи количеств *n*-плетов в геномах являются кандидатами на роль универсальных генетических правил. Данные алгебраические инварианты глобально-геномного характера сохраняются неизменными на протяжении миллионов лет биологической эволюции, по ходу которой отмирают миллионы видов живых организмов и возникают новые (хотя локально геномные последовательности изменяются из-за действия мутаций, пресса естественного отбора и пр.). Это затрагивает вопрос о происхождении жизни и информационной связи всех организмов, коль скоро данные алгебраические структуры и правила характеризуют геномы даже бактерий и архей.

Отметим, что гиперболические правила геномов кардинально отличаются от известного гиперболического правила Ципфа, не связанного с кодированием информации и сформулированного в лингвистике для процентов встречи отдельных слов в длинных текстах. Геномные правила олигомерных сумм в многослойном представлении кодирующих ДНК-текстов говорят не о проценте встречи отдельных фрагментов текста, а о суммах множества различных *n*-плетов, распределенных в геномном ДНК-тексте, в котором каждый нуклеотид является составной частью многих олигомерных наборов, существующих одновременно (каждый нуклеотид A, T, C и G является распределенным участником многих членов соответствующих последовательностей сумм *n*-плетов A-, T-, C- и G-группировок одновременно и делает вклад в каждую из них). Кратко говоря, эти правила - геномов и Ципфа - относятся к существенно разным структурам.

Малая часть ДНК-текстов геномов (у человека всего 1,5%) является генами, кодирующими белки (в первой хромосоме человека содержится 4505 генов). Оказывается, что представление ДНК-текстов как многослойных по авторскому методу *n*-плетных сумм в слоях дают важные результаты при анализе не только геномов, но также относительно коротких генов. Например, анализ гена TTN, содержащего 81940 нуклеотидов и кодирующего длинный мышечный белок Титин, выявил закономерные триплет-отклонения в его последовательностях *n*-плетных сумм от модельных гиперболических последовательностей. Эти отклонения происходят при значениях *n* кратных 3 *(n* = 3*m*) (в генетическом коде именно триплеты кодируют аминокислоты белков).

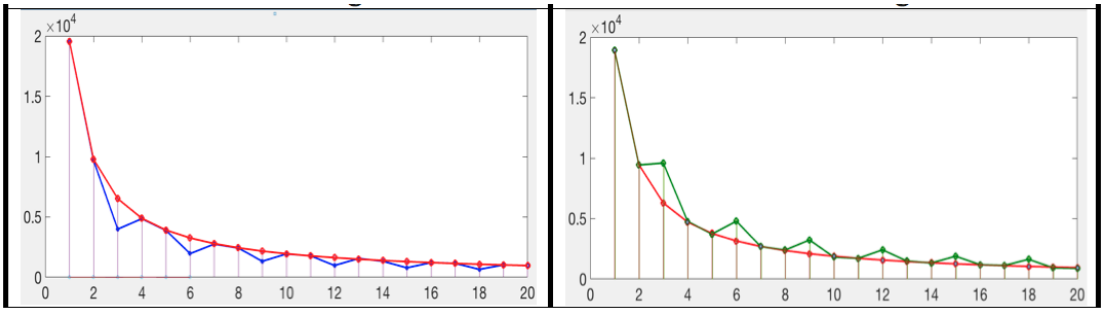


Рис. 8.4. Триплет-отклонения реальных количеств *n*-плетов относительно их модельных гиперболических последовательностей в 20 слоях ДНК-текста гена TTN, кодирующего длинный мышечный белок Титин. Слева показаны данные по группировке T_1-олигомеров, а справа – по группировке G_1-олигомеров (из препринта [Petoukhov, 2021a]).

Эти триплет-отклонения при $n = 3m$ не случайны и также лежат на гиперболах, но закономерно сдвинутых. Поражает регулярность и согласованность триплет-отклонений с величиной до 50% в противоположных направлениях для разных классов *n*-плетов (Рис. 8.5).

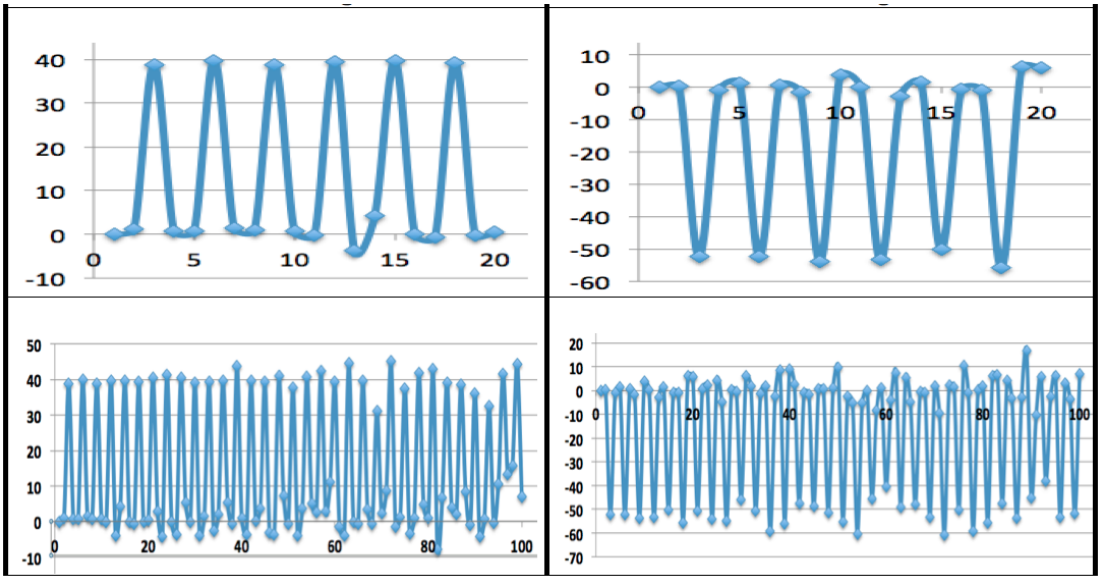


Рис. 8.5. Процентные представления высокорегулярных ритмических последовательностей триплетных отклонений реальных количеств *n*-плетов, которые принадлежат к группировкам T_1-олигомеров (слева) и G_1-олигомеров (справа), от соответствующих модельных гиперболических значений 19569/n и 18901/n в гене TTN мышечного белка Титина. Здесь n = 1, 2, 3,…, 20 (верхний ряд) и n = 1, 2, 3,…, 100 (нижний ряд), как показано по осям абсцисс. Оси ординат показывают процент отклонений от значений модели, принятых за 100% (взято из препринта [Petoukhov, 2021a], содержащего дополнения).

Модельные гиперболические последовательности $S_A(1)/n$, $S_T(1)/n$, $S_C(1)/n$ и $S_G(1)/n$ (Рис. 8.2) выступают математическими стандартами для гиперболических правил сумм количеств *n*-плетов в разных геномах. Эти гиперболические

последовательности отличаются только величиной числителей. Нормировка этих последовательностей делением на величины числителей приводит их к исторически известной **гармонической прогрессии**:

$$1, \quad 1/2, \quad 1/3, \quad 1/4, \quad 1/5, \quad \ldots, \quad 1/n \tag{8.1}$$

Свое имя она получила в силу связи с гармониками в музыке и длинами стоячих волн в колеблющейся струне.

Гармоническая прогрессия (8.1) изучалась Пифагором, Лейбницем, Ньютоном, Эйлером, Фурье, Дирихле, Риманом в связи с разными задачами математики и физики. Ныне гармоническая прогрессия неожиданно проявила себя во всех протестированных геномах, то есть в информационных основах живой материи.

Гармоническая прогрессия со времен Пифагора связывается с гармониками музыки и стоячими волнами в резонаторах. Это позволяет предполагать, что связь гармонической прогрессии с численной фрактальной структурой спиральных ДНК-текстов имеет резонансную природу и отражает существование некой системы стоячих волн. Данное предположение согласуется с гипотезами о том, что спиральные ДНК являются спиральными антеннами электромагнитных волн, а потому, как и другие спиральные хиральные биомолекулы, эмитируют и адсорбируют электро-магнитные волны определенной круговой поляризации, которая обеспечивает биомолекулам возможность обмена радиоволнами избирательной поляризации. Это также согласуется с авторской концепцией мульти-резонансной генетики [Petoukhov, 2016].

Аналог спиральной антенны обнаружен шведскими учеными в хвостовой части сперматозоида. Высказано мнение, что эти спирали выступают в роли камеры видеонаблюдения для обхождения сперматозоидом препятствий и поиска яйцеклетки [Zabeo et al., 2018].

Антенно-подобные спиральные структуры участвуют в биомеханике координированных движений. Например, одноклеточный организм *Mixotricha paradoxa* движется за счет находящихся на его поверхности 250 тысяч спиральных бактерий *Treponema spirochetes*, спиральные жгутики которых согласованно скручиваются как единое целое, обеспечивая целенаправленное движение.

Электрические и механические колебания в живых телах тесно связаны, поскольку многие их структуры являются пьезоэлектриками: нуклеиновые кислоты, актин, дентин, сухожилия, кости, и пр. Тем самым, электро-магнитные феномены сопровождаются феноменами вибрационной механики.

К теме о давно предполагаемой структурной связи живых тел с музыкальной гармонией, сопряженной с гармонической последовательностью (8.1), можно привести следующую цитату из книги [Шноль, 1989, с.75]: «*Из возможных следствий взаимодействия макро-молекул ферментов, осуществляющих [циклические] конформационные колебания, рассмотрим пульсации давления – звуковые волны. Диапазон чисел оборотов большинства ферментов соответствует слышимым звуковым частотам. Некую, пока еще фантастическую картину «музыкальных взаимодействий» биохимических систем, клеток, органов, и возможную роль этих взаимодействий в жизнедеятельности мы рассмотрим подробнее… Это наводит на приятные размышления о природе слуха, происхождении музыкального восприятия и о*

многом другом, что уже принадлежит области биохимической эстетики».

