корреляционно-спектрального анализа, позволит исключить вероятность немотивированного брака и сократить время диагностики в 20 раз.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Шахматов Е. В., Прохоров С. А., Крючков А. Н. и др.* Определение звукоизолирующей способности элементов конструкций / Проблемы и перспективы развития двигателестроения. Сб. тр. междунар. науч.-техн. конф.: Самара: СГАУ, 2003. С. 318–323.
2. *Макарьянц Г. М., Гаспаров М. С., Графкин А. В., Шахматов Е. В.* Использование методов корреляционно-спектрального анализа в виброакустической диагностике автомобильных амортизаторов// Проблемы и перспективы развития двигателестроения. Мат. докл. междунар. науч.-техн. конф. Самара: СГАУ, 2006. Ч. 1. С. 237-238.
3. *Тюрин Ю.Н., Макаров А.А.* Анализ данных на компьютере / Под ред. В.Э. Фигурнова. 3-е изд., перераб. и доп. М.: ИНФРА-М, 2003. 544 с.
4. *Прохоров С. А., Иващенко А. В., Графкин А. В.* Автоматизированная система корреляционно-спектрального анализа случайных процессов / *Под ред. С.А. Прохорова* Самара: СНЦ РАН, 2003. 286 с.
5. *Прохоров С. А., Графкин А. В.* Программный комплекс корреляционно-спектрального анализа в ортогональных базисах. Самара: СНЦ РАН, 2005. 198 с.
6. *Графкин А. В.* Моделирование автоматизированных систем научных исследований с использованием семантических сетей// Повышение эффективности средств обработки информации на базе математического моделирования. VIII Всерос. науч.-техн. конф. Тамбов: ТВАИИ, 2006. С. 65–75.
7. *Венгеров И. А., Карев А. В.* Инструментальный контроль автомобилей. М.: За рулем, 1999. 31 с.
8. *Прохоров С. А.* Аппроксимативный анализ случайных процессов/ 2-е изд., перераб. и доп. Самара: СНЦ РАН, 2001. 380 с.
9. *Графкин А. В.* Аппроксимативный анализ взаимных корреляционно-спектральных характеристик временных рядов с помощью ортогональных функций Лежандра, Дирихле/ Дисс. … канд. тех. наук. Самара, 2006. 147 с.

Поступила 31.07.2006 г.

УДК 57.01+577.4

Л. С. Бекасов, А. А. Тверетин

МЕТОДЫ ПРЕДСТАВЛЕНИЯ ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ

Рассмотрен принцип формального представления элементарных звеньев ДНК. Предложены критерии физической природы, на основе которых получена графическая модель фрагмента 16S-рРНК. Формализация нуклеотидных последовательностей обуславливает возможность анализа характера упаковки её элементов. Показаны несколько вариантов такой модели, в том числе, иллюстрирующие её чувствительность на имитацию мутационных процессов.

Четыре органические молекулы — аденин (A), гуанин (G), цитозин (C) и тимин/урацил (T/U) являются алфавитом, с помощью которого осуществляется кодирование всей генетической информации. Эти четыре молекулы могут образовывать между собой комплементарные (согласованные) связи. Аденин и гуанин называются пуринами, тогда как тимин и цитозин это пиримидины. Аденин связывается с тимином (урацилом), а гуанин с цитозином. Поскольку все белки строятся из аминокислот, а каждая аминокислота образуется из трёх нуклеотидов (триплетов, кодонов), то моделирование отдельных компонент генетики, в частности белков и нуклеиновых кислот, можно осуществлять, используя аминокислотный алфавит.

Не лишне напомнить, что 21 аминокислота кодируется 64-мя триплетами — структурами, представляющими собой три последовательно расположенных нуклеотида. При этом каждой аминокислоте может соответствовать не один кодон. Всего известно семнадцать вариантов, регламентирующих соответствие между аминокислотами и составом кодонов. Таким образом, переход от нуклеотидного алфавита к аминокислотному обуславливает, в первую очередь, формализацию как самих AGTC-элементов, так и простейших структур.

