

АКАДЕМИЯ НАУК СССР
НАУЧНЫЙ ЦЕНТР БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ

Н.Л.ЛУНИНА
СИСТЕМА ОБРАБОТКИ НУКЛЕОТИДНЫХ
ПОСЛЕДОВАТЕЛЬНОСТЕЙ HEID

МИНИ-ЭВМ

МАТЕРИАЛЫ ПО
МАТЕМАТИЧЕСКОМУ
ОБЕСПЕЧЕНИЮ ЭВМ

9

ПУЩИНО • 1984

АКАДЕМИЯ НАУК СССР
НАУЧНЫЙ ЦЕНТР БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ
ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР

МАТЕРИАЛЫ ПО МАТЕМАТИЧЕСКОМУ
ОБЕСПЕЧЕНИЮ ЭВМ

Серия Мини-ЭВМ, Выпуск 9

Н.Л.Лунина

СИСТЕМА ОБРАБОТКИ НУКЛЕОТИДНЫХ
ПОСЛЕДОВАТЕЛЬНОСТЕЙ NEID

ПУЩИНО•1984

УДК 681.3.06

Описана диалоговая система обработки нуклеотидных последовательностей HEID, приведены инструкция по эксплуатации и пример использования.

В качестве данных используются последовательности Гейдельбергского банка. В системе принята определенная их организация и содержатся средства преобразования данных к нужному виду. Система содержит также средства поиска и отбора нужного для работы подмножества последовательностей. Для каждой из отображенных последовательностей система позволяет ответить на ряд часто возникающих при исследовании вопросов: наблюдение статистических закономерностей в распределении нуклеотидов, поиск заданных участков, различного рода повторов, рестриктов и открытых рамок считывания, определение считываемых с последовательности белков и т.п.

Система реализована на языке Фортран и функционирует в рамках RSX-11M на CM-4.

1. ВВЕДЕНИЕ

В настоящем выпуске описана диалоговая система обработки нуклеотидных последовательностей HEID, приведены инструкции по эксплуатации и пример использования.

В качестве данных используются последовательности Гейдельбергского банка. В системе принята определенная их организация и содержатся средства преобразования данных к нужному виду. Система содержит также средства поиска и отбора нужного для работы подмножества последовательностей. Для каждой из отобранных последовательностей система позволяет ответить на ряд часто возникающих при исследовании вопросов: наблюдение статистических закономерностей в распределении нуклеотидов, поиск заданных участков, различного рода повторов, рестриктов и открытых рамок считывания, определение считываемых с последовательности белков и т.п.

Система была разработана в НИВЦ АН СССР по инициативе А.С.Кондрашова. В ходе разработки системы был проведен ряд ценных обсуждений с сотр. ИБ АН СССР Л.А.Ворониным, значительно способствовавших пониманию задачи. Полезные замечания по отдельным программам были высказаны сотрудниками ИБФМ АН СССР В.В.Вельковым (задача рестрикции) и В.М.Крюковым (задача трансляции). Последовательности, представленные В.М.Крюковым, неоднократно использовались в качестве тестового материала для отладки программ системы. Алгоритм вычисления изоэлектрической точки белка был подсказан сотрудником ИБ АН СССР А.В.Финкельштейном.

Первая версия системы была реализована на ЭВМ ЕС-1040. Совместно с М.А.Исаевым был выполнен перенос ее на СМ-4 в операционную систему RSX-11M. Дальнейшее развитие системы получила именно на СМ-4.

Неизменную помощь в деле системного обеспечения оказывал Л.И.Подольский, технического – В.П.Кузнецов. Тщательное редактирование рукописи проведено В.В.Левитиным.

Всем названным товарищам автор выражает глубокую благодарность.

2. КОНЦЕПЦИИ И ВОЗМОЖНОСТИ

2.1. Общее описание

Система обработки нуклеотидных последовательностей HEID предназначена для решения ряда задач, часто возникающих при исследовании первичной структуры рибонуклеиновых кислот.

