

11
0287



Ф.Е. Френкель М.Б. Чалей Е.В. Коротков

011-2002

**MRFGS И MRFPS — БАЗЫ ДАННЫХ
ПО СКРЫТОЙ ПЕРИОДИЧНОСТИ
В ПОСЛЕДОВАТЕЛЬНОСТЯХ ДНК И БЕЛКОВ**

Москва 2002

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

МИНИСТЕРСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ ПО АТОМНОЙ ЭНЕРГИИ

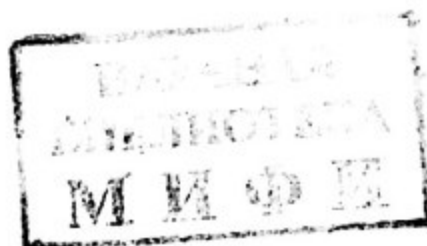
МОСКОВСКИЙ ИНЖЕНЕРНО-ФИЗИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

Ф.Е. Френкель М.Б. Чалей Е.В. Коротков

**MRFGS И MRFPS — БАЗЫ ДАННЫХ
ПО СКРЫТОЙ ПЕРИОДИЧНОСТИ
В ПОСЛЕДОВАТЕЛЬНОСТЯХ ДНК И БЕЛКОВ**

Препринт 011-2002

Москва 2002



УДК 577.29:00465

ББК 28.072

Ф87

Ф.Е. Френкель, М.Б. Чалей, Е.В. Коротков.

**MRFGS и MRFPS – БАЗЫ ДАННЫХ ПО СКРЫТОЙ ПЕРИОДИЧНОСТИ
В ПОСЛЕДОВАТЕЛЬНОСТЯХ ДНК И БЕЛКОВ.** Препринт 011-2002. М.:
МИФИ, 2003. 20 с.

В результате анализа баз данных(БД) GenBank-116 и SwissProt-38 оригинальным пакетом программ поиска периодичности созданы первые версии БД по периодичности в нуклеотидных (MRFGS) и аминокислотных (MRFPS) последовательностях (Miscellaneous Repetitive Features in Genetic and Protein Sequences). БД MRFGS содержит периодические последовательности ДНК/РНК с длиной периода кратной трем: 6, 9, 12, ..., 120 нуклеотидов. БД MRFPS содержит последовательности белков с периодичностью от 2 до 1046 аминокислотных остатков, причем последовательности с длинами периода более 100 аминокислотных остатков составляют только 14% полного числа периодических последовательностей БД MRFPS. Число записей в БД MRFGS и MRFPS равно, соответственно, 12067 и 1072. Обе БД могут быть полезны для изучения связи между характерной периодичностью гена или белка и их функциональной ролью.

Рекомендовано к изданию редсоветом МИФИ

Проект создания БД MRFGS и MRFPS был поддержан РФФИ, грант № 99-07-90154.

© Ф.Е. Френкель, М.Б. Чалей, Е.В. Коротков, 2003

© Московский инженерно-физический институт
(государственный университет), 2003

Введение

Еще 10 лет назад научное сообщество не представляло в полной мере, каких огромных усилий потребует систематизация знаний о последовательностях геномов и белков, чтобы послужить надежной основой для теоретических предсказаний структуры и механизмов взаимодействий молекул клетки. Вступив в эпоху геномики и протеомики, когда почти экспоненциально растет число разнообразных биологических баз данных и появляются новые методы анализа последовательностей ДНК/РНК и белков, мы все еще находимся только у основания огромной пирамиды знаний о том, как работает молекулярно-генетический аппарат клетки.