Формальное представление AGTC-элементов и связей между двумя элементами комплементарного и последовательного характера ранжируется с учётом некоторых критериев. В связи с тем, что модель биологической структуры должна наиболее достоверно отражать её информационную сущность, желательно, чтобы критерий был бы не один, а физическая природа критериев не должна быть одинаковой.

Ниже в качестве таковых используются особенности физического характера. Довольно сложно учесть все возможные параметры и получаемые на их основе формальные описания, характеризующие потенциальную сторону обязательно будут содержать определенную погрешность. Так или иначе, модель будет представлять нуклеотидную последовательность числовым рядом комплексной значимости с пространственными и амплитудными параметрами. Это представление позволит осуществлять анализ конкретной биоструктуры с точки зрения принципов её построения, в частности, характера очерёдности расположения нуклеотидов.

В качестве количественных критериев были взяты молекулярная масса и силы водородных связей между комплементарными нуклеотидами с учётом их пространственной ориентации, как наиболее очевидные и простые.

В [1] принцип организации аминокислоты интерпретируются с точки зрения алгебры-логики, т.е. все аминокислоты представляются через конъюнктивные нормальные формы, элементами которых являются нуклеотиды. Такая формализация с использованием аппарата комбинаторики позволяют получить трехмерные модели гипотетических белков (кодонограммы), Тем не менее, столь оригинальный подход в принципе, на наш взгляд, не в состоянии оценить структурные особенности системного характера, т.е. выявить закономерности взаимосвязей компонентов между собой, в том числе особенности водородных связей, которые имеют не химическую, а полевую природу. Применяющийся в этих методах алфавит качественно не изменяется. Формализуются лишь обозначения кодонов и возможности их взаимного расположения, тогда как качество характера их взаимосвязей с учётом особенностей организации непосредственно гена (например, взаимные расположения экзонов, интронов, регуляторов) игнорируется, более того, логика взаимоотношений между нуклеотидами здесь, как минимум, будет четырехзначная.

С. В. Петуховым представлен метод формализации, касающийся непосредственно комплементарных связей между четырьмя основными AGTC-элементами с применением семейства биериодических матриц, простейшая из которых представлена на рис. 1.

Каждому нуклеотиду в матрице присваивается (принцип присваивания не раскрывается) определенное значение двоичного кода. Оригинальность подхода состоит в попытке учесть в модели такой физический параметр, как количество водородных связей между пуриновыми (G, A) и пиримидиновыми (C, T) основаниями, которое для пары GC равно 3 и AT соответственно 2.

Однако многообразие и сложность процессов порождения белковых структур не позволяют известными методами однозначно выявить динамические особенности этих процессов. С целью избежания недостатков предыдущей модели, предлагается кодировать не один нуклеотид, а его сочетание с последующим в цепи, что позволит учесть разницу между нуклеотидами и уйти от линейности системы (рис. 2) [2]. Нуклеотиды в сочетании упорядочены по возрастанию своей молекулярной массы.

	1	0
1	**C**(11)x3	**A**(10)x2
0	**T**(01)x2	**G**(00)x3

Р и с. 1. Простейшая матрица кодирования нуклеотидов

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

Р и с. 2. Простейшая матрица сочетаний нуклеотидов

Таким образом, каждому сочетанию присваивается весовой коэффициент от 0 до 15 в соответствии с его молекулярной массой, причем сочетание **AG** не эквивалентно сочетанию **GA**, т.к. даже такое малое изменение последовательности может существенно скорректировать функцию нуклеотидной последовательности [3]. Например: сочетание **AG** и **CT** имеют весовые коэффициенты 1011_2 и 0001_2 или 11 и 1 соответственно.