Вопросами сродства генетической гармонии организмов с музыкальной гармонией и ее математикой занимается наша группа в «Центре междисциплинарных исследований музыкального творчества» Московской государственной консерватории им. П.И.Чайковского, существующего с 2012 года (http://www.mosconsv.ru/ru/groups.aspx?id=135683); научный руководитель Центра – декан композиторского факультета проф. А.А.Кобляков). Глубокая связь музыки с математикой давно утверждалась мыслителями. Например, Г. Лейбниц полагал, что «*музыка есть арифметика души, которая вычисляет, сама того не сознавая*».

Характерно, что в живой природе информация на разных уровнях организации передается в виде цепевидных текстов: цепи нуклеотидов ДНК, белковые цепи аминокислот, последовательности импульсов в нейронах, литературные тексты, музыкальные тексты. Геномные ДНК-тексты с их гештальт-правилами могут рассматриваться как метатексты при анализе биологических текстов различного вида.

9. О геометриях, физиологии и искусственном интеллекте.

В отношении генетически наследуемой мозговой деятельности математик и специалист по когнитивным наукам Д.Хофштадтер пишет [Hofstadter, 1980, p. 612]:

«*Каждый математик чувствует, что в математике существует некая метрика, объединяющая идеи, – что вся математика есть сеть результатов, которые соединены между собой огромным количеством связей. ... Имей мы возможность ввести это высокоразвитое ощущение математической близости – «ментальную метрику математика»… – в программу, мы могли бы создать примитивного «искусственного математика*».

В.В.Налимов, заместитель А.Н.Колмогорова по лаборатории теории вероятностей мехмата МГУ, указывает в книге [Налимов, 2015, §4]: «*В связи с проблемой искусственного интеллекта обострился интерес к представлению о метрике пространств математического мышления. ... Искусственный интеллект мог бы быть сближен с математическим мышлением, если бы оказалось возможным осознать метрические свойства пространства мышления человека…. мы готовы идти дальше и говорить, что само сознание структурировано геометрически: экзистенциально человек геометричен. … Желая усилить аргументацию этого утверждения, мы обратим здесь внимание на геометрическую обусловленность зрительного восприятия. Наше зрите-льное восприятие – это не автоматическое перенесе-ние внешнего Мира в наше сознание, а его сложное воспроизведение, отвечающее определенным геометри-ям. Эта тема обстоятельно рассмотрена в книге [Петухов, 1981] …. Несколько перефразируя суждение Петухова, мы можем сказать, что в нашем сознании при построении текстов, через которые мы воспринимаем Мир, происходит что-то очень похожее на то, что происходит в морфогенезе. Мы готовы увидеть в глубинах сознания те же геометрические образы, которые раскрываются в морфогенезе».*

Налимов также добавляет в связи со сказанным о геометричности живых тел: «*Опираясь на приведенные в книге [Петухов, 1981] материалы, отметим прежде всего, что еще в сороковых годах Р. Лунебург [Luneburg, 1947, 1948, 1950] высказал экспериментально обоснованное утверждение о том, что пространство зрительного восприятия у человека характеризуется геометрией Лобачевского. Позднее это высказывание нашло широкий и благоприятный отклик. Особенно обстоятельная проверка концепции Лунебурга была осуществлена в работе [Kienle, 1964] в 1960-х годах. Серьезное применение геометрический подход нашел в цветоведении. Известный физик Э. Шрёдингер [Schrödinger, 1920], занимавшийся и теорией зрения, опирался на представления проективной геометрии при изучении физиологических законов смешения цветов. Г. фон Шеллинг [von Shelling, 1955; 1956 a, b; 1960; 1964] ввел неевклидову метрику для описания цветовосприятия и построил перцепционную теорию относительности по аналогии с представлением о пространственно-временном многообразии в специальной теории относительности. Г. Резников [Reznikoff, 1974] обратился к дифференциальной геометрии при изучении цветовосприятия. Даже из такого беглого обзора следует, что зрительное восприятие Мира – это его отображение через тексты нашего сознания, построенные на совсем нетривиальных геометриях*».

Вопросам применения неевклидовых геометрий и алгебры Клиффорда для моделирования специфической организации работы мозга посвящены публикации [Labunets, 2003; Labunets, Rundblad, Astola, 2002].

Обнаружение описанных выше закономерностей связи генетической системы с широким множеством бинарных тетрагруппировок вызывает ассоциации с концепцией Ю.И. Кулакова о бинарных эйдосах (черных и белых, мужских и женских) как основе законов физики и геометрии [Кулаков, 1968]. Продолжение работ Ю.И.Кулакова и его соратника Г.Г.Михайличенко в бинарной геометрофизике Ю.С.Владимирова [Владимиров, 2009, 2020, 2021] связано с важным понятием предгеометрии в теории систем отношений. По отношению к этой предгеометрии понятия пространства и времени являются вторичными. Очень похоже на то, что генетическая система с ее первичными системами бинарных отношений построена именно на их предгеометрии. При таком понимании из этой бинарной генетической предгеометрии должны вытекать в качестве вторичных биологические время и пространство организмов. К теме бинарных отношений в генетике напомним, что физиология активных тканей строится на фундаментальном бинарном законе «все или ничего»: нервная клетка и мышечная единичка дают ответы только «да» или «нет» на действия различных стимулов. На подпороговые стимулы они не реагируют, а на надпороговые реагируют полной амплитудой. Крокодильчики и черепашки, вылупившись из яйца и не имея опыта работы с пространством и временем, сразу ползут к воде вполне координированными движениями за счет генетически наследуемой согласованной активности миллионов их бинарно работающих нервных и мышечных клеток, напоминающей согласованную работу триггеров при выполнении компьютерных программ. Вопросы такой наследуемой согласованности на основе гено-логического кодирования в свете задач алгебраической биологии рассматриваются в статье [Petoukhov, Petukhova, 2017].

10. Теория Фрелиха о коллективных квантовых эффектах в биологических системах по аналогии с квантовым конденсатом Бозе-Эйнштейна.

Изложенные выше и другие результаты авторского исследования системы генетического кодирования свидетельствуют о наличии дальнодействующих связей в составе геномных ДНК-текстов, а также о квантовой информационной сущности генетической системы и живых организмов в целом [Петухов, 2020a,b,c]. Автор полагает, что эти результаты связаны с популярной колебательно-резонансной теорией Г. Фрёлиха о дальнодействующей когерентности в биологических системах, то есть о коллективных квантовых эффектах в биологических системах [Frohlich, 1978, 1980, 1988; Frohlich, Kremer, 1983].

Нобелевский лауреат Р. Пенроуз использует теорию Фрёлиха для обоснования своих представлений о мозге как квантовом компьютере и пишет о нем следующим образом [Пенроуз, 2005, с. 540, 561, 568]. Выдающийся физик Герберт Фрёлих (который в 1930-х годах совершил один из фундаментальных прорывов в понимании низкотемпературная сверхпроводимость) предположил возможную роль коллективных квантовых эффектов в биологических системах. В 1968 году Фрёлих предложил полагать, что внутри активных клеток должны быть вибрационные эффекты, которые будут резонировать с микроволновым электромагнитным излучением с частотой 10^{11} Гц в результате феномена биологической квантовой когерентности. Вместо того, чтобы требовать низкой температуры, эффекты возникают из-за наличия большой энергии метаболического влечения. В настоящее время существуют некоторые достоверные данные наблюдений во многих биологических системах, подтверждающие именно тот эффект, который Фрелих предсказал в 1968 году. Он утверждал, что пока энергия метаболического влечения достаточно велика, а диэлектрические свойства соответствующих материалов являются достаточно экстремальными, то существует возможность крупномасштабной квантовой когерентности, аналогичной той, которая возникает в явлениях сверхпроводимости и сверхтекучести - иногда называемых конденсацией Бозе-Эйнштейна - даже при относительно высоких температурах, присутствующих в биологических системы. В конденсате Бозе-Эйнштейна (который также возникает при действии лазера) большое количество частиц коллективно участвует в одном квантовом состоянии. Для этого состояния существует волновая функция, подходящая для отдельной частицы, но теперь она применяется сразу ко всей совокупности частиц, которые участвуют в этом состоянии. В конденсате Бозе-Эйнштейна вся система, содержащая большое количество частиц, ведет себя как единое целое по аналогии с квантовым состоянием отдельной частицы, за исключением того, что все масштабируется соответствующим образом. Существует согласованность в большом масштабе, где многие странные особенности квантовых волновых функций реализуются на макроскопическом уровне.

Ф. Фрёлих (сын Герберта Фрёлиха) написал статью «Генетический код как язык» о состояниях квантовой когерентности и дальнодействующей коммуникации в геномах. В ней он отметил: «*Помимо хромосомы, геном в целом должен содержать какую-то связь на большом расстоянии, чтобы не производить на одном участке антигены, продуцируемые на других участках. Таким образом, существует сложность сущностей, которые по-разному используют генетический язык. Можно сказать, что существует логика клеток*» [Fröhlich F., 1988, p. 194]. Гипотеза Г. Фрёлиха «*предполагает, что дальнодействующие когерентные колебания приведут к резонансу между*

дифференцированной клеткой с ее собственными характерными колебаниями и хромосомой, так что специфическая для хромосомы область, реагирующая на эту характеристическую частоту, будет активирована или открыта, чтобы она могла производить соответствующие белки. Такой резонанс мог бы транспортировать эмбриологические, уже частично индуцированные клетки, к их мишени, и там они будут дополнительно зафиксированы для производства правильных белков для этого органа посредством наложенного резонанса, и поэтому, если бы он был слишком зрелым, он не мог бы так адаптироваться. Степени резонанса могут быть разными» [там же, с. 199]. «*Это было сделано более конкретно в случае деления клеток во время митоза, где была выдвинута гипотеза о том, что соответствующие хромосомы выстраиваются в линию посредством резонанса с одинаковой частотой, обнаруживая друг друга из очевидной путаницы генома посредством связи на большие расстояния* [Holland, 1972]. *Резонансные колебания притягивают подобное к подобному. В более общем плане может существовать какое-то связное дальнодействующее взаимодействие между членами генома, создающее самомаркировку. Если это ломается в результате последующих мутаций, возникают аутоиммунные заболевания*» [там же, с. 203]. Данные цитаты из статьи Ф. Фрёлих о роли резонансов в геномах согласуются с концепцией мультирезонансной генетики в статье [Petoukhov, 2016] и описанными выше авторскими данными о геномных гештальт-правилах.