В течение последних пяти лет основные усилия научной группы сконцентрированы на изучении периодичности нуклеотидных и аминокислотных последовательностей и ее возможной взаимосвязи с пространственной структурой и функциями нуклеотидных последовательностей и белков [1-6]. Ранее скрытая периодичность в 21 основание была найдена у многих генов хеморецепторов различных бактерий [4] в областях, соответствующих цитоплазматическому домену МСР-белков, который, как полагают, состоит из нескольких α -спиралей [7]. Было отмечено, что скрытая периодичность в 19 аминокислот типична для NAD^+ -связывающих сайтов различных белков, периодичность в две аминокислоты может быть характерна для трансмембранных доменов разного рода рецепторов [5]. Причинами существования скрытой периодичности в нуклеотидных последовательностях могут быть либо многократные дубликации некоторого фрагмента ДНК/РНК, сопровождаемые дивергенцией оснований, либо процесс внутренней конвергенции, поддерживаемый адаптивной эволюцией структуры либо ДНК, либо кодируемого ею белка. Вероятное биологическое значение скрытой периодичности в том, чтобы служить своего рода масштабной разметкой последовательностей при формировании комплексных структур с другими белками и нуклеиновыми кислотами.

Методы и источники данных

Разумеется, что наиболее полное изучение того, как часто встречается скрытая периодичность в белках и нуклеотидных последовательностях, а также ее значение для характеристики определенных функциональных районов, возможно только с созданием специальных баз данных (БД), доступных для анализа широкому кругу молекулярных биологов. Нами создана первая версия БД по периодичности в нуклеотидных и аминокислотных последовательностях (Miscellaneous Repetitive Features in Genetic (MRFGS) and Protein Sequences (MRFPS)). Результаты, составляющие содержание этих БД, были получены в результате анализа последовательностей из БД GenBank-116 и SwissProt-38. Периодические последовательности в белках и ДНК были выявлены оригинальным методом поиска скрытой периодичности (далее метод информационной декомпозиции), который выявляет широкий спектр периодичности от совершенной и сильно размытой, хотя все еще видимой, до скрытой периодичности [5,6]. Периодичность, которую мы называем скрытой, можно идентифицировать лишь на основе неслучайной статистики оснований (или аминокислот) в отдельных сайтах периода. Причем, вовсе не обязательно преобладание какого-то одного типа основания или аминокислоты в таких сайтах. В случае нуклеотидных последовательностей это могут быть и два и три типа оснований, а в аминокислотных последовательностях – даже и более многочисленные группы из 20 возможных аминокислот.

Статистическим критерием значимости выделяемого скрытого периода является величина $Z = (I(1) - I'_m) / \sigma$, называемая Score, подробное описание которой можно найти в работах [4-6]. Для нуклеотидных и аминокислотных последовательностей значение Z более 5.0 соответствует случайной вероятности обнаружения скрытого периода менее чем 10^{-6} [4,5].

При создании первой версии БД скрытой периодичности белковых и генетических последовательностей мы выбрали критический уровень Z выше минимального уровня значимости найденной периодичности ($Z = 5$): $Z = 8$ как для нуклеотидных, так и аминокис-

лотных последовательностей, ограничив тем самым данные о периодичности в первой версии БД MRFGS и MRFPS периодичностью совершенной и почти совершенной. Поэтому скрытая периодичность составила незначительную часть обеих БД. Это ограничение было сделано сознательно, так как мы полагаем, что подойти к пониманию биологического смысла скрытой периодичности можно будет только после исчерпывающего исследования связи между сильно размытой периодичностью и ее функциональной, структурно-организующей ролью в генетических последовательностях и белках.

На сегодня первая версия БД MRFGS и MRFPS доступна на компакт-дисках. Интерфейс реализован по технологии Web, позволяющей в дальнейшем сделать банки данных доступными через сеть Интернет.

Структура записей баз данных

БД MRFGS и MRFPS имеют подобную друг другу структуру записей. Каждый случай периодичности, найденный в БД GenBank или SwissProt, получает свой уникальный код доступа (ключ) "Periodicity accession number", при этом также фиксируется Accession number исходной последовательности из GenBank или SwissProt, в которой была найдена периодичность. Записываются координаты области периодичности, его полная последовательность, уровень статистической значимости Z (Score) найденной периодичности и длина периода.

В БД MRFGS собраны случаи периодичности в последовательностях ДНК/РНК с длиной периода кратной трем: 6, 9, 12, ..., 120 оснований. Всего в БД MRFGS содержатся 12 067 таких случаев. На рис.1 показан график распределения численности случаев периодичности в зависимости от длины периода. Как можно видеть, наиболее многочисленна периодичность в последовательностях ДНК/РНК с периодом в 6 оснований (2555 последовательностей). Численность остальных случаев распределена в среднем равномерно, с некоторым преимуществом для периодов в 9 оснований (516

случаев), 12 оснований (690 случаев), 15 оснований (528 случаев), 21 основание (569 случаев) и 84 основания (616 случаев). Полный объем первой версии БД MRFGS занимает немногим более 42 Мбайт на жестком диске.