Представим нуклеотидную последовательность в виде числового ряда $Y_i = F(X_i)$, где X_i — нуклеотид в цепи, i=0, 1, … , n:

$$F(X_i)=\begin{cases}3, \text{если } X_i = G;\\ 2, \text{если } X_i = A;\\ 1, \text{ если } X_i = T;\\ 0, \text{ если } X_i = C.\end{cases} \quad (1)$$

Запишем функцию Y_i для модифицированной матрицы (рис. 2.):

$$Y_i = F(X_{i+1}) + 4 \cdot F(X_i). \quad (2)$$

Вышеприведенный метод имеет существенный недостаток: присвоение сочетаниям нуклеотидов дискретных значений весовых коэффициентов реализует лишь возможность их ранжирования, тогда как более точное описание характера взаимосвязи отсутствует. Цитозин, гуанин, тимин, аденин имеют молекулярные массы 110,11; 150,14; 125,12; 134,14 соответственно. Введем массив M, элементами которого являются относительные молекулярные массы азотистых оснований нуклеотидов:

$$M = \left|0,733 \;\; 0,833 \;\; 0,893 \;\; 1\right|. \quad (3)$$

С помощью (3) можно дать более точную информацию. Весовые коэффициенты для сочетаний можно получить, используя либо аддитивную (в виде разности), либо мультипликативную (как отношение или произведение) процедуры для соответствующих нуклеотидов.

Функции Y_i в этом случае запишутся следующим образом:

$$Y_i = M_{F(X_i)} - M_{F(X_{i+1})}, \quad (4)$$

$$Y_i = M_{F(X_i)} \cdot M_{F(X_{i+1})}, \quad (5)$$

$$Y_i = \frac{M_{F(X_i)}}{M_{F(X_{i+1})}}. \quad (6)$$

В [4] получены средние энергии взаимодействия для A–T и G–C связей с учётом их пространственной ориентации, которые равны 13 ккал/моль и 21 ккал/моль соответственно. Энергии определены с точностью до 1–2 ккал/моль. Используя относительную силу водородных связей между комплементарными парами нуклеотидов, она рассчитывается как отношение сил связей между A–T и G–C. В результате получим массив коэффициентов K:

$$K = \left|-1,615 \;\; -1 \;\; 1 \;\; 1,615\right|, \quad (7)$$

элементами которого являются относительные силы связей между комплементарными нуклеотидами, кроме того, введен знак, который отражает пространственную ориентацию пары комплементарных нуклеотидов относительно третьего или пятого углеродного атомов дезоксирибозы.

В качестве элементов массива K можно также взять удельную силу связи между комплементарными нуклеотидами, которая рассчитывается как отношение средних энергий взаимодействия и количества водородных связей соответствующих комплементарных пар:

$$K = \left|-7 \;\; -6,5 \;\; 6,5 \;\; 7\right|. \quad (8)$$

Таким образом, соотношения (2), (4), (5), (6) соответственно запишутся в следующем виде:

$$Y_i = (F(X_{i+1}) + 4 \cdot F(X_i)) \cdot K_i, \quad (9)$$

$$Y_i = (M_{F(X_i)} - M_{F(X_{i+1})}) \cdot K_i, \quad (10)$$

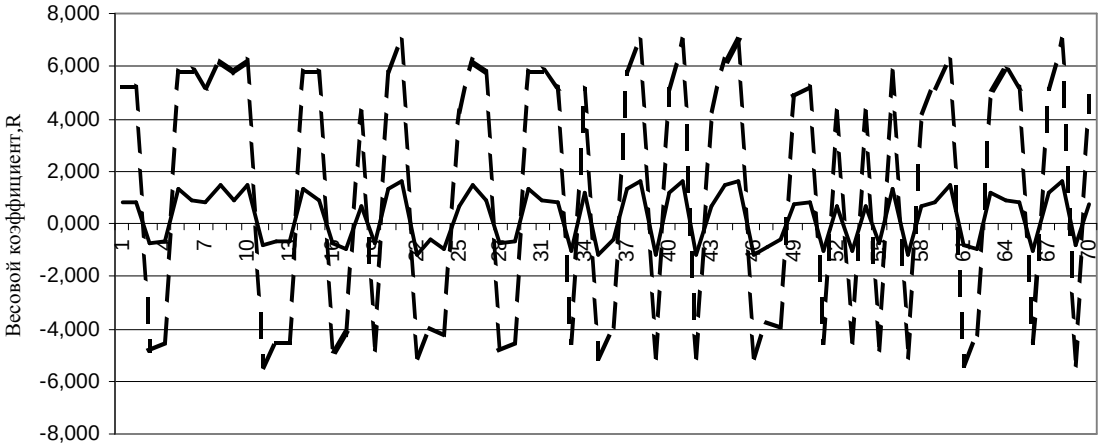
$$Y_i = (M_{F(X_i)} \cdot M_{F(X_{i+1})}) \cdot K_i, \quad (11)$$