В системе хранятся все последовательности Гейдельбергского банка данных (EMBL Nucleotide Sequence Data Library.. Release 1.1. European Molecular Biology Laboratory). Однако пользователь может обрабатывать и другие последовательности, заранее подготовив их на диске в нужном формате. Обычно объектом исследования является одна или несколько последовательностей. Для отбора этих последовательностей пользователь специальным образом формулирует условия, которым они должны удовлетворять (например, "все транспортные РНК, выделенные из такой-то группы организмов", или "все последовательности, напечатанные в таком-то журнале за такой-то год"). Если заранее известно, какая именно последовательность должна быть подвергнута обработке, можно просто указать ее имя; поиск в банке в этом случае не производится.

С отобранным множеством последовательностей можно провести следующие виды обработки:

- 1) печать последовательности;
- 2) просмотр последовательности на дисплее;
- 3) просмотр описания последовательности;
- 4) сравнение двух последовательностей;
- 5) обработка первичной структуры, а именно:
 - подсчет плотности распределения нуклеотидов и пар нуклеотидов;
 - построение таблицы ближайших соседей;
 - подсчет сочетаний нуклеотидов по два и по три;
- 6) поиск участков, а именно:
 - поиск участков данного вида;
 - поиск участков со слабой гомологией;
 - поиск прямых повторов;

- поиск инвертированных повторов;
- поиск участков двойной симметрии ("шпилек");
- 7) поиск рестриктов при совместной рестрикции;
- 8) поиск открытых рамок считывания.

Многие из указанных видов обработки требуют от пользователя дополнительной информации, которую можно задавать в диалоговом режиме. Существует возможность заранее подготовить информацию такого рода на диске, что является в ряде случаев более удобным (например, при одинаковой обработке большого числа последовательностей).

Выходную информацию всегда можно посмотреть на дисплее и по желанию сохранить на диске и напечатать на АЦПУ.

2.2. Организация и преобразование данных в системе

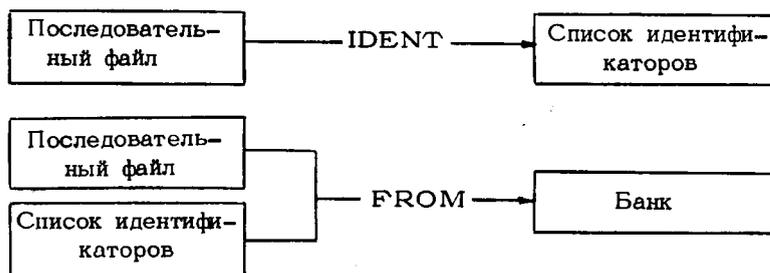
Логической единицей данных в системе является одна нуклеотидная последовательность, представленная в формате Гейдельбергского банка (см. 3.1).

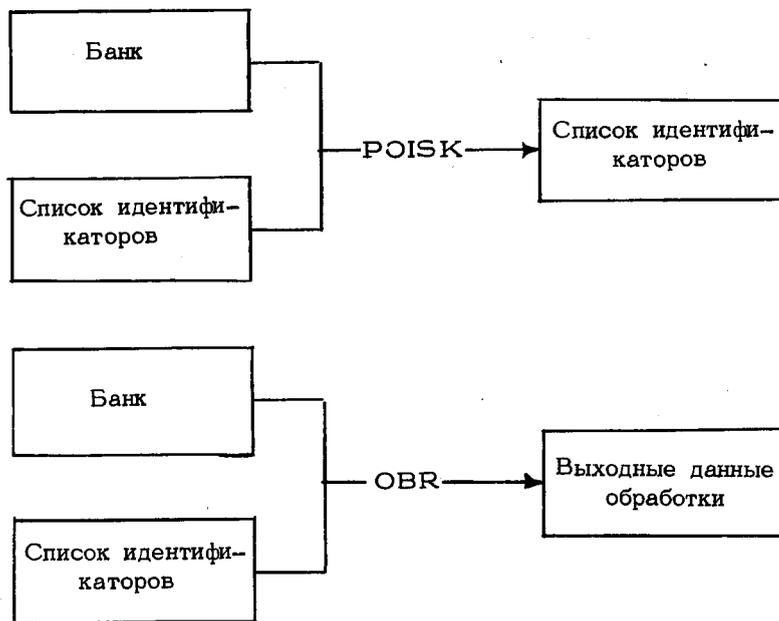
Чтобы обеспечить эффективное обращение к нужной последовательности, в системе принята определенная их организация: одна последовательность - один файл (в смысле операционной системы RSX-11M); имя файла однозначно связано с идентификатором последовательности.