Синхронные крупномасштабные коллективные колебания Фрёлиха подразумевают межклеточное микроволновое излучение, которое будет представлять собой нехимическое и нетепловое взаимодействие между клетками [Vasconcellos et al, 2012]. Сообщалось о некоторых доказательствах нетеплового влияния когерентного микроволнового излучения на конформационное состояние генома у E. coli [Hyland, 1998], которые могут указывать на то, что «*хромосомная ДНК может быть целью миллиметрового микроволнового излучения внутри этой системы. Кроме того, низкоинтенсивное микроволновое облучение лейкоцитов приводит к значительному увеличению биофотонного излучения в оптическом диапазоне, источником которого считается ДНК. Также стоит отметить возможное влияние концепции биокогерентности на очень конкретную диполярную систему, которой является вода. Можно рассматривать возможность того, что биологическая вода сама может поддерживать когерентные диполярные возбуждения, распространяющиеся на мезоскопические области; таким образом, вода, вместо пассивного растворителя, заполняющего пространство, могла бы занять важное особое положение, полное значение которого еще не выяснено*» [Vasconcellos et al, 2012].

Живые организмы на Земле подвергаются воздействию целого спектра модальностей электромагнитного излучения, связанных с физическим строением Земли и положением нашей планеты в солнечной и галактической системах. В статье [Geesink, Meijer, 2016] представлены результаты обширного обзора литературы, включающего 175 статей с 1950 по 2015 год, посвященных влиянию электромагнитного излучения на живые системы in vitro и in vivo. Эти результаты указывают на то, что некоторые типичные дискретные когерентные частоты электромагнитных волн способны стабилизировать клетки, тогда как другие вызывают явную дестабилизацию. Эти авторы пишут: «*Мы находим поддержку в отношении гипотезы Г. Фрёлиха о том, что управляемый набор осцилляторов конденсируется в широком диапазоне энергий и может*

активировать колебательный режим в живых организмах при комнатной температуре. Принимая во внимание поддерживающие жизнь частоты, взятые из литературы, был выведен алгоритм когерентных частот стоячих волн для устойчивости биологического порядка. Интересно, что мы обнаружили, что к происхождению конкретного биологического алгоритма можно математически приблизиться с помощью выбранной эталонной «темперированно пифагорейской» акустической шкалы. Этот алгоритм выражает одномерные волновые уравнения, известные для колеблющихся струн». Эти результаты, сопряженные со стоячими волнами, можно рассматривать как имеющие тесную корреляцию с описанными выше результатами автора о связи гармонической прогрессии (8.1) с геномами высших и низших организмов.

Кроме того, статья [Geesink, Meijer, 2016] напоминает, что Г. Фрёлих заявлял о колебательных зарядах в термальной ванне, в которой большое количество квантов может конденсироваться в одно состояние, известное как Бозе-конденсат, который может составлять физический и нетепловое взаимодействие между клетками. Эти авторы отмечают, что связь между «живой» и неживой материей также была установлена Вольфгангом Паули. Он заявил, что ментальная и материальная области регулируются общими принципами упорядочения и могут быть поняты как «*взаимодополняющие аспекты одной и той же реальности*» [Pauli, 1994]. Другими словами, аналогию с конденсатом Бозе-Эйнштейна можно распространить и на мысленную область.

Развитие представлений о живых телах как особых аналогах квантового конденсата Бозе-Эйнштейна может привести к новому пониманию древних представлений восточных культур об энергии ци (или ци), циркулирующей в живых телах. Дополнительное Приложение в конце данной статьи посвящено связям генетической системы с древними доктринами восточной культуры, отображенными в схемах Инь-Ян древнекитайской книги «И Цзин» и 4-секционных мандалах буддистов, индуистов и представителей других религий.

Заключение.

Задачи развития систем искусственного интеллекта, воспроизводящих в той или иной степени свойства естественного интеллекта человека и животных, требуют изучения генетически наследуемой способности живых организмов к интеллектуальной деятельности [Петухов, 2021; Петухов, Толоконников, 2020]. Представленные в статье результаты анализа системы молекулярно-генетического кодирования отражают стремление автора понять алгебраические особенности закономерно структурированной системы генетического кодирования.

Описанные закономерности процентных сумм *n*-плетов в слоях ДНК-текстов геномов высших и низших организмов согласуются с утверждением П.Йордана, автора термина “квантовая биология”, о том, что пропущенные наукой законы жизни являются правилами вероятностей квантового мира [Jordan, 1932; McFadden, Al-Khalili, 2018]. Полученные автором результаты показывают существование ранее неизвестных генетических закономерностей. Эти результаты были получены на основе новых методов анализа и моделирования ДНК-текстов, связанных с математическими формализмами квантовой механики и квантовой информатики, алгебрами многомерных

гиперкомплексных чисел, метрическим анализом и теорией помехоустойчивого кодирования информации. Изложенные результаты могут быть использованы, в частности, для углубления наши знаний о мозговой деятельности.

Учитывая слова Йордана и Шредингера о диктаторской роли структурированной информатики генетических молекул для всего организма [McFadden, Al-Khalili, 2018], естественно думать, что структурные особенности ДНК-информатики накладывают свой отпечаток на все генетически унаследованные биологические системы и явления. Это согласуется с тем фактом, что все физиологические системы должны быть структурно согласованы с генетическим кодированием, чтобы передаваться в генетически закодированной форме потомству. Это также согласуется с точкой зрения, что основная задача живых организмов - передача генетической информации по цепи поколений. Описанные алгебро-генетические результаты свидетельствуют о том, что система генетического кодирования основана на методах вероятностного кодирования, которые сейчас изучаются в лаборатории автора.

Шредингер писал: "«*Для организма следует ожидать новых законов. ... все известное нам о структуре живой материи заставляет ожидать, что деятельность живого организма нельзя свести к проявлению обычных законов физики ... потому, что его структура отличается от всего изученного нами до сих пор в физической лаборатории*» [Шредингер, 1972 с. 77].

Для сравнения, ферменты в живых организмах работают в миллионы раз эффективнее, чем лучшие катализаторы в лабораториях. То, что делает фермент в живом теле за 1 секунду, катализатор в лаборатории может сделать только за 100 тысяч лет. Насколько мы понимаем, найденные гештальт-правила «диктаторских» ДНК-текстов не выводятся из известных законов физики, а потому относятся к тем законам структуризации живого, о которых говорил Шредингер, отмечая дополнительно следующее: «*Именно со статистической точки зрения структура важнейших частей живого организма полностью отличается от структуры любого вещества, с которым мы, физики и химики, имели до сих пор дело практически в наших лабораториях и теоретически за письменным столом*» [Шредингер, 1972, с. 14].

Эта статья посвящена алгебраическим гештальт-феноменам статистического характера в генетике, то есть генетическим феноменам, в которых реализуется один и тот же целостный паттерн для множества различных наборов его составных частей. Попутно можно отметить, что физика и другие науки основаны на открытии присущих им гештальт-феноменов. Например, закон Кулона в физике показывает существование устойчивой картины силового взаимодействия в системах, состоящих из различных наборов заряженных частиц. Другими словами, человеческое познание мира происходит через открытие в нем гештальт-феноменов.

В целом представленные результаты авторских исследований структурных закономерностей генетической информатики свидетельствуют об эффективности модельного подхода к живым организмам как квантово-информационным алгебро-гармоническим сущностям на модульных принципах. Исследования по алгебраической биологии и матричной генетике раскрывают связи генетики с алгебрами многомерных чисел и геометриями, которые полезны для совершенствования систем искусственного интеллекта.

Системы искусственного интеллекта можно пытаться создавать на десятках разных путей с разной степенью плодотворности разработок. Но здесь уместно вспомнить слова Декарта в его "Трактате о животном": "*Нет более*

плодотворного занятия чем познание самого себя".

Желая усилить участие научного сообщества в развитии алгебраической биологии, напомним еще известное утверждение Тейяра де Шардена о том, что человек как «*предмет познания*» - «*это ключ ко всей науке о природе*» [Тейяр де Шарден, 1987].

Выводы.

1) Композитные ДНК-тексты геномов обладают описанными универсальными алгебраическими особенностями и неизвестными ранее симметриями;
2) Эти алгебраические особенности геномов позволяют развивать гештальт-генетику и выявлять структурные связи генетики с формализмами квантовой информатики;
3) Полученные результаты позволяют рассматривать живые организмы как квантово-информационные алгебро-гармонические сущности на блочных принципах;
4) Появляются новые подходы к анализу биологических текстов различного вида по аналогии с геномными;
5) Согласно полученным данным, биоинформатика базируется на вероятностных методах многослойного кодирования и обработки информации, формирующих из вероятностных последовательностей закономерные целостные образы (гештальт паттерны).
6) Результаты алгебраической биологии и матричной генетики желательно учитывать при совершенствовании систем искусственного интеллекта.

Приложение I. **Пример гештальт правил суммарных процентов *n*-плетов в бактериальном геноме**.

В этом Приложении представлены данные о процентном содержании *n*-плетов в штамме ризобактерий Bradyrhizobium japonicum E109, полный геном, длиной 9224208 пар нуклеотидов (исходные данные взяты из https://www.ncbi.nlm.nih.gov/nuccore/CP010313.1?report = genbank (длина этой последовательности более чем в 20 раз короче длины выше рассмотренной последовательности хромосомы человека №1). Рис. ПI.1-ПI.5 показывают матрицы феноменологических процентов 4 нуклеотидов, 16 дублетов, 64 триплетов и 256 тетраплетов в соответствующих слоях этой геномного ДНК-текста. Эти процентные матрицы для бактерий и процентные матрицы (рис. 2.2-2.5) для хромосомы человека №1 имеют сходные свойства для *m*-позиционных тетрагруппировок *n*-плетов, которые выражаются равенствами (3.1) и (3.2) .

C	A	➔	%C	%A	=	0.3184	0.1819
T	G		%T	%G		0.1815	0.3182

Рис. ПI.1. Преобразование символьной матрицы 4 нулеотидов в численную матрицу процентов 4 нуклеотидов в случае ДНК-текста генома *Bradyrhizobium japonicum* strain E109.