$$Y_i = \frac{M_{F(X_i)}}{M_{F(X_{i+1})}} \cdot K_i. \quad (12)$$

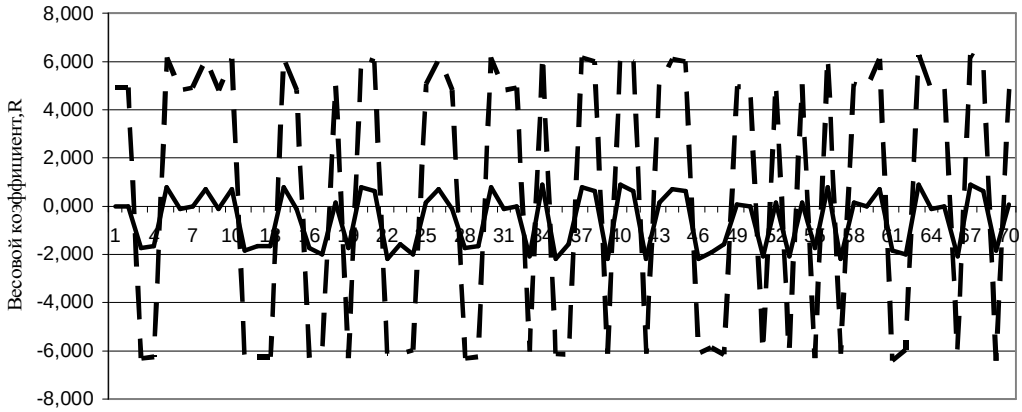
Анализируемая последовательность ДНК, кодирующая сегмент 16S–рРНК E.coli., представлена на рис. 3, а ее табуляция по обозначенным выше физическим параметрам позволила получить графические представления соотношений (9)–(12) (без коэффициента K). Учитывая коэффициенты K по (7) и (8) соответственно, получим (рис. 4, 5, 6) графические зависимости по формулам (10)–(12).

AAATTGAAGAGTTTGATCATGGCTCAGATTGAACG
CTGGCGGCAGGCCTAACACATGCAAGTCGAACGGT

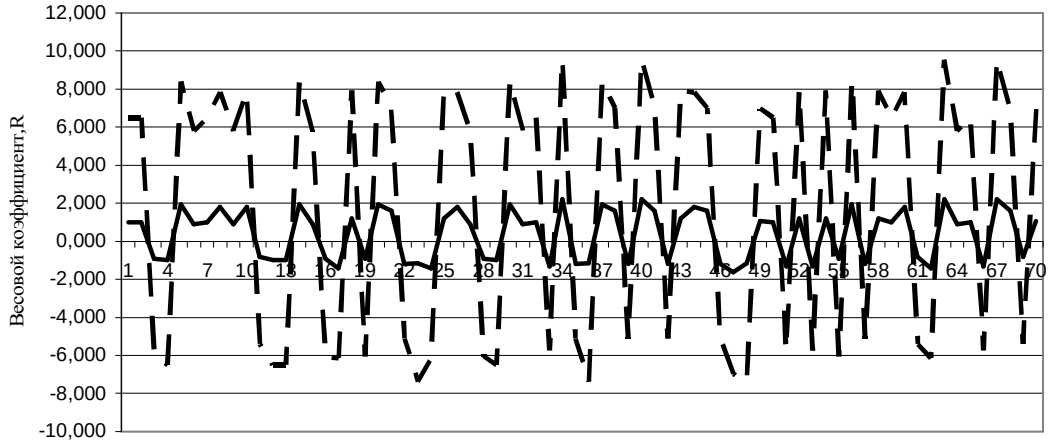
Р и с. 3. Нуклеотидная последовательность ДНК, кодирующая сегмент 16S-рРНК E. coli., N=70