Такая организация данных не является традиционной. В то же время распространен способ обмена данными в виде последовательного файла. Поэтому система содержит средства преобразования данных из последовательного файла к нужному виду. Эту функцию выполняют задачи IDENT и FROM.

Задача POISK выполняет функцию поиска и отбора нужного для работы подмножества последовательностей. Задача OBR выполняет обработку отобранного множества последовательностей (в частности, одной последовательности).

Средством описания нужного подмножества последовательностей является список их идентификаторов. Задачи FROM, POISK, OBR используют такой список в качестве входных данных; задачи IDENT и POISK создают такой список в качестве выходных данных. Схема движения данных в задачах системы такова:





Задача IDENT просматривает последовательный файл и получает список идентификаторов всех содержащихся в нем последовательностей.

Задача FROM просматривает последовательный файл и "вырезает" из него те последовательности, которые указаны во входном списке идентификаторов. В частном случае, таким списком может быть выходной список задачи IDENT.

Задача POISK для каждой последовательности, принадлежащей указанному подмножеству, отвечает на вопрос: удовлетворяет ли она поисковому запросу. Если да, идентификатор последовательности заносится в выходной список.

Задача OBR для каждой последовательности, принадлежащей указанному подмножеству, выполняет указанные виды обработки; выходные данные можно посмотреть на дисплее и напечатать на АШПУ.

2.3. Возможности диалога

Система HEID ведет диалог с пользователем на русском языке. При необходимости получения от пользователя ответа в виде текста система напоминает допустимые ответы. Ответы в виде текста (за исключением собственно нуклеотидов) можно вводить на любом регистре. Названия режимов можно сокращать до одного символа. При ответах типа "да/нет" пустой ответ всегда считается утвердительным.

Ответы числового типа вводятся в свободном формате (то есть количество позиций не имеет значения), проверяются на присутствие недопустимых символов и удовлетворение ограничениям, вытекающим из смысла задачи (например, номер нуклеотида должен находиться в пределах от единицы до длины последовательности, значение водородного показателя pH – от 2 до 12 и т.п.).

При получении некорректного ответа вопрос повторяется.

3. ПОИСК

3.1. Структура последовательности в Гейдельбергском банке

Последовательности в Гейдельбергском банке хранятся вместе со своими описаниями. В описании указана информация разного типа: например, дата ввода последовательности в банк, автор, литературный источник, организм, из которого последовательность была выделена, его систематика и т.п. Эта информация хранится в разных полях описания последовательности; каждое поле имеет двухбуквенное наименование:

ID - поле идентификации;
DT - дата;
DE - описание;
KW - ключевые слова;
OS - спецификация организма;
OC - систематика организма;
RN - поле ссылок;
RA - поле авторов;
RT - название работы;
RL - печатный источник;
CC - поле комментариев;
FH - поле разметки;
FT - таблица особенностей;
SQ - поле с указанием длины и структуры последовательности.

Пример последовательности:

```
ID  SCRNA6  SACCHARO. CEREVI. 17SRRNA. 6SRRNA; DNA; 423 BP.
DT  25-MAR-1981  FIRST ENTRY
DE  YEAST GENES FOR 17S RIBOSOMAL RNA AND 5.8S RIBOSOMAL RNA.
KW  RIBOSOMAL RNA
OS  SACCHAROMYCES CEREVISIAE (YEAST, , HEFE)
OC  EUKARYOTA; PLANTA; MYCOPHYTA; PHICOMYCETES.
RN  [1] (BASES 1-423)
RA  VELDMAN G. M., BRAND R. C., KLOOTWIJK J., PLANTA R. J.;
RT  "SOME CHARACTERISTICS OF PROCESSING SITES IN
RT  RIBOSOMAL RNA OF YEAST";
RL  NUCL. ACIDS REE. 8:2907-2920(1980).
CC  KST SCE. 17S. 5. 8SRRNA
FH  KEY      FROM      TO      DESCRIPTION
```

```

FT  RRNA      <1    24      175 RRNA
FT  RRNA      <1    233     185 RRNA
FT  RRNA      380   >423     5.8 S RRNA
SQ  SEQUENCE  423 BP; 127 A; 70 C; 139 T; 86 G; 1 -
GGTGAAC TCG CGGAAGGATC ATTAAGAAA TTTAATAATT TTGAAATGG ATTTTTTTTG
TTTTGGC AAG AGCATGAGAG CTTTTACTGG GCAAGAAGAC AAGAGATGGA GA-TCCAGCC
GGGCCTG CGC TTAAGTGCGC GGTCTTGCTA GGCTTGTAAG TTTCTTTCTT GCTATCCAA
ACGGTGAGAG ATTTCTGTGC TTTTGTATA GGACAATTAA AACCGTTTCA ATACAACACA
CTGTGAGAT TGCATATCTT TGCAACTTTT TCTTTGGGCA TTCGAGCAAT CGGGGCCCAI
AGGTAACAAA CACAACAAT TTTATTTATT CATTAAATTT TTGTCAAAAA CAAGAATTT
CGTAAC TGG AATTTTAAAA TATTAARAC TTTCAACAAC GGATCTCTTG GTTCTCGCA
CGA