%CC	%CA	%AC	%AA
%CT	%CG	%AT	%AG
%TC	%TA	%GC	%GA
%TT	%TG	%GT	%GG

=

0.0781	0.0587	0.0467	0.0367
0.0515	0.1302	0.0469	0.0517
0.0720	0.0145	0.1216	0.0719
0.0365	0.0584	0.0468	0.0779

Рис. III.2. Матрица феноменологических процентов 16 дуплетов в геномной ДНК последовательности дуплетов *Bradyrhizobium japonicum* strain E109 (сравнить с Рис. 2.3).

%CCC	%CCA	%CAC	%CAA	%ACC	%ACA	%AAC	%AAA
%CCT	%CCG	%CAT	%CAG	%ACT	%ACG	%AAT	%AAG
%CTC	%CTA	%CGC	%CGA	%ATC	%ATA	%AGC	%AGA
%CTT	%CTG	%CGT	%CGG	%ATT	%ATG	%AGT	%AGG
%TCC	%TCA	%TAC	%TAA	%GCC	%GCA	%GAC	%GAA
%TCT	%TCG	%TAT	%TAG	%GCT	%GCG	%GAT	%GAG
%TTC	%TTA	%TGC	%TGA	%GTC	%GTA	%GGC	%GGA
%TTT	%TTG	%TGT	%TGG	%GTT	%GTG	%GGT	%GGG

=

0.0131	0.0154	0.0136	0.0129	0.0145	0.0083	0.0097	0.0068
0.0148	0.0344	0.0145	0.0185	0.0053	0.0182	0.0075	0.0126
0.0174	0.0040	0.0446	0.0328	0.0205	0.0043	0.0205	0.0106
0.0126	0.0183	0.0185	0.0343	0.0075	0.0144	0.0052	0.0146
0.0149	0.0133	0.0044	0.0021	0.0353	0.0210	0.0195	0.0153
0.0105	0.0323	0.0043	0.0039	0.0210	0.0443	0.0209	0.0171
0.0151	0.0020	0.0206	0.0133	0.0195	0.0044	0.0356	0.0150
0.0069	0.0125	0.0082	0.0153	0.0098	0.0133	0.0148	0.0133

Рис. III.3. Матрица феноменологических процентов 64 триплетов в геномной ДНК последовательности триплетов *Bradyrhizobium japonicum* strain E109 (сравниить с Рис. 2.4).

.0023	.0025	.0033	.0030	.0052	.0017	.0035	.0021	.0022	.0037	.0018	.0019	.0023	.0023	.0016	.0015
.0026	.0057	.0038	.0055	.0014	.0051	.0027	.0045	.0029	.0058	.0024	.0022	.0015	.0036	.0019	.0018
.0048	.0011	.0119	.0082	.0066	.0011	.0081	.0030	.0014	.0006	.0057	.0050	.0025	.0010	.0045	.0028
.0039	.0052	.0048	.0096	.0021	.0044	.0018	.0052	.0014	.0019	.0029	.0047	.0018	.0022	.0015	.0039
.0036	.0024	.0016	.0004	.0141	.0072	.0094	.0057	.0042	.0044	.0010	.0006	.0053	.0043	.0026	.0026
.0020	.0090	.0015	.0004	.0073	.0162	.0085	.0091	.0035	.0085	.0012	.0016	.0039	.0072	.0035	.0019
.0056	.0005	.0068	.0037	.0082	.0016	.0173	.0055	.0024	.0005	.0048	.0034	.0018	.0006	.0058	.0036
.0019	.0043	.0022	.0055	.0035	.0051	.0059	.0056	.0019	.0027	.0024	.0039	.0015	.0013	.0030	.0027
.0027	.0031	.0035	.0031	.0012	.0009	.0006	.0005	.0059	.0062	.0048	.0047	.0058	.0036	.0042	.0029
.0035	.0057	.0034	.0036	.0006	.0016	.0005	.0005	.0059	.0173	.0049	.0067	.0018	.0081	.0025	.0057
.0040	.0010	.0103	.0092	.0017	.0004	.0013	.0010	.0069	.0013	.0167	.0102	.0098	.0017	.0070	.0040
.0027	.0030	.0050	.0082	.0011	.0011	.0006	.0010	.0045	.0080	.0058	.0118	.0024	.0066	.0014	.0047
.0036	.0031	.0007	.0004	.0065	.0033	.0035	.0031	.0037	.0036	.0010	.0007	.0094	.0065	.0037	.0035
.0026	.0058	.0006	.0004	.0043	.0069	.0044	.0024	.0026	.0095	.0010	.0016	.0053	.0141	.0041	.0036
.0028	.0005	.0047	.0030	.0036	.0009	.0061	.0032	.0040	.0005	.0048	.0034	.0059	.0011	.0059	.0027
.0015	.0021	.0020	.0029	.0024	.0016	.0036	.0026	.0016	.0034	.0018	.0033	.0023	.0052	.0021	.0024

Рис. П.4. Матрица феноменологических процентов 256 тетраплетов в геномной ДНК последовательности тетраплетов у *Bradyrhizobium japonicum* strain E109 (сравнить с Рис. 2.5).

%C ≈ 0.3184	%G ≈ 0.3182	%A ≈ 0.1819	%T ≈ 0.1815
Σ%CN ≈ 0.3185	Σ%GN ≈ 0.3182	Σ%AN ≈ 0.1820	Σ%TN ≈ 0.1814
Σ%NC ≈ 0.3183	Σ%NG ≈ 0.3182	Σ%NA ≈ 0.1818	Σ%NT ≈ 0.1816
Σ%CNN ≈ 0.3197	Σ%GNN ≈ 0.3201	Σ%ANN ≈ 0.1806	Σ%TNN ≈ 0.1797
Σ%NCN ≈ 0.3166	Σ%NGN ≈ 0.3173	Σ%NAN ≈ 0.1836	Σ%NTN ≈ 0.1825
Σ%NNC ≈ 0.3189	Σ%NNG ≈ 0.3173	Σ%NNA ≈ 0.1815	Σ%NNT ≈ 0.1824
Σ%CNNN ≈ 0.3184	Σ%GNNN ≈ 0.3183	Σ%ANNN ≈ 0.1816	Σ%TNNN ≈ 0.1817
Σ%NCNN ≈ 0.3184	Σ%NGNN ≈ 0.3185	Σ%NANN ≈ 0.1816	Σ%NTNN ≈ 0.1815
Σ%NNCN ≈ 0.3185	Σ%NNGN ≈ 0.3181	Σ%NNAN ≈ 0.1823	Σ%NNTN ≈ 0.1811
Σ%NNNC ≈ 0.3182	Σ%NNNG ≈ 0.3180	Σ%NNNA ≈ 0.1820	Σ%NNNT ≈ 0.1818

Рис. П.5. Проценты нуклеотидов C, G, A, T и сумм процентов *n*-плетов с этими нуклеотидами на их атрибутивных позициях в соответствующих тетра-группировках *n*-плетов в случае геномной ДНК последовательности у *Bradyrhizobium japonicum* strain E109 (сравнить с Рис. 3.4).

Что касается проекторных тетрагруппировок *n*-плетов в разных слоях геномной ДНК штамма E109 Bradyrhizobium japonicum, то, как было отмечено выше в выражениях (7.9) и (7.10), следующие равенства выполняются во всех 7 мозаичных матрицах дублетов и триплетов на рис. 7.3:

$$\Sigma\%\heartsuit \approx \Sigma\%\diamondsuit \approx 0.277, \qquad \Sigma\%\spadesuit \approx \Sigma\%\clubsuit \approx 0.223$$

Например, в случае мозаичной (4*4)-матрицы с тетрагруппировкой №1 для 16 дублетов (см. Рис. 7.3) для этой геномной ДНК появляются следующие процентные суммы, основанные на процентах дублетов на Рис. П.2:

$\Sigma\%\heartsuit \approx \%CA + \%CG + \%AA + \%AG \approx 0.0587 + 0.1302 + 0.0367 + 0.0517 \approx 0.277;$
$\Sigma\%\diamondsuit \approx \%TC + \%TT + \%GC + \%GT \approx 0.0720 + 0.0365 + 0.1216 + 0.0468 \approx 0.277;$
$\Sigma\%\spadesuit \approx \%CC + \%CT + \%AC + \%AT \approx 0.0781 + 0.0515 + 0.0467 + 0.0469 \approx 0.223;$
$\Sigma\%\clubsuit \approx \%TA + \%TG + \%GA + \%GG \approx 0.0145 + 0.0584 + 0.0719 + 0.0779 \approx 0.223.$

Приложение II. О Инь-Ян схемах древнекитайской “Книги перемен” и 4-секционных мандалах в древних традициях.

В этом Приложении описаны аналогии между мозаичными генетическими матрицами на Рис. 7.9-7.11, которые связаны с описанными генетическими гештальными феноменами и которые представляют расположение *m*-позиционных тетрагрупп из 64 триплетов, и знаменитой таблицей 64 гексаграмм из древнекитайской “Книги (циклических) перемен” (“И-Цзин”).

Система генетического кодирования содержит бинарно-оппозиционные структуры на разных уровнях своей организации. Как известно, древняя китайская «Книга перемен», написанная несколько тысяч лет назад,

использовала систему символов Инь и Ян (эквиваленты чисел 0 и 1). Эта книга оказала мощное влияние на культуру, медицину и науку древнего Китая и ряда других стран. Система И-Цзин представлена схемами с 4 биграммами, 8 триграммами и 64 гексаграммами. Подобно этому, генетический код построен на молекулах ДНК с использованием 4 азотистых оснований, 16 дублетов и 64 триплетов. Структурные аналогии инь-янских схем И-Цзин с алфавитами ДНК давно отмечались различными авторами, в том числе видными генетиками: Стентом Г.С. в 1969 г. и Нобелевским лауреатом Ф. Джейкобом в 1977 г. Подробнее о таких аналогиях можно прочитать в публикациях [Петухов, 2008; Hu, Petoukhov, Petoukhova, 2017; Petoukhov, 1999; Petoukhov, He, 2010].