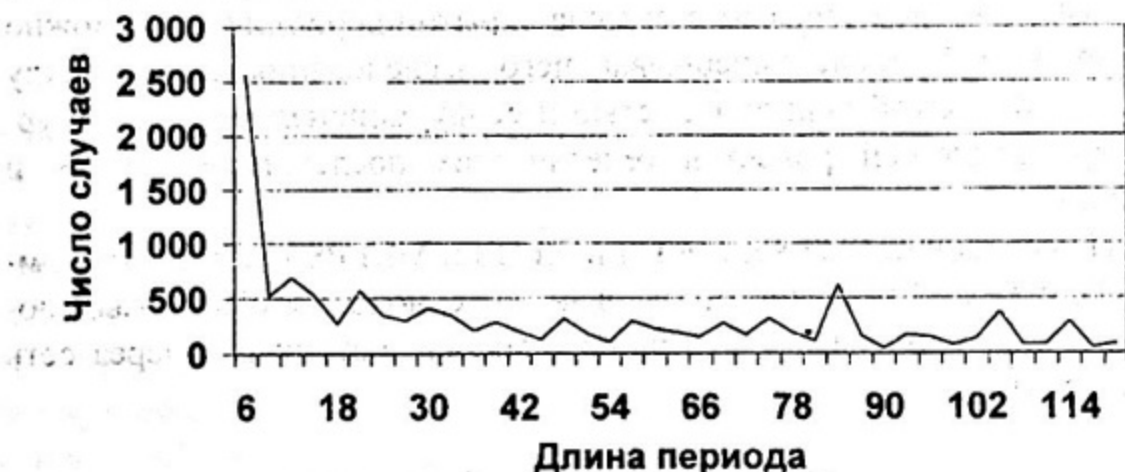


Рис.1. Распределение численности случаев периодичности в БД MRFGS в зависимости от длины периода

В первой версии БД MRFPS собраны 1072 случая периодичности в последовательностях белков, которые составляют около 4.5% от общего числа последовательностей в БД SwissProt-38. Периодичность в белках представлена длинами периодов в 2, 3, ..., 1046 аминокислотных остатков. Рис.2 показывает спектр распределения численности случаев периодичности в БД MRFPS в зависимости от длины периода. Наиболее распространенными являются случаи периодичности в 28 аминокислот (145 случаев), три аминокислоты (116 случаев), а также в две аминокислоты (79 случаев). Имеются только шесть случаев с длиной периода более 1000 аминокислотных остатков. Вероятнее всего, что периодичность более 100 аминокислотных остатков обусловлена дубликациями в последовательностях генов соответствующих белков.



Рис.2. Распределение численности случаев периодичности в БД MRFPS в зависимости от длины периода

Интерфейс пользователя

На домашней странице (рис.3) можно выбрать интересующую БД (по периодичности в нуклеотидных последовательностях (MRFGS) и белках (MRFPS)) или предварительно изучить описания к ним на соответствующих страницах. Обе БД предоставляют одинаковый интерфейс для запроса информации. Начальные страницы БД (рис.4) предназначены для оформления запроса пользователя путем заполнения соответствующих полей ввода. Чтобы узнать, есть ли периодичность в какой-либо последовательности из GenBank (или SwissProt), достаточно ввести соответствующий accession number в поле "Accession number substring", оставив остальные поля не заполненными. Если пользователь уже работал с БД MRFGS и MRFPS, он может послать запрос, указав Periodicity accession number нужной последовательности в соответствующей БД. Можно оформить запрос о наличии периодических последовательностей с фиксированной длиной периода или заданной в определенном интервале. Рис.5 демонстрирует список последовательностей с периодичностью в 102, полученный в ответ на запрос.