```

//

3.2. Возможности поиска

При поиске пользователь имеет возможность отобрать из всего банка или из указанного его подмножества те последовательности, которые удовлетворяют поисковому запросу. Далее мы будем говорить о них как об отобранном множестве последовательностей (ОМП).

Поисковый запрос представляет собой составленное по определенным правилам логическое выражение из требований к отдельным полям описания последовательности. Требование к одному полю описания последовательности — это выражение вида (имя-поля операция текст)

В качестве имени поля допускаются все существующие в банке поля описания последовательности (ID, DE, DT и т.д.).

В качестве операции допускаются две операции: "=" в смысле "содержит" и "≠" в смысле "не содержит".

В качестве текста указывается тот текст, наличие или отсутствие которого нужно проверить в соответствующем поле описания последовательности. Текстом считается все от знака операции до круглой скобки; в частности, в нем могут содержаться пробелы.

Примеры требований к отдельным полям:

```

(DE = RNA)
(DT = 1982)
(OC ≠ COLI)
(FT = READING FRAME)

```

В поисковом запросе требования к отдельным полям T_i соединяются логическими функциями F_i, каждая из которых есть "и" или "или", образуя выражения следующего вида:

```

[T1 F1 T12 F1 . . . T1N]
FO [T21 F2 T22 F2 . . . T2N2] FO . . .

```

При этом следует соблюдать следующие правила:

1. Выражение должно содержать по крайней мере одну пару квадратных скобок.
2. Квадратные скобки соединяются между собой функциями одного (любого) типа.

3. Внутри пары квадратных скобок требования к отдельным полям соединяются между собой функциями одного (любого) типа.

4. Логические функции внутри разных квадратных скобок могут как отличаться друг от друга, так и совпадать.

5. Логические функции внутри скобок могут как отличаться от функций между скобками, так и совпадать.

В случае нарушения этих правил выдается диагностическое сообщение; запрос игнорируется.

Примеры правильно составленных запросов:

```
[(DE = RNA)]
[(DE = DNA) ИЛИ (ID = DNA) ИЛИ (OC = DNA)]
[(DE=RNA) И (DE=RIBOSOMA)] ИЛИ [(DE=TRNA)]
[(DT=1982)] И [(OS=ADENOVIRUS) ИЛИ (OS=PLANTA)]
```

Ограничения:

1. Общая длина запроса - не более 320 символов.
2. Допускается не более 5 пар квадратных скобок.
3. Внутри скобок допускается не более 6 требований к отдельным полям.
4. Текст, который ищется в указанном поле, не должен быть длиннее 20 символов.

При нарушении п. 1-3 выдается сообщение об ошибке и запрос игнорируется. При нарушении п. 4 сообщение об ошибке не выдается, но в поиске участвуют только первые 20 символов.

4. ОБРАБОТКА

4.1. Общее описание

Отбранное множество последовательностей (ОМП) можно тем или иным образом обработать. Система HEID содержит следующие виды обработки последовательности:

- печать;
- просмотр последовательности на дисплее;
- просмотр описания последовательности;
- формирование новой последовательности из фрагментов;
- сравнение двух последовательностей;
- обработка первичной структуры;
- поиск участков;
- поиск рестриктов при совместной рестрикции;
- поиск открытых рамок считывания.

Сеанс работы с программами обработки устроен так, что для каждой из содержащихся в ОМП последовательностей в начале работы с ней и после завершения очередного вида обработки предоставляется возможность выбрать любой из допустимых видов обработки или кончить обработку данной последовательности.