И-Цзин посвящено большое количество литературных источников. Многие из этих источников называют И-Цзин одним из величайших и самых загадочных творений человека. С точки зрения китайской культуры (самой древней из всех культур, продолжающих свое существование на Земле) И-Цзин представляет собой нечто еще более значительное: творение Сверхчеловека, воплотившего тайну Вселенной в специальных символах и знаках. Символы и принципы И-Цзин проникли во все сферы жизни традиционного Китая, от теоретических представлений и высокого искусства до предметов быта и украшений. Конфуций сказал: «*Если бы можно было продлить свои годы, я бы отдал пятьдесят из них на изучение И-Цзин*» [Щуцкий, 1997]. И-Цзин провозглашает универсальность циклического принципа организации в природе. Традиционная восточная медицина с ее акупунктурой, тибетской диагностикой по пульсу и пр. основана на положениях, изложенных в этой книге.

Многие западные ученые изучали и использовали И-Цзин. Например, создатель аналитической психологии К. Юнг разработал свое учение о коллективном бессознательном в связи с этой книгой. По словам Юнга и его соратника, лауреата Нобелевской премии по физике В. Паули, триграммы и гексаграммы И-Цзин «*фиксируют универсальный набор архетипов (врожденных психических структур)*» [Щуцкий, 1997, с. 12]. Нильс Бор выбрал символ Инь-Ян в качестве своего герба. Многие современные физики, ощущающие единство мира, связывают свои теории с идеями традиционной восточной культуры, объединяющей всю природу. Например, это отражено в названии «*восьмеричный путь*» известной книги [Gell-Mann, Ne'eman, 1964]. Интенсивное развитие самоорганизующейся и нелинейной динамики сложных систем способствует усилению внимания западных ученых к традиционному восточному мировоззрению (например, см. [Капра, Луизи, 2000]). Специальные группы изучают И-Цзин во многих восточных и западных университетах. Несколько конференций «И-Цзин и современные науки» были проведены в Москве. Подобным исследованиям посвящено огромное количество сайтов в Интернете. Влияние И-Цзин широко представлено в современной жизни восточных стран. Например, на государственном флаге Южной Кореи изображены триграммы. Во многих специализированных школах преподаватели обучают студентов методическим аспектам практического применения актуальных древних знаний в медицине и других областях.

Согласно древним китайцам, «*триграммы, гексаграммы и их компоненты во всех возможных комбинационных комбинациях образуют универсальную иерархию классификационных схем. Эти схемы в визуальных образцах охватывают любые аспекты реальности - пространственные части, временные интервалы, элементы, числа, цвета, органы тела, социальные и*

семейные условия и т. д.» [Щуцкий, 1997, с. 10]. Древние китайцы утверждали, что эта таблица из 64 гексаграмм (Рис. ПП.1) является универсальным архетипом природы. Древние китайцы ничего не знали о генетическом коде, но генетический код во многих аспектах устроен в соответствии с И-цзин.

Следует отметить, что И-Цзин был написан за несколько тысяч лет до появления современных Академий наук. Эта книга представляет собой набор утверждений неясного происхождения. С точки зрения современной науки книга обосновывает универсальное соответствие ее утверждений устройству природы без их соответствующего обоснования. Исторически сложилось так, что западная академическая наука и ее научные законы (например, законы Ньютона) развивались без какой-либо связи с И-Цзин на основе другой методологии, хотя состояния озарения, в которых человек внезапно получает полную картину ответа на его вопросы, хорошо известны и в западной науке.

В И-Цзин прерывные и непрерывные (сплошные) линии символизируют Инь и Ян соответственно. Основная таблица в И-Цзин - знаменитая квадратная таблица из 64 гексаграмм в порядке Фу-Си (Рис. ПП.1), которую китайская традиция считает универсальным природным архетипом. Каждая гексаграмма представляет собой стопку из шести прерывных и непрерывных линий. Согласно западной традиции, эти прерывные и непрерывные линии показаны в виде двоичных символов «0» и «1», и каждая гексаграмма представляет собой последовательность таких шести двоичных символов. На Рис. ПП.1 показаны гексаграммы в обеих этих формах. Каждая позиция во всех гексаграммах имеет свой индивидуальный номер: в западном числовом представлении позиции двоичных символов гексаграмм нумеруются слева направо числами от 1 до 6; в китайском графическом представлении нумерация линий каждой гексаграммы читается в последовательности снизу вверх.

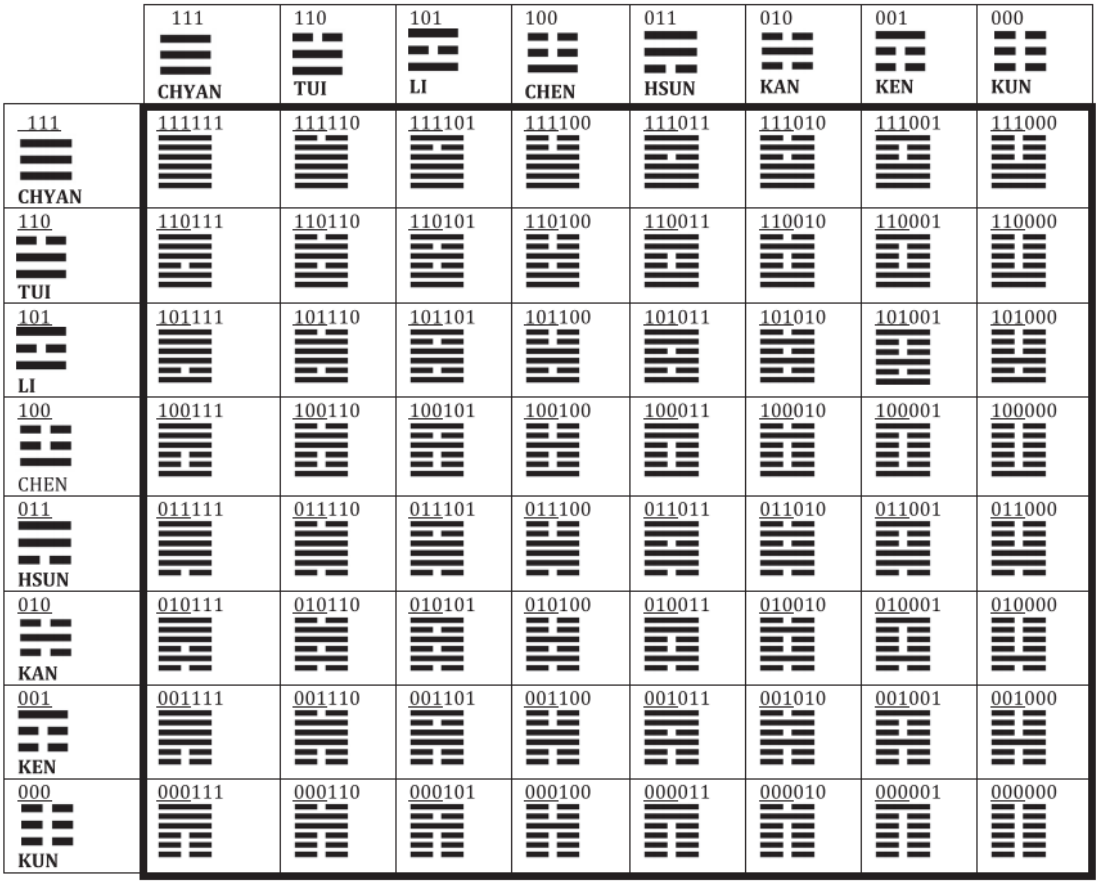


Рис. ПП.1. Таблица 64 гексаграмм в порядке Фу-Си, где каждая гексаграмма показана в символической китайской традиции, а также в двоичном европейском представлении: каждая сплошная или прерывная линия соответствует цифрам 1 или 0. Также показаны китайские названия триграмм.

В таблице 64 гексаграмм (Рис. ПП.1) 8 триграмм (трехразрядные двоичные числа или стопки из трех черт) обозначают ее строки и столбцы. Каждая из гексаграмм представляет собой соединение двух триграмм: триграммы, нумерующей ее строку, и триграммы, нумерующей ее столбец. Следует подчеркнуть, что именно по аналогичной схеме нумерации столбцов и строк троичными цифровыми символами автором построена (8*8)-матрица 64 триплетов генетического кода на Рис. 2.2 в данной статье. По этой аналогии каждый из 64 генетических триплетов на Рис. 2.2 может быть обозначен 6-битным двоичным числом; например, триплет CAC обозначается двоичным числом 111101 [Петухов, 2008; Petoukhov, He, 2010]. Традиционная китайская точка зрения на гексаграммы гласит: «*Гексаграммы - это не триграммы, сплавленные вместе, а две триграммы, расположенные вертикально одна над другой*» [Щуцкий, 1997, с. 101].

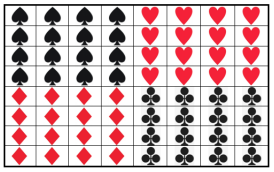
Изучим некоторые аналогии между таблицей 64 гексаграмм (Рис. ПП.1) и генетической мозаикой (8*8)-матриц (Рис. 6.1-6.3), которые представляют расположение в них *m*-позиционных тетрагруппировок 64 триплетов. Ключевым элементом для выявления этих аналогий является следующая древняя китайская доктрина о важной роли пар корреляционных позиций 1-4, 2-5 и 2-6 в китайских гексаграммах: «*Теория И-Цзин считает, что дно триграмма касается внутренней жизни… а верхняя триграмма относится к*

внешнему миру…. Подобные позиции в верхней триграмме и в нижней триграмме имеют самое близкое отношение друг к другу. Из-за этого первая позиция относится по аналогии к четвертой позиции, вторая позиция - к пятой позиции, а третья позиция по аналогии с шестой позицией…. Если эти корреляционные позиции (1-4, 2-5, 3-6) заняты разными линиями, считается, что между ними «существует соответствие», а в случае, когда эти корреляционные позиции заняты идентичными линиями, считается считал, что между ними «нет соответствия»» [Щуцкий, 1997, с. 86].