DataBases of Miscellaneous Repetitive Features in Genetic(MRFGS) and Protein(MRFPS) Sequences

MRFGS DB 1.0

MRFPS DB 1.0

About MRFGS database

About MRFPS database

Рис.3. Домашняя страница БД MRFGS и MRFPS

Enter the request conditions

Accession number substring:

Minimal period length:

Maximal period length (leave zero to skip period length matching):

Periodicity accession number (leave zero to skip period length matching):

Submit

Рис.4. Страница формирования запросов к БД MRFGS

For the description of table columns see [the following page](#)

Periodicity Accession Number	GenBank Accession	Period length	Score	Left coordinate	Right coordinate
7	AB021363	102	51.99	1347	2673
8	AB021364	102	51.94	1345	2678
9	AB021365	102	41.58	1281	2270
10	AB021365	102	49.72	2240	3529
62	D90900	102	14.75	130743	131321
66	D90903	102	30.64	65341	66735
67	D90903	102	21.6	67041	68120
105	U23947	102	49.81	7301	8697
259	M58002	102	19.31	1589	2259
276	M80803	102	16.41	1031	1398
277	L06634	102	49.07	1423	2680
278	X68781	102	54.23	1281	2678
279	X16130	102	47.74	1487	2678
283	U50552	102	53.62	1281	2680
303	AE000710	102	14.5	12393	13436
307	AE000798	102	19.69	5329	5930
308	AE000798	102	25.9	7005	7832
309	AE000799	102	24.99	3246	3980
356	AF012325	102	45.68	3001	4151
362	AF016221	102	48.68	1281	2511

NEXT

Рис.5. Список всех записей в БД MRFGS с длиной периода в 102 основания

Гиперссылка на GenBank accession number (или SwissProt accession number) ведет к списку всех периодичностей, найденных в исходной последовательности из БД GenBank (SwissProt), представляемому в описанном выше виде (см. рис.5). Гиперссылка на Periodicity accession number индивидуальной последовательности в БД MRFGS (MRFPS) ведет на страницу Periodicity information (рис.6), где представлена полная информация об этой последова-

тельности. На получаемой странице перечислены: Periodicity accession number данной последовательности; accession number исходной последовательности из GenBank, в которой была найдена периодичность; длина периода; уровень значимости найденной периодичности Score (Z); левая и правая координаты выявленного района периодичности в исходной последовательности; полная последовательность оснований этого района.

Periodicity information

For the description of the periodicity information see [the following page](#).

Periodicity accession number: 64
 GenBank accession number: D90903
 Period length: 6
 Significance (score): 34.89
 Left coordinate: 3052
 Right coordinate: 3679

Sequence:

```
GAACCCGAGC CGGAACCTGA ACCCGAGCCG GAACCTGAGC CGGAACCTGA ACCCGAGCCG
GAACCTGAGC CGGAACCTGA ACCCGAGCCG GAACCTGAAC CCGAGCCGGA ACCTGAACCC
GAGCCGGAAC CTGAACCCGA GCCGGAACCT GAACCCGAGC CGGAACCTGA ACCCGAGCCG
GAACCTGAAC CTGAGCCTGA ACCTGAGCCG GAACCTGAGC CGGAACCTGA ACCCGAGCCG
GAACCTGAGC CGGAACCTGA ACCCGAGCCG GAACCTGAGC CGGAACCTGA ACCCGAGCCG
GAACCTGAGC CGGAACCTGA ACCCGAGCCG GAACCTGAAC CCGAGCCGGA ACCTGAACCC
GAGCCGGAAC CTGAGCCGGA ACCTGAACCC GAGCCGGAAC CTGAGCCGGA ACCTGAACCC
GAGCCGGAAC CTGAGCCGGA ACCTGAACCC GAGCCGGAAC CTGAACCCGA GCCGGAACCT
GAGCCGGAAC CTGAACCCGA GCCGGAACCT GAGCCGGAAC CTGAACCCGA GCCGGAACCT
GAACCCGAGC CGGAACCTGA GCCGGAACCT GAACCCGAGC CGGAACCTGA GCCGGAACCT
GAACCCGAGC CGGAACCTGA GCCTGAGC
```

Frequency matrix:

	1	2	3	4	5	6
A	.	104	64	.	.	.
T	42
C	.	.	.	104	104	24
G	104	.	40	.	.	38

Рис.6. Начало страницы Periodicity information, содержащей полное описание периодической последовательности в БД MRFGS

Помимо этого, на странице Periodicity information приводится матрица встречаемости оснований А, Т, С, G в каждом из сайтов периода для выявленной периодической последовательности. На основе матрицы распределения оснований по сайтам периода строится график Chi-squared graphics, показывающий, насколько набор частот оснований в каждом сайте отличается от ожидаемого набора частот для данной последовательности. Отличие измеряется в соответствии с критерием χ^2 , имеющим число степеней свободы, равное длине периода. Далее на этой же странице строится спектр возможных периодичностей в данном районе (рис.7). Спектр показывает значение величины Z (уровня значимости) в зависимости от анализируемой длины периода. Наибольшее значение Z в спектре и соответствующая ему длина периода могут не совпадать со значением Score и длиной периода, указанными в начале страницы Periodicity information. Такое несоответствие возникает из-за того, что спектр лишь позволяет отобрать для анализа наибольшие значения Z и соответствующие им длины периода. Окончательный вывод о наличии периода определенной длины и статистической значимости периодичности в последовательности (Score) делается после проведения испытаний каждой отобранной длины периода методом Монте-Карло.

Поскольку, как уже сообщалось выше, структура записей (входов), а также процедура исполнения запроса пользователя в обеих БД практически одинакова, далее ограничимся лишь краткими замечаниями относительно БД MRFPS.

Аминокислотная последовательность периодического участка показана на странице Periodicity information периодичности в однобуквенном обозначении. Далее на странице Periodicity information находится матрица, содержащая количественное распределение аминокислот во всех сайтах периода, что позволяет наглядно представить, какая периодичность существует в данной последовательности белка: совершенная, размытая, сильно размытая или скрытая. Приводится также спектр значений Z для всех возможных длин периодов для данной последовательности.

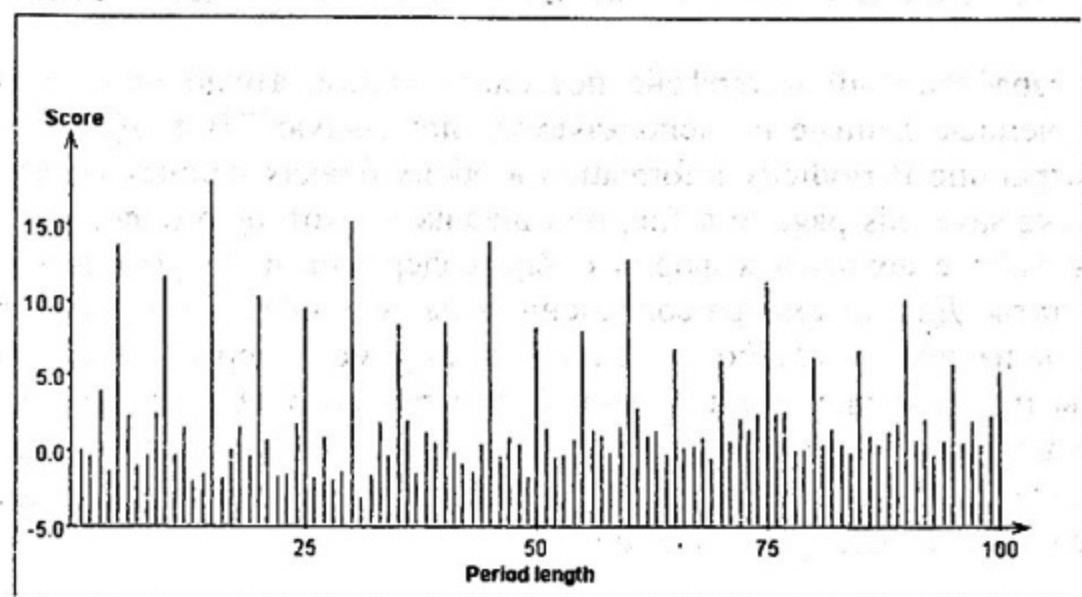


Рис.7. Спектр значений Z-статистической значимости всех возможных длин периодов для периодической последовательности №65 в БД MRFGS