Параметры обработки (вид обработки, длина участков разбиения, количество допустимых несовпадений и т.п.), как правило,

вводятся с дисплея. Если это неудобно (например, при необходимости одинаковым образом обработать много последовательностей), то можно подготовить параметры обработки в некотором файле на диске и указать его имя в ответ на соответствующий вопрос системы HEID. Этот файл можно подготовить либо текстовым редактором независимо от задач системы, либо средствами системы, указав режим записи параметров обработки в файл для последующего использования и имя этого файла.

Выходные данные каждой программы обработки всегда выводятся на дисплей. Параллельно с этим они накапливаются в протоколе текущей обработки. В конце каждого вида обработки у пользователя выясняется, нужно ли сохранить протокол обработки. При положительном ответе текущий протокол дописывается к общему протоколу; при отрицательном – уничтожается. Это свойство системы следует использовать при задании неудачных параметров обработки.

После завершения обработки последней последовательности из ОМП пользователь имеет возможность напечатать общий протокол обработки или не делать этого. В любом случае на диске эта информация остается и может быть использована в дальнейшем (например, при необходимости получения нескольких экземпляров протокола).

В процессе работы некоторых программ обработки получают схемы расположения на последовательности различных участков, сайтов рестрикции и открытых рамок считывания. Эти схемы накапливаются друг под другом в графическом протоколе. Графический протокол также может быть напечатан в конце сеанса.

4.2. Печать последовательности

Программа печати последовательности позволяет пользователю получить последовательность напечатанной таким образом, чтобы легко мог быть найден нуклеотид с нужным номером. Для этого под нуклеотидами печатаются их номера через 10. В процессе диалога выясняется, нужно ли печатать последовательность по 80 или по 100 символов в строке (этот параметр определяется возможностями АЦПУ), а также нужно ли печатать комплементарную нить ДНК.

Информация попадает в текущий протокол обработки и не отображается на дисплей. Последовательность будет напечатана вместе с общим протоколом обработки в конце сеанса.

4.3. Просмотр последовательности на дисплее

Программа просмотра последовательности на дисплее позволяет просматривать последовательность в процессе сеанса обработки между различными другими видами обработки. Последователь-

ность выводится на дисплей по 50 нуклеотидов в строке с пробелом между каждыми десятью нуклеотидами; слева и справа указываются номера первого и последнего нуклеотида в строке.

4.4. Просмотр описания последовательности

Пользователю предоставляется возможность просмотреть как все описание последовательности, так и его отдельные поля.

4.5. Формирование новой последовательности из фрагментов

Пользователю предоставляется возможность для решения ряда теоретических вопросов поработать с последовательностью, реально не существующей, составленной из фрагментов различных существующих в банке последовательностей. При этом в качестве первого фрагмента рассматривается обрабатываемая последовательность; для каждого следующего фрагмента запрашивается идентификатор последовательности и начальная и конечная позиции нуклеотидов.

Сформированная таким образом последовательность существует только во время сеанса; никакой записи в банк при этом не происходит.

4.6. Сравнение двух последовательностей

Программа сравнения двух последовательностей позволяет сравнить две последовательности, из которых первой является обрабатываемая последовательность, а идентификатор и границы второй запрашиваются во время работы программы. Программа печатает две последовательности друг под другом, отмечая несовпадения типа замены. В конце печатается общее число несовпадений типа замены.

Такая программа сравнения бывает полезна тогда, когда заранее известна близость определенных участков двух последовательностей (например, с них считываются родственные белки). Более глубокому изучению этого вопроса посвящена работа /1/.

4.7. Обработка первичной структуры

4.7.1. Общее описание

Система HEID предлагает пользователю следующие виды обработки первичной структуры последовательности:

- подсчет плотности распределения нуклеотидов и пар нуклеотидов;
- построение таблицы ближайших соседей;

– подсчет сочетаний нуклеотидов по два и по три.

После завершения каждого вида обработки пользователю предоставляется возможность снова выбрать любой из этих видов или закончить обработку первичной структуры.