Рис. ПII.2-ПII.4 показывают таблицу 64 гексаграмм (из Рис. ПII.1) с европейскими обозначениями гексаграмм двоичными числами. Пары двоичных цифр, которые находятся в гексаграммах в соответствующих корреляционных позициях 1–4, 2–5 и 3–6, выделены 4 цветами: синий цвет соответствует паре цифр 1 и 1; красный цвет соответствует паре чисел 1 и 0; желтый соответствует паре цифр 0 и 1; зеленый цвет соответствует паре чисел 0 и 0. Эти 4 пары цифр, стоящие в указанных позициях, определяют принадлежность каждой гексаграммы к той или иной группировке гексаграмм в их соответствующей тетрагруппировке 64 гексаграмм:

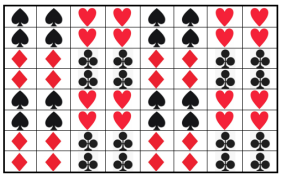
- 1-позиционная тетрагруппировка гексаграмм и 16 членов каждой из ее 4 группировок определяются по корреляционным позициям 1-4 гексаграмм, на которых расположены указанные атрибутивные пары чисел 0-0, 0-1, 1-0 и 1-1 (Рис. ПII.2);
- 2-позиционная тетрагруппировка гексаграмм и 16 членов каждой из ее 4 группировок определяются по корреляционным позициям 2-5 гексаграмм, на которых расположены указанные атрибутивные пары чисел 0-0, 0-1, 1-0 и 1-1 (Рис. ПII.3);
- 3-позиционная тетрагруппировка гексаграмм и 16 членов каждой из ее 4 группировок определяются по корреляционным позициям 3-6 гексаграмм, на которых расположены указанные атрибутивные пары чисел 0-0, 0-1, 1-0 и 1-1 (Рис. ПII.4).

На каждом из Рис. ПII.2-ПII.4, члены каждой из 4 группировок названных *m*-позиционных тетрагруппировок выделены соответственно 4 цветами. Расположение этих 4 цветов в соответствующих корреляционных позициях (1-4, 2-5 и 3-6) в 64 гексаграммах полностью совпадает с тем расположением 4 символов ♠, ♥, ♦ и ♣ в показанных генетических матрицах 64 триплетов, которое связано с описанными геномными гештальт-феноменами вероятностей и которое представляет расположение 64 триплетов в случаях 1-позиционной тетрагруппировки (Рис. ПII.2), 2-позиционной тетрагруппировки (Рис. ПII.3) и 3-позиционной тетрагруппировки (Рис. ПII.4). Точнее, во всех этих взаимно соответствующих *m*-позиционных тетрагруппировках из 64 триплетов и 64 гексаграмм символ ♠ соответствует гексаграммной паре цифр 1 и 1; символ ♥ соответствует паре гексаграмм 1 и 0; символ ♦ соответствует паре 0 и 1; символ ♣ соответствует паре 0 и 0.



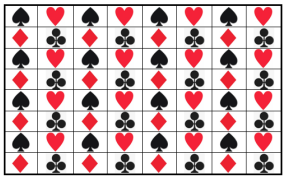
111111	111110	111101	111100	111011	111010	111001	111000
110111	110110	110101	110100	110011	110010	110001	110000
101111	101110	101101	101100	101011	101010	101001	101000
100111	100110	100101	100100	100011	100010	100001	100000
011111	011110	011101	011100	011011	011010	011001	011000
010111	010110	010101	010100	010011	010010	010001	010000
001111	001110	001101	001100	001011	001010	001001	001000
000111	000110	000101	000100	000011	000010	000001	000000

Рис. ПП.2. Идентичность расположения 1-позиционной тетрагруппировки 64 генетических триплетов в генетической матрице (из Рис. 6.1) и расположения 1-позиционной тетрагруппировки 64 гексаграмм в древнекитайской таблице И-Цзин; эта тетрагруппировка определяется четырьмя бинарными парами (1 и 1, 1 и 0, 0 и 1, 0 и 0), стоящими внутри гексаграмм в их традиционных корреляционных положениях 1-4. Члены каждой из 4 группировок отмечены своим цветом.



111111	111110	111101	111100	111011	111010	111001	111000
110111	110110	110101	110100	110011	110010	110001	110000
101111	101110	101101	101100	101011	101010	101001	101000
100111	100110	100101	100100	100011	100010	100001	100000
011111	011110	011101	011100	011011	011010	011001	011000
010111	010110	010101	010100	010011	010010	010001	010000
001111	001110	001101	001100	001011	001010	001001	001000
000111	000110	000101	000100	000011	000010	000001	000000

Рис. ПП.3. Идентичность расположения 2-позиционной тетрагруппировки 64 генетических триплетов в генетической матрице (из Рис. 6.2) и расположения 2-позиционной тетрагруппировки 64 гексаграмм в древнекитайской таблице И-Цзин; эта тетрагруппировка определяется четырьмя бинарными парами (1 и 1, 1 и 0, 0 и 1, 0 и 0), стоящими внутри гексаграмм в их традиционных корреляционных положениях 2-5. Члены каждой из 4 группировок отмечены своим цветом по аналогии с предыдущим рисунком.



111111	111110	111101	111100	111011	111010	111001	111000
110111	110110	110101	110100	110011	110010	110001	110000
101111	101110	101101	101100	101011	101010	101001	101000
100111	100110	100101	100100	100011	100010	100001	100000
011111	011110	011101	011100	011011	011010	011001	011000
010111	010110	010101	010100	010011	010010	010001	010000
001111	001110	001101	001100	001011	001010	001001	001000
000111	000110	000101	000100	000011	000010	000001	000000

Рис. ПП.4. Идентичность расположения 3-позиционной тетрагруппировки 64 генетических триплетов в генетической матрице (из Рис. 6.3) и расположения 3-позиционной тетрагруппировки 64 гексаграмм в древнекитайской таблице И-Цзин; эта тетрагруппировка определяется четырьмя бинарными парами (1 и 1, 1 и 0, 0 и 1, 0 и 0), стоящими внутри гексаграмм в их традиционных корреляционных положе-

ниях 3-6. Члены каждой из 4 группировок отмечены своим цветом.

Параллели, показанные на Рис. ПП.2-ПП4, дают свидетельства того, что обнаруженные геномные гештальт феномены вероятностей структурно связаны со схемами И-Цзин, включая таблицу 64 гексаграмм. Здесь можно снова вспомнить, что создатель аналитической психологии К. Юнг и его соратник, лауреат Нобелевской премии по физике В. Паули, которые были экспертами по И-Цзин, считали, что триграммы и гексаграммы И-Цзин «*фиксируют универсальный набор архетипов (врожденных психических структур)*» [Щуцкий, 1997, с. 12]. Соответственно, все эти данные подтверждают, в частности, что обнаруженные генетические гештальт-феномены и описанные генетические гештальт-правила имеют глубокое отношение к свойствам генетически унаследованных психологических феноменов и к богатой теме архетипов природы.

Дополнительно можно упомянуть, что матрицы из 4 биграмм, 16 тетраграмм и 64 гексаграмм И-Цзин могут быть представлены как взаимосвязанные члены тензорного семейства матриц по аналогии с тензорным семейством генетических матриц *n*-плетов на Рис. 2.2 (подробнее см. в книге [Петухов, 2008]). Такие тензорные семейства различных (2^n*2^n)-матриц позволяют искать новые «*врожденные психические структуры*» и архетипы природы.

И-Цзин еще называют «Книгой циклических перемен». Автором отмечается существование связи между различными ансамблями двоичных чисел в таблице 64 гексаграмм, основанной на логической операции сложения по модулю 2. Например, ансамбли двоичных чисел из разных частей тетра-группировки, показанной на рис. ПП.2, преобразуются друг в друга на основе этой логической операции. В частности, те два ансамбля двоичных чисел таблицы 64 гексаграмм, расположение которых соответствует расположению символов ♠ и ♦ в мозаичной матрице, представляющей 1-позиционную тетра-группировку 64 генетических триплетов на рис. ПП.2, преобразуются друг в друга логической операцией сложения по модулю 2 двоичного числа 100000 с членами ансамблей. Отдельной интересной темой является изучение циклических бинарных отношений между различными тетрагруппировками в матрицах, представляющих генетические гештальт-феномены.

Интересно отметить, что числа 2 и 3, которые в ДНК и РНК выступают числами водородных связей комплементарных нуклеотидов, являлись основой древнекитайской арифметики.

Остановимся еще на давно отмечаемых структурных аналогиях ДНК алфавитов с 4-секционными мандалами, сопряженными со схемами И-Цзин (Рис. ПП.5). На протяжении тысячелетий миллионы буддистов, индуистов и других верующих создавали мандалы как инструмент медитации для достижения «просветления» и исцеления. Создатель аналитической психологии К. Юнг и его соратник В. Паули считали мандалу врожденным архетипом бессознательного и сопряжением космограммы с психограммой, способным гармонизировать психику приведением к тождеству всех жизненных оппозиций. Юнг говорил о психической целостности как «*Самости, которая проявляется во множественности, так называемой четверичности… Мандала – изображение психического процесса, воссоздание нового центра личности. Символически изображается при помощи круга, симметричного расположения некоего четырехкратного количества. В ламаизме и тантра-йоге является*

инструментом созерцания, местом рождения и пребывания божества. ... Четверичность (также кватерность). Универсальный архетип, являющий собой логическую предпосылку всякого целостного суждения» [Вер, 2010, с. 88, 200-202].

Тема мандал издревле связана с темой музыкальной гармонии устройства мира. Древние учения «просветленных» мудрецов Индии, Китая и других стран утверждали взгляд на музыкальную гармонию как на нечто изначальное. В силу этого тысячелетия практик медитации с использованием мандал сопряжены для буддистов с идеей музыкальной гармонии, отраженной в рядах масштабируемых колоколов в древних буддийских монастырях (Рис. ПП.5).

Рис. ПП.5. Верхний ряд: иллюстрация аналогий между алфавитами ДНК и древними 4-секционными мандалами (изображения взяты с сайтов: www.kozlenkoa.narod.ru, https://blog.humdes.com/teoreticheskie-osnovy-chast-3-geksagrammy-i-dzin/, https://www.dreamstime.com/kamala-yantra-hinduism-symbol-vector-icon-kamala-yantra-hinduism-symbol-image154056053).
Нижний ряд: примеры ансамблей колоколов в древних буддистских монастырях (фото с сайтов https://www.fotex.biz/countries/thailand/chiangmai/6653001550/ , https://ru.depositphotos.com/stock-photos/brass-bells.html?offset=200&sorting=newest&qview=307409376).