Р и с. 4. Модель представления нуклеотидной последовательности ДНК, кодирующей сегмент 16S-рРНК E.coli. С использованием (11) при N=70: – $(M_1 \cdot M_2) \cdot K_1$, – – – $(M_1 \cdot M_2) \cdot K_2$



Р и с. 5. Модель представления нуклеотидной последовательности ДНК, кодирующей сегмент 16S-рРНК E.coli. С использованием (10) при N=70: – $(M_1 - M_2) \cdot K_1$, – – – $(M_1 - M_2) \cdot K_2$



Р и с. 6. Модель представления нуклеотидной последовательности ДНК, кодирующей сегмент 16S-рРНК E.coli. С использованием (12) при N=70: – $(M_1 \cdot M_2) \cdot K_1$, – – – $(M_1 - M_2) \cdot K_2$

Применение коэффициента K позволяет получить информацию в графическом виде о силе водородных связей между комплементарными нуклеотидами и их пространственной ориентации относительно третьего или пятого углеродного атома дезоксирибозы. Следует отметить,

что применение (7) более предпочтительно, чем применение (8), так как дает информацию об относительной суммарной силе связи между комплементарными нуклеотидами, которая всего на 0,003 отличается от числа j — величины золотого сечения.

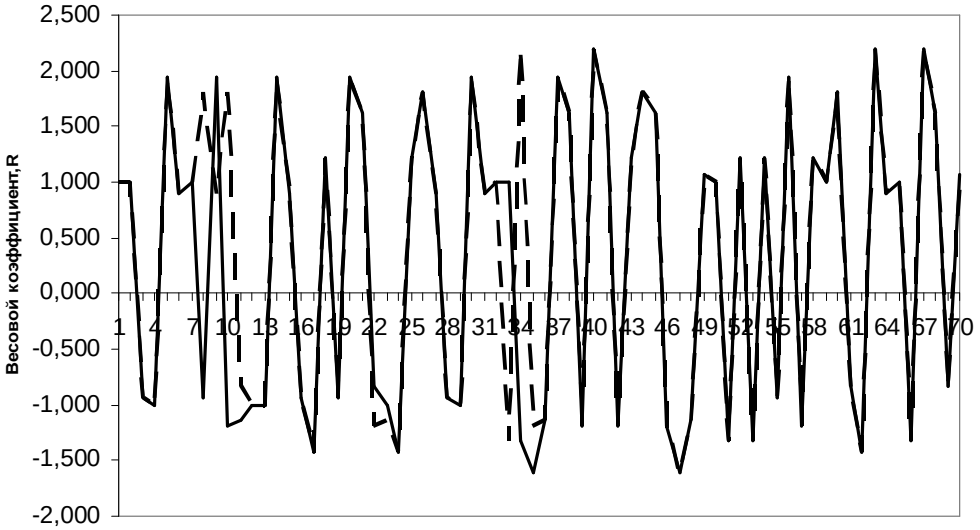
Условие (9) не позволяет судить о количественном соотношении молекулярных масс нуклеотидов. Соотношение (11) не позволяет выяснить саму разницу между нуклеотидами в их сочетаниях. Таким образом, условия (10) и (12) с массивом коэффициентов (7) дают наиболее чувствительный метод кодирования.

Модификация исходной последовательности выразилась в том, что элементы 9,10 и 11 (GAG) были заменены на элементы TGC, нуклеотид 22 (C) заменён на T, а пара 33,34 (CG) изменена на AC. Старая нуклеотидная последовательность (рис. 3) и новая (рис. 7) представлена на рис. 8.

Внесем возмущения в исходную последовательность путем замены нуклеотидов 9, 10 и 11 (G, A, G на T, G, C); нуклеотида 23 (C на T); нуклеотидов 34,35 (C, G на A, C) соответственно. Значения функции (12) с коэффициентом (7), полученные из исходной нуклеотидной последовательности (рис. 3) и модифицированной последовательности (рис. 7), представлены на рис. 8, соответственно. Видно, что с возрастанием области возмущения, кривые расходятся на все большем интервале.