4.7.2. Подсчет плотности распределения нуклеотидов и пар нуклеотидов

Программа подсчета плотности распределения нуклеотидов и пар нуклеотидов разбивает последовательность на участки указанной длины и для каждого участка вычисляет количество нуклеотидов каждого вида, а также суммарное количество нуклеотидов А и Т(У), G и С. Вычисляется также процентное содержание на участке нуклеотидов каждого вида и указанных пар.

4.7.3. Построение таблицы ближайших соседей

Программа построения таблицы ближайших соседей предназначена для изучения локального порядка в последовательности. Для каждого из видов нуклеотидов подсчитывается, сколько раз на 1, 2, 3, 4 местах слева и справа от него встретились нуклеотиды другого вида.

Результаты печатаются в виде четырех матриц, поименованных в соответствии с видами нуклеотидов. В матрице первый индекс соответствует положению по отношению к рассматриваемому нуклеотиду (меняется от -4 до 4), второй – номеру в списке А, Т, G, С.

Пример:

A	A	T	G	C
-4	136	106	98	81
-3	127	131	95	69
-2	137	140	80	65
-1	168	97	76	81
0	422	0	0	0
1	168	113	70	71
2	137	117	87	81
3	127	126	82	87
4	136	130	87	69

Из этой таблицы видно, что нуклеотид "Т" встретился на 4-м слева месте от нуклеотида "А" 106 раз. Видно также, что сочетание ТА (97) меньше, чем АТ (113).

4.7.4. Подсчет сочетаний нуклеотидов по два и по три

Программа подсчета сочетаний по два и по три также предназначена для изучения локального порядка последовательности. Она подсчитывает количество всех пар и троек нуклеотидов, реально встретившихся в последовательности, а также печатает математическое ожидание этих величин в предположении, что нуклеотиды вида i встречаются в последовательности независимо друг от друга с частотой $p(i)$, равной

$$p(i) = N(i) / L,$$

где $N(i)$ - количество нуклеотидов i -того вида в последовательности, L - длина последовательности.

Математическое ожидание пар нуклеотидов i -того и j -того видов $MO(i, j)$ вычисляется при этом по формуле

$$MO(i, j) = p(i) \times p(j) \times L,$$

троек нуклеотидов вида i, j, k -

$$MO(i, j, k) = p(i) \times p(j) \times p(k) \times L.$$

4.8. Поиск участков на последовательности

4.8.1. Общее описание

Система HEID предлагает пользователю следующие виды поиска участков:

- поиск участков данного вида (при этом допускаются несовпадения только типа замены);
- поиск участков со слабой гомологией (допускаются несовпадения типа замены, вставки, выпадения);
- поиск прямых повторов;
- поиск инвертированных повторов;
- поиск участков двойной симметрии ("шпилек").

После завершения каждого вида поиска пользователю предоставляется возможность снова выбрать любой из указанных видов поиска участков или закончить поиск участков.

Общей для всех видов поиска является возможность прервать выполнение программы после нахождения заданного пользователем числа участков.

4.8.2. Поиск участков данного вида

Программа поиска участков данного вида находит и печатает (точнее, выводит на дисплей и в протокол) все номера позиций в

последовательности, с которых начинается указанный пользователем участок с точностью до заданного числа несовпадений типа замены. В графический протокол выводится схема расположения этих участков на последовательности.

Если участок не является "шпилькой", то поиск ведется также и на комплементарной нити ДНК.

В конце протокола печатается общее число найденных участков на каждой нити.

4.8.3. Поиск участков со слабой гомологией

Программа поиска участков со слабой гомологией находит на последовательности все участки, которые имеют по сравнению с указанным заданное число несовпадений типа замены, вставки или выпадения. Параметром программы является число замен, эквивалентное одной вставке или выпадению. Найденные участки печатаются вместе с указанным в 4 строки: во 2-й и 3-й – совпавшие нуклеотиды, в 1-й и 4-й – несовпавшие.

Для каждого найденного участка печатается число несовпадений, из них – типа замены. В конце печатается общее число найденных участков.

4.8.4. Поиск прямых повторов

Прямым повтором длины N с промежутком длины L называется участок последовательности длиной $2N+L$, где N крайних нуклеотидов в точности повторяют друг друга, а между последним нуклеотидом левой симметричной части и первым нуклеотидом правой части находятся L неважно каких нуклеотидов. Например, участок `ATGCTTTATGC` является прямым повтором длины 4 с промежутком длины 3. По аналогии с задачей о "шпильках" промежуток называется также "головой".