Обнаружение описанной выше связи многослойных ДНК-текстов с гештальт правилами тетрагруппировок и гармонической прогрессией 1/*n*, сопряженной с музыкальной гармонией, позволяет автору строить «геномные» 4-секционные мандалы алгебро-гармонического числового характера (а не просто картинки с символами). В этих 4-секционных диаграммах все сектора и все их кольца взаимно связаны гармонической прогрессией, сопряженной с гармониками в музыке. На Рис. ПП.6 показан поясняющий пример к теме таких 4-секционных диаграмм. На этом примере каждый из четырех секторов содержит данные о суммарном количественном составе соответствующей 1-позиционной группировки *n*-плетов в слоях геномного ДНК-текста, то есть группировки, *n*-плеты которой начинаются с одного из нуклеотидов A, T, C или G (см. выше разделы 3 и 8, откуда заимствуются обозначения на этом рисунке).

При нормировке этих количеств - в каждом из секторов - делением на количество соответствующего нуклеотида в геноме, последовательность колец каждого из 4 секторов предстает как последовательность членов гармонической прогрессии или как последовательность музыкальных гармоник.

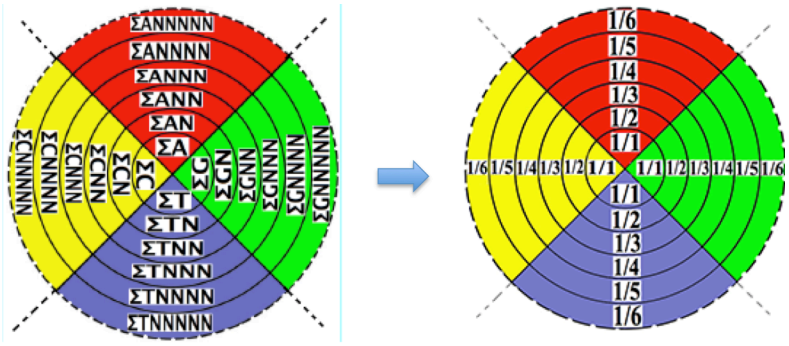


Рис. ПП.6. Пример “геномной” мандалы первого порядка.

Для длинных ДНК-текстов кроме описанной численной мандалы первого порядка, основанной на *n*-плетах, начинающихся с одного из 4 моноплетов A, T, C и G, существуют численные мандалы второго, третьего и более высоких порядков, основанные на тетрагруппировках *n*-плетов, начинающихся с одного из 16 дуплетов, или одного из 64 триплетов, и т.д. Это связано с описанным выше в конце раздела 3 эстафетным принципом гештальт связи соседних слоев геномных ДНК-текстов, представленном равенствами (3.5). Он касается связи каждого *n*-плета с четверкой (*n*+1)-плетов очередного (*n*+1)-го слоя, которая напоминает общебиологический процесс учетверения при мейозе, в ходе которого из одной половой клетки образуется четыре половые клетки.

Тем самым, имеется огромная фрактало-подобная иерархия численных ДНК-мандал, завязанная на геномные гештальт-правила и музыкальные гармоники. Этот «хор» ДНК-мандал можно озвучивать или представлять в цвете, вибрациях, электро-стимулах и пр. в технологических целях, включая развитие древних технологий мандал-музыкальных медитаций для просветления и оздоровления. Данные вопросы разрабатываются в нашей лаборатории и «Центре междисциплинарных исследований музыкального творчества Московской государственной консерватории им. П.И. Чайковского.

Благодарности

Представленные в статье результаты стали возможны во-многом благодаря многолетнему сотрудничеству Российской и Венгерской Академий наук по теме “Нелинейные модели и симметрологический анализ в биомеханике, биоинформатике и теории самоорганизующихся систем”, по которой автор был научным руководителем от Российской Академии наук. Автор выражает глубокую благодарность Ю.Дарвашу, А.А. Коблякову, Ю.И. Манину, В.И. Свирину, И.В. Степаняну, Г.К. Толоконникову, Е.Фиммель, М. Хи, З.Б. Ху за их сотрудничество.

Литература

Бернштейн Н.А. *Очерки по физиологии движений и физиологии активности*. Москва: Медицина, (1966).
Вер Г. *Карл Густав Юнг*. Урал Лтд (1998). ISBN: 5-88294-073-7
Владимиров Ю.С. *Метафизика*. (Второе издание) М.: БИНОМ. (2009).
Владимиров Ю.С. *Реляционная картина мира. Книга 1: Реляционная концепция геометрии и классической физики*. М.: ЛЕНАНД (2020).
Владимиров Ю.С. *Реляционная картина мира. Книга 2: От бинарной пред-геометрии микромира к геометрии классического пространства-времени*. М.: ЛЕНАНД (2021).
Гурвич А.Г. *Избранные труды*. Москва, Медицина (1977).
Капра Ф., Луизи П.Л. *Системный взгляд на жизнь: целостное представление*. Перевод с английского А.А. Сидоровой-Бирюковой, В.В. Шуликовской; под редакцией В. В. Шуликовской. М., URSS, 502 с. (2020).
Конопельченко Б.Г., Румер Ю.Б. Классификация кодонов в генетическом коде. *ДАН СССР*, т. 223, №2, с. 471-474 (1975).
http://odasib.ru/OpenArchive/DocumentImage.cshtml?id=Xu1_pavl_635766969644249164_22519&eid=R1_0001_0374
Кулаков Ю.И. *Теория физических структур*. М., 2004, 847 с.
Манин Ю.И. Вычислимое и невычислимое», М., Сов.радио, 1980, 128 с.
Налимов В.В. *Разбрасываю мысли*. М., Центр гуманитарных инициатив (2015).
Орен Гоберман. О чём думают грибы? – *Калкалист,* Израиль, 11.08.2011, https://efimbog.livejournal.com/601646.html.
Пенроуз Р. *Тени разума*. М., Институт компьютерных исследований (2005).
Петухов С.В. Биомеханика, бионика и симметрия. М., Наука (1981).
Петухов С.В. *Матричная генетика, алгебры генетического кода, помехоустойчивость*. М, Регулярная и хаотическая динамика, 316 с. (2008).
Петухов С.В. Генетические основы алгебраической биологии, гештальт-генетика и тетра-эйдосы Ю.И. Кулакова. *Презентация доклада на семинаре «Основания фундаментальной физики» физического факультета МГУ, руководимом Ю.С.Владимировым*, 04.03.2021. https://drive.google.com/file/d/18G0vmhPw82s4BmdLHvNNiLNGhroYkPQ0/view?usp=sharing.
Петухов С.В., Свирин В.И., Хазина Л.В. Бионика спиральных структур. *Проблемы машиностроения и надёжности машин*, №3, 2015, с.70-76, http://petoukhov.com/bionica.pdf
Петухов С.В., Толоконников Г.К. Алгебраическая биология и матричная генетика. – *Видео-презентация доклада на международном междисциплинарном семинаре "Алгебраическая биология и теория систем"* 13.02.2020, https://www.youtube.com/watch?v=H2dNtvTMl1M&t=330s.
Прасолов В. В. *Задачи и теоремы линейной алгебры*. 2-е изд. М., 2008.
Тейяр де Шарден. *Феномен человека*. М., Наука (1987).
Франк-Каменецкий М.Д. *Самая главная молекула*. М, Наука (1988).
Шноль С.Э. *Физико-химические факторы биологической эволюции*. М., Мир (1989).
Шредингер Э. Основные принципы теории метрики цветов в дневном свете. – В кн: *Новые пути в физике*. М., Наука, с. 135-192 (1971).
Шредингер Э. *Что такое жизнь?* М., Атомиздат (1972).
Шульц Е. Организм как творчество. - В кн.: *Вопросы теории и психологии творчества*, т.7, Харьков, с.108-190 (1916).
Щуцкий Ю.К. *Китайская классическая «Книга перемен»*. М., Восточная

литература (1997).
Abbott D., Davies P.C.W., Pati A.K. (eds.), Foreword by Sir Roger Penrose. *Aspects of Life* (2008). ISBN-13: 978-1-84816-253-2
Cook T.A. *The Curves of Life*. London: Constable and Co, 1914.
Fimmel E., Petoukhov S.V. Genetic Code Modelling from the Perspective of Quantum Informatics. In: Hu Z., Petoukhov S., He M. (eds) Advances in Artificial Systems for Medicine and Education II. AIMEE2018 2018. Advances in Intelligent Systems and Computing, vol 902, pp. 117-125. Springer, Cham (2020a).
Fimmel E., Petoukhov S.V. Development of Models of Quantum Biology Based on the Tensor Product of Matrices. In: Hu Z., Petoukhov S., He M. (eds). Advances in Intelligent Systems and Computing, v. 1126, p.126-135. Springer (2020b), DOI https://doi.org/10.1007/978-3-030-39162-1_12.
Fröhlich F. Genetic code as Language. – In: Fröhlich H. Editor. *Biological Coherence and Response to External Stimuli*, Springer, p. 192-204 (1988). ISBN 978-0-387-18739-6.
Fröhlich H. Coherent electric vibrations in biological systems and cancer problem. // IEEE Trans. Microwave Theory Tech., v. 26, p. 613–617 (1978).
Fröhlich H. The biological effects of microwaves and related questions. In: *Advances in Electronics and Electron Physics*, v. 53, p. 85–152. New York, USA: Academic Press (1980).
Fröhlich H. Introduction. Theoretical Physics and Biology. – In: Herbert Fröhlich editor, *Biological Coherence and Response to External Stimuli*. Springer, p. 3-24 (1988), ISBN 978-0-387-18739-6.
Fröhlich H., Kremer F. *Coherent Excitations in Biological Systems*. Springer-Verlag, (1983). ISBN 978-3-642-69186-7.
Geesink H.J.H., Meijer D.K.F. Quantum wave information of life revealed. *NeuroQuantology*, vol. 14, issue 1 (March 2016), doi:10.14704/nq.2016.14.1.911.
Gell-Mann M., Ne'eman Y. *The Eightfold Way*. N.Y., W.A.Benjamin (1964).
Hofstadter D.R. *Gödel, Escher, Bach. An Eternal Golden Braid*. N.Y.: Vintage Books, 777 p. (1980).
Holland B.W. Dynamic specificity and the pairing of homologous chromosomes in meiosis. *J. Theor. Biol*., 35(2), p. 395-397 (1972).
Hyland G.J. Coherent GHz and THz excitations in active biosystems, and their implications. In: The Future of Medical Diagnostics? Proceeding of Matra Marconi UK, Directorate of Science, Internal Report, Portsmouth, UK, pp. 14–27 (25 June 1998).
Hu Z.B., Petoukhov S.V., Petukhova E.S. I-Ching, dyadic groups of binary numbers and the geno-logic coding in living bodies. *Progress in Biophysics and Molecular Biology*, vol. 131, pp. 354-368 (December 2017).
Igamberdiev A.U. Quantum mechanical properties of biosystems: a framework for complexity, structural stability, and transformations. *Biosystems*, 31(1), p. 65–73 (1993).
Jordan P. Die Quantenmechanik und die Grundprobleme der Biologie und Psychologie. *Naturwissenschaften* 20, 815–821 (1932), (doi:10.1007/BF01494844).
Kienle G. Experiments concerning the non-euclidian structure of the visual space. – In: *Bioastronautics*. N.Y., London, p. 386-400 (1964).
Kienle G. *Die optishen Wahrnehmungsstorungen und die nichteuklidische Struktur*