Для всех последовательностей белков с периодичностью 34 аминокислотных остатков из БД MRFPS, за исключением одной (SwissProt accession P13983), в БД MRFGS содержатся соответствующие последовательности их генов. Даже на примере такой небольшой выборки пар периодический ген – периодический белок можно видеть, что почти совершенная периодичность в гене соответствует почти совершенной периодичности в белке (MRFGS 4569 – MRFPS 269, MRFGS 279 – MRFPS 645). Сильно размытая периодичность в гене обуславливает сильно размытую периодичность в белке (MRFGS 9884 – MRFPS 369). И наконец, скрытая периодичность в гене переходит в очень сильно размытую периодичность в белке (MRFGS 522 – MRFPS 653).

Перенос данных по периодичности на ЭВМ пользователя

Разработанный интерфейс позволяет пользователю переносить полученные данные на используемую локальную ЭВМ. Для этого на странице Periodicity information в обоих банках данных имеется ссылка save this page to a file, при щелчке на которую будет загружен файл с архивом в формате Zip, содержащим полученные результаты. Для просмотра сохраненных данных необходимо извлечь файлы из архива в любой каталог используемой операционной системы и с помощью установленной программы Web-навигации открыть в выбранном каталоге файл page.html. Представленные в нем данные полностью совпадают с содержанием соответствующей страницы Periodicity information.

Сравнительный анализ созданной БД с существующими аналогами

Для сравнения используемого алгоритма информационной декомпозиции с существующими методами выявления периодичности в генетических последовательностях были исследованы:

- преобразование Фурье на примере одной из его реализаций [8],
- оригинальный метод и программа поиска тандемных повторов Tandem Repeat Finder д-ра Г. Бенсона [9].

Сопоставление результатов анализа последовательностей преобразованием Фурье с данными, полученными методом информационной декомпозиции, выявило снижение чувствительности первого при больших длинах периода. При преобразовании Фурье данные о длинных периодах аккумулируются в коротких обертонах, которые, в свою очередь, не имеют значимого уровня сигнала, так как для анализа символьных последовательностей их необходимо преобразовывать в числовые, что при длине периода, значительно превосходящей размер алфавита, приводит к появлению ложных коротких периодов для некоторых символов. Для примера возьмем искусственную последовательность ДНК (алфавит 'A', 'T', 'C', 'G') с явным

периодом в 6 оснований – **ACGATTACGATTACGATTACGATT...**
Результат анализа данной последовательности представлен на рис.8.



Рис.8. График веса гармоник для различных длин периодов искусственной последовательности с периодом ACGATT

Преобразование Фурье выявляет период в три основания (при заданных шести). Это объясняется тем, что периодичность в три символа для основания 'А' уменьшает значимость периода в 6 символов для остальных оснований.

Таким образом, можно отметить, что при анализе преобразованием Фурье символьных последовательностей с короткими алфавитами наблюдается эффект ослабления гармоник длинных периодов в пользу гармоник коротких. Данный эффект усиливается при наличии модификаций в периодах.

Метод поиска тандемных повторов д-ра Бенсона основан на статистически значимом сходстве между периодами, выявляемом при наличии последовательно расположенных идентичных слов длиной в несколько символов. Алгоритм позволяет обнаруживать периоды, имеющие вставки и делеции, наличие которых в последовательности может быть критичным при поиске в ней периодичности. Происходящий при таких мутациях сдвиг фазы периода приводит к неспособности алгоритмов, не учитывающих вставки и делеции, об-

наружить периодичность. Однако использование в данном методе в качестве меры подобия периодов гомологии не позволяет выявлять случаи скрытой периодичности. Поэтому результаты анализа генетических последовательностей данным методом и описываемого здесь метода информационного подобия образуют пересекающиеся (в области явной периодичности без вставок и делеций) множества, каждое из которых не содержит другое.

Приведем для примера случай скрытой периодичности, не обнаруживаемой ни преобразованием Фурье, ни методом поиска тандемных повторов:

LOCUS AE001716 (длина 16931 пар нуклеотидов (н.п.)).

Координаты области с периодичностью: 1-523.