AAATTGAATGCTT TGATCATGGTTCAGATTGAAAC
CTGGCGGCAGGCCTAACACATGCAAGTCGAACGGT

Р и с. 7. Нуклеотидная последовательность ДНК с возмущениями, кодирующая сегмент 16S-рРНК E. coli., N=70



Р и с. 8. Возмущения в нуклеотидной последовательности ДНК, кодирующей сегмент 16S-рРНК E.coli. с использованием (12) при N=70: – исходная последовательность, – – – последовательность с возмущениями

Предложенный в статье метод формализации нуклеотидных последовательностей позволяют сделать следующие выводы:

1. Критерии различной физической природы, положенные в основу формального представления AGTC-последовательностей, с высокой степенью идентичности графического представления формально отражают характер их последовательностной взаимосвязи, и, тем самым подтверждают объективность представленных данных.

2. Метод графической формализации удовлетворяет по разрешающей способности требованиям распознавания, что особенно важно применительно к сайтам «узнавания».

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Эйнгорин М. Я.* Основы кодирования и управления в молекулярной биологии. Н. Новгород: Нижегор. медиц. акад. 2001. 116 с.
2. *Бекасов Л. С.* О стратегиях моделирования генетических структур // Вестн. Сам. гос. техн. ун-та. Сер.: Физ.–мат. науки, 2005. Вып. 38. С. 158–162.
3. *Сингер М., Берг П.* Гены и геномы. М.: Мир. 1998. 373 с.
4. *Полтев В. И., Дерябина А. С., Гонзалез Э, Грохлина Т. И.* Взаимодействия между основаниями нуклеиновых кислот. Новые параметры потенциалов и новые минимумы энергии // Биофизика, 2002. Т. 47. С. 996–1004.

Поступила 20. 08.2006 г.

УДК 681.51:629.78

А. А. Тюгашев

АВТОМАТИЗИРОВАННОЕ ОЦЕНИВАНИЕ ВРЕМЕННЫХ ХАРАКТЕРИСТИК ПРОГРАММ УПРАВЛЕНИЯ КОМПЛЕКСНОГО ФУНКЦИОНИРОВАНИЯ В ТЕХНОЛОГИИ ГРАФКОНТ

Описывается один из элементов технологии автоматизированного проектирования управляющих программ реального времени для космических аппаратов ГРАФКОНТ – метод автоматизированного оценивания временных характеристик управляющих программ комплексного функционирования, и дальнейшее использование полученных данных. Под управляющими программами комплексного функционирования понимаются программы, предназначенные для координации работы бортовой аппаратуры, агрегатов и систем космических аппаратов при решении целевой задачи.

При проектировании бортового комплекса управления космического аппарата (БКУ КА) подсчет временных характеристик исполнения управляющих алгоритмов реального времени является одной из критически важных задач обеспечения работоспособности и надежности всего комплекса бортовой аппаратуры (БА).

Ключевой особенностью организации процесса функционирования бортовой вычислительной системы (БВС) является тот факт, что данная система является системой управления реального времени, и выполняющиеся в ней процессы реализуют управляющие алгоритмы реального времени (УА РВ).

Дополнительной отличительной особенностью управляющих бортовых цифровых вычислительных машин (БЦВМ) является то, что они взаимодействуют не с пользователем, а с физическими объектами [1], моменты поступления сигналов от них асинхронны и, в общем случае, носят труднопредсказуемый случайный характер.

Как известно [2], при организации БВС реального времени необходимо обеспечить выполнение следующих требований:

1) ограничение времени отклика, то есть гарантия того, что после наступления события реакция на него последует не позднее истечения заданного интервала времени $t_{кр}$;

2) одновременность обработки, то есть даже если одновременно наступает более одного события, все временные ограничения для всех событий должны быть выдержаны.

Это с необходимостью приводит к тому, что БВС реального времени должны поддерживаться решением параллельной работы программ, что достигается применением многозадачного подхода.

По условной классификации систем реального времени на «жесткие» и «мягкие», БВС КА должны быть отнесены к «жестким» системам реального времени, поскольку цена несоблюдения временного ограничения может быть неприемлемо высокой и приводить к невыполнению функциональной задачи, например, срыву слежения за астроориентиром, в более тяжелом случае — к потере дорогостоящего аппарата, а в случае пилотируемой экспедиции — к опасности для жизни экипажа.