des Sehraumes. Stuttgart: Thieme, 172 S. (1968).
Labunets V.G. Clifford algebras as unified language for image processing and pattern recognition. Computational Noncommutative Algebra and Applications , eds. J. Byrnes &G. Osteimer, Kluwer: Dordrecht and Boston and London, vol. 136 of NATO Science Series , pp. 197–225 (2003).
Labunets V.G., Rundblad E., Astola J. Is the brain a clifford algebra quantum computer? (Chapter 25). Applications Of Geometrical Algebras In Computer Science And Engineering , eds. L. Dorst, C. Doran & A. Lasenby, Birkhauser: Boston, pp. 285–295 (2002).
Luneburg R.K. *Mathematical analysis of binocular vision*. Princeton: Univ. Press, 194 p. (1947).
Luneburg R.K. Metric methods in bimocular visual perception. – In: Studies and essays. N.Y.: Intersci. Publ., Courant anniv. vol., p. 215-239 (1948).
Luneburg R.K. The metric of binocular visual space. *J. Opt. Soc. Amer*., v. 40, #10, p. 627-642 (1950).
Matsuno K. Cell motility as entangled quantum coherence. *BioSystems*, 51, p. 15–19 (1999).
Matsuno K., Paton R.C. Is there a biology of quantum information? *BioSystems,* 55, p. 39–46 (2000).
McFadden J., Al-Khalili J. The origins of quantum biology. *Proceedings of the Royal Society A*, Vol. 474, Issue 2220, p. 1-13, 12 December 2018, https://doi.org/10.1098/rspa.2018.0674. https://royalsocietypublishing.org/doi/full/10.1098/rspa.
Nakagaki T., Yamada H., Toth A. Intelligence: Maze-solving by an amoeboid organism. *Nature* **407**, 470 (28 September 2000), doi:10.1038/35035159.
Patel A. Quantum algorithms and the genetic code. *Pramana – J. Phys*. 56(2–3), p. 367–381, arXiv:quant-ph/0002037 (2001a).
Patel A. Testing quantum dynamics in genetic information processing. *J. Genet*. 80(1), p. 39–43 (2001b).
Patel A. Why genetic information processing could have a quantum basis. *J. Biosci.* 26(2), p. 145–151 (2001c).
Pauli W. *Writings on Physics and Philosophy*. Ed. by C.P. Enz and K. von Meyenn, Springer, Berlin (1994).
Petoukhov S.V. Genetic code and the Ancient Chinese «Book of Changes». *Symmetry:Culture and Science*, vol. 10, №3-4, pp. 211-226 (1999).
Petoukhov S.V. The system-resonance approach in modeling genetic structures. *Biosystems*, v. 139, p. 1-11 (January 2016), http://dx.doi.org/10.1016/j.biosystems.2015.11.001, http://petoukhov.com/PETOUKHOV_ARTICLE_IN_BIOSYSTEMS.pdf.
Petoukhov S.V. The Genetic Coding System and Unitary Matrices. *Preprints* 2018, 2018040131 (2018), doi: 10.20944/preprints201804.0131.v2.
Petoukhov S.V. Nucleotide Epi-Chains and New Nucleotide Probability Rules in Long DNA Sequences. Preprints 2019, 2019040011, 17 pages (2019), doi: 10.20944/preprints201904.0011.v1, https://www.preprints.org/manuscript/201904.0011/v1
Petoukhov S.V. Hyperbolic Rules of the Cooperative Organization of Eukaryotic and Prokaryotic Genomes. *Biosystems*, 198, 104273 (2020a).
Petoukhov S.V. Hyperbolic Rules of the Oligomer Cooperative Organization of Eukaryotic and Prokaryotic Genomes. *Preprints* 2020, 2020050471 (2020b),

doi:10.20944/preprints202005.0471.v2, https://www.preprints.org/manuscript/202005.0471/v2.

Petoukhov S.V. The rules of long DNA-sequences and tetra-groups of oligonucleotides. arXiv:1709.04943v6 , 6th version from 22.05.2020 (2020c).

Petoukhov S.V. Algebraic Rules for the Percentage Composition of Oligomers in Genomes. Preprints 2021, 2021010360 (2021a), (doi:10.20944/preprints202101.0360.v2). https://www.preprints.org/manuscript/202101.0360/v2

Petoukhov S.V. Modeling inherited physiological structures based on hyperbolic numbers, *BioSystems*, Vol. 199, 104285 (2021b), ISSN 0303-2647, https://doi.org/10.1016/j.biosystems.2020.104285.

Petoukhov S.V. Algebraic Rules for the Percentage Composition of Oligomers in Genomes. Preprints 2021, 2021010360 (2021a), (doi:10.20944/preprints202101.0360.v2). https://www.preprints.org/manuscript/202101.0360/v2

Petoukhov S.V., He M. *Symmetrical Analysis Techniques for Genetic Systems and Bioinformatics: Advanced Patterns and Applications*. Hershey, USA, IGI Global (2010).

Petoukhov S.V., Petukhova E.S. Symmetries in genetic systems and the concept of geno-logical coding. *Information*, *8*(1), 2 (2017), doi:10.3390/info8010002, http://www.mdpi.com/2078-2489/8/1/2/pdf

Petoukhov S.V., Petukhova E.S., Svirin V.I. Symmetries of DNA alphabets and quantum informational formalisms. *Symmetry: Culture and Science,* Vol. 30, No. 2, p.161-179 (2019), https://doi.org/10.26830/symmetry_2019_2_161.

Resnikoff H.L. Differential geometry and color perception. *J. Math. Biol*., v. 1, #2, p. 97-131 (1974).

Schelling H. von. Advanced color geometry. J. Opt. Soc. Amer., v. 45, #12, p. 1072-1079 (1955).

Schelling H. von. Concept of distance in affine geometry and its application in theories of vision. J. Opt. Soc. Amer., v. 46, p. 309-315 (1956a).

Schelling H. von. Reply to P.J van Heerden. J.Opt. Soc. Amer., v. 46, #11, p. 1000-1001 (1956b).

Schelling H. von. Die Geometrie des beidenaugigen Sehens. – Optik, Bd. 17, H. 7, S. 345-364 (1960).

Schelling H. von. Experienced space and time. – In: *Bioastronautics*. N.Y., L., p. 361-385 (1964).

Vasconcellos A.R., Vannucchi F.S., Mascarenhas S., Luzzi R. Frohlich Condensate: Emergence of Synergetic Dissipative Structures in Information Processing Biological and Condensed Matter Systems. *Information*, v. 3(4), p. 601-620 (2012), doi:10.3390/info3040601. https://www.mdpi.com/2078-2489/3/4/601/htm.

Zabeo D., Heumann J.M., Schwartz C.L., Suzuki-Shinjo A., Morgan G., Widlund P.O., Höög J.L. A lumenal interrupted helix in human sperm tail microtubules. *Scientific Reports*,; 8(1), 2727. (2018 Feb 9). doi: 10.1038/s41598-018-21165-8.

Algebraic Biology, Gestalt Genetics, Hereditary and Artificial Intelligence

S.V. Petoukhov
Laboratory of Biomechanical Systems, Mechanical Engineering Research Institute, Russian Academy of Sciences

http://petoukhov.com/

Abstracts. The results of the analysis of the structural features of the molecular genetic coding system are presented. The regularities of DNA sequences of nucleotides in the genomes of higher and lower organisms have been revealed. Their detection is based on the presentation of genomic DNA sequences in the form of multilayer algebraic structures and a comparative analysis of the percentage of *n*-plets in different layers. The rules for the constancy of the sums of percentages in special tetragroupings of *n*-plets in different text layers of genomic DNA are described. These constancy of the sums are realized with very different quantities and values of the summands. These rules for the constancy of the percentages are accompanied by hyperbolic rules for the interrelation of quantities in *n*-plets groupings in different layers of genomic DNA texts. The discovered rules are candidates for the role of universal genomic rules. The study uses a tensor family of matrix representations of DNA alphabets to identify the algebraic features of the mentioned rules. The analogies of the discovered genetic phenomena with the phenomena of Gestalt psychology are noted, which allows developing the ideas of Gestalt genetics. The presented results indicate: 1) living organisms can be interpreted as quantum-informational algebraic-harmonic entities; 2) bioinformatics is based on probabilistic methods of multilayer coding and information processing.

Keywords: DNA alphabets, binary oppositions, genomes, multilayered texts, probabilities, tetragroupings, gestalt, matrices, tensor product, hyperbolic numbers, projectors, matrix genetics, algebraic and quantum biology, intelligence