Периодичность в 67 н.п. найдена в следующей подпоследовательности (показаны три первых периода):

[CAATGGGTTTCAATAATTCSTTAGAGGTATGGAAACAGACAATCGTGGATCAGGTCATCAATGCCC
T]

[GAATTAGTTTCAATAATTCSTTAGAGGTATGGAAACACTTGTGGAGATTGAGGAACATCTAGGAAA
G]

[AACATCGTTTCAATAATTCSTTAGAGGTATGGAAACAGCGATAACATTCALTTCSTTCGACGAGGTA
C]...

Статистическая значимость периодичности: 8.23 (минимум - 5.00).

При поиске преобразованием Фурье для вычисления статистической значимости полученных данных (веса гармоник для различных длин периодов) были проведены серии испытаний случайных последовательностей с символьным составом, идентичным данной последовательности. При генерации случайных последовательностей распределение символов было сохранено с учетом их положения в триплетах. Таким образом в случайных последовательностях воспроизводилась имеющаяся триплетная периодичность, что исключало ее влияние на результаты исследований. Так как в оригинальной статье [8] указывается, что распределение весовой функции гармоник близко к нормальному, то для вычисления статистической значимости вычислялось стандартное отклонение веса гармоники от случайного.

Графики периодичности на указанном выше участке последовательности AE001716 методом информационной декомпозиции и методом преобразования Фурье представлены на рис.9.

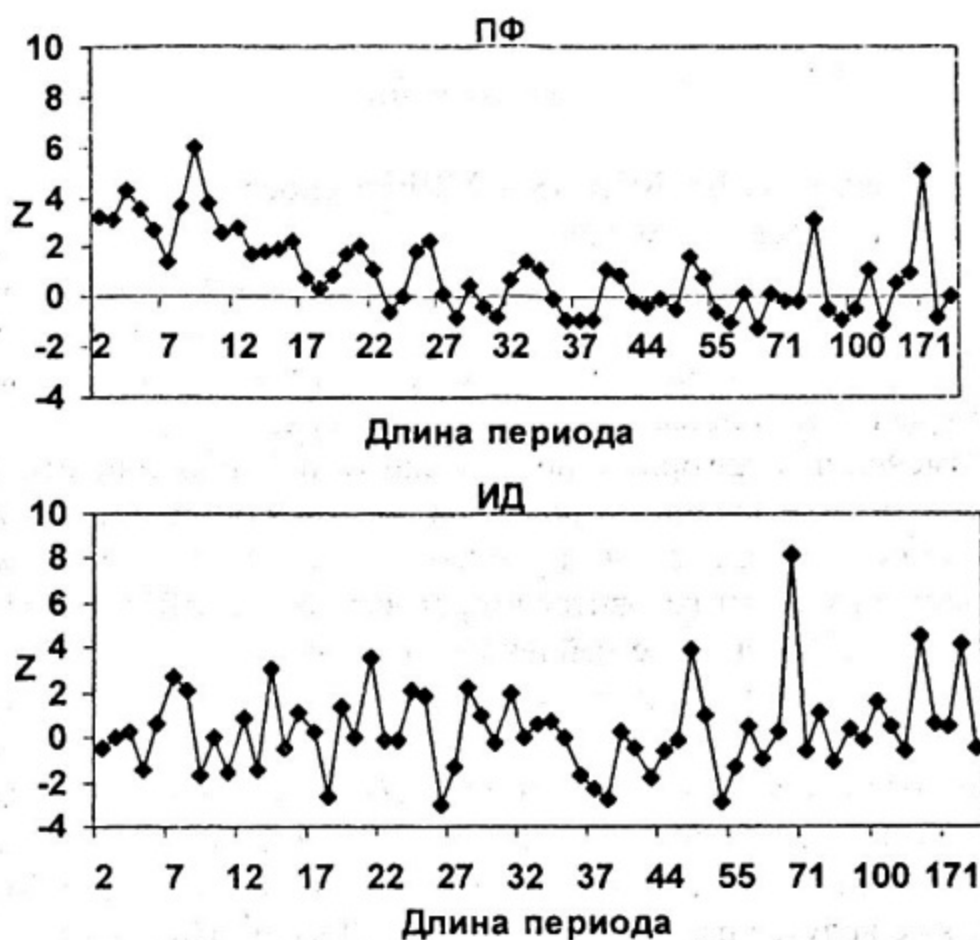


Рис.9. Графики периодичности в последовательности АЕ001716 (участок 1-523), найденной методом преобразования Фурье (график ПФ) и методом методом информационной декомпозиции (график ИД). Здесь Z – стандартное отклонение значения весовой функции периодичности для данной длины периода от случайного

Программа поиска тандемных повторов также не находит периодичность ввиду отсутствия между периодами прямой гомологии.

Заключение

Созданная пара БД MRFGS – MRFPS способствует пониманию того, в каких пределах может существовать связь между совершенной, размытой и скрытой периодичностью гена и периодичностью белка. Кроме того, обе БД могут быть полезны в выявлении характерных периодичностей, связанных с элементами пространственной структуры и функциональных сайтов в белках.

Примененный алгоритм поиска скрытой периодичности имеет существенные отличия от аналогов, позволяющие ему выявлять скрытую периодичность, без прямого подобия между периодами.

В данный момент готовится вторая версия БД MRFGS и MRFPS, куда будут включены все найденные аминокислотные и нуклеотидные последовательности, содержащие скрытую периодичность с уровнем значимости Z больше 5.0. Число периодических нуклеотидных последовательностей в выпуске GenBank-116 составляет немногим более 600 000, а число периодических аминокислотных последовательностей в выпуске БД SwissProt-38 составляет более 20 000. Также ведутся работы по переносу БД в сеть Интернет.

По вопросам получения банков данных MRFGS и MRFPS можно обращаться по адресу: katrin22@mtu-net.ru.

Список литературы

1. Korotkov E.V., Korotkova M.A. DNA regions with latent periodicity in some human clones // DNA Seq. 1995. V.5. P.353-358.
2. Korotkov E.V., Korotkova M.A. Enlarged similarity of nucleic acid sequences // DNA Res. 1996. V.3. P.157-164.
3. Korotkov E.V., Korotkova M.A., Tulko J.S. Latent sequence periodicity of some oncogenes and DNA-binding protein genes // CABIOS. 1996. V.13. P.37-44.
4. Chaley M.B., Korotkov E.V., Skryabin K.G. Method revealing latent periodicity of the nucleotide sequences modified for a case of small samples // DNA Res. 1999. V.6. P.153-163.
5. Korotkova M.A., Korotkov E.V., Rudenko V.M. Latent periodicity of protein Sequences // J. Mol. Model. 1999. V.5. P.103-115.
6. Korotkov E.V., Korotkova M.A., Rudenko V.M., Skryabin K.G. Latent Periodicity Regions in Amino Acid Sequences // Molecular Biology (Russian). 1999. V. 33(4). P.688-695.
7. Mowbray S.L., Sandren M.O.J. Chemotaxis Receptors: A progress report on structure and function // J. Struct. Biol. 1998. V.124. P.257-275.
8. Issac B, Singh H, Kaur H, Raghava GP. Locating probable genes using Fourier transform approach // Bioinformatics. V.18(1). P.196-197.
9. G. Benson. Tandem repeats finder: a program to analyze DNA sequences // Nucleic Acid Research. 1999. V.27(2). P.573-580.

Феликс Ефимович Френкель

Мария Борисовна Чалей

Евгений Владимирович Коротков

**MRFGS И MRFPS — БАЗЫ ДАННЫХ
ПО СКРЫТОЙ ПЕРИОДИЧНОСТИ
В ПОСЛЕДОВАТЕЛЬНОСТЯХ ДНК И БЕЛКОВ**

Препринт 011-2002

Рукопись поступила в РИО 06.12.2002

Ответственный за выпуск Ф.Е. Френкель

Редактор Н.Н. Антонова

Подписано в печать 13.03.2003. Формат 60x84 1/16.

Печ.л. 1,25. Уч.-изд.л. 1,25. Тираж 100 экз.

Изд. № 011-2002. Заказ № 261

*Московский инженерно-физический институт
(государственный университет).*

Типография МИФИ.

115409, Москва, Каширское ш., 31