Программа поиска прямых повторов находит на последовательности все прямые повторы, удовлетворяющие ограничениям на минимальную длину повтора и минимальную и максимальную длину промежутка и имеющие не более заданного числа несовпадений типа замены в симметричных частях. На дисплей и в протокол выводятся начальные позиции симметричных частей, длина промежутка, сами симметричные части; в конце сообщается общее число найденных повторов и максимальная длина.

4.8.5. Поиск инвертированных повторов

Инвертированным повтором длины N с промежутком длины L называется участок последовательности длиной $2N+L$, где в симметричных частях длиной N на одинаковых от концов позициях находятся одинаковые нуклеотиды. Например, участок

ATGCTTTCGTA является инвертированным повтором длины 4 с промежутком длины 3.

Программа имеет те же параметры и выдает результаты в таком же виде, как программа поиска прямых повторов.

4.8.6. Поиск участков двойной симметрии ("шпильки")

Участком двойной симметрии ("шпилькой") длины N с промежутком ("головкой") длины L называется участок последовательности длиной $2N + L$, где в симметричных частях длиной N на одинаковых от концов позициях находятся комплементарные нуклеотиды. Например, участок ATGCTTTCGAT является "шпилькой" длины 4 с головкой длины 3.

Программа имеет те же параметры и выдает результаты в таком же виде, как программа поиска прямых повторов.

4.9. Поиски рестриктов при совместной рестрикции

Если на последовательность подействовать ферментом (рестриктазой), он будет находить на последовательности определенные участки (сайты рестрикции; для каждой рестриктазы они свои) и расщеплять последовательность в определенных местах этих участков на фрагменты (рестрикты).

Программа рестрикции позволяет получить ответ на вопрос: какие рестрикты мы получим, действуя на данную последовательность определенным набором рестриктаз, при условии полного расщепления. В графический протокол выводятся как схемы расщепления каждой рестриктазой, так и схема совместной рестрикции; на дисплей и в протокол для каждого рестрикта, полученного в результате с о в м е с т н о й рестрикции, выводятся:

- номер начальной позиции;
- номер конечной позиции;
- длина рестрикта в нуклеотидах;
- какими рестриктазами образован;
- молекулярный вес рестрикта;
- длина рестрикта в ангстремах;
- процент содержания GC-пар.

При этом можно использовать как стандартные, широко известные рестриктазы, так и новые, нестандартные. В первом случае достаточно указать название рестриктазы; во втором случае необходимо указать также сайт рестрикции. Для реализации первой возможности в файле RESLST.DAT содержится список стандартных рестриктаз с их сайтами рестрикции, соответствующий документу /2/.

Система предоставляет определенный сервис при работе со списком стандартных рестриктаз. Пользователь имеет возможность - посмотреть весь список;

- выяснить сайт заданной рестриктазы;
- выяснить, какие рестриктазы обладают заданным сайтом рестрикции.

Пользователь, который регулярно работает с определенным набором рестриктаз, имеет возможность записать их в некоторый файл на диске (завести личный набор рестриктаз) и в дальнейшей работе указывать не названия рестриктаз, а имя соответствующего файла.

При вводе названий рестриктаз и сайтов рестрикции нужно следовать определенным правилам. Эти правила могут быть сообщены пользователю по его запросу.

При подсчете молекулярного веса рестриктов рассматриваются 3 варианта:

- однонитевая РНК;
- однонитевая ДНК;
- двунитевая ДНК.

Решение о том, является ли рассматриваемая последовательность РНК или ДНК, принимается автоматически на основании первичной структуры последовательности (последовательность считается РНК, если в ней встречается нуклеотид U). Является ли молекула ДНК двунитевой, выясняется у пользователя.

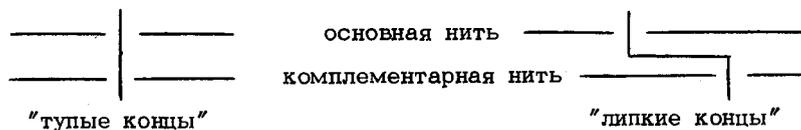
При подсчете используются следующие значения молекулярных весов нуклеотидов:

	A	T(U)	G	C
в ДНК	313	302	328	288
в РНК	329	305	344	304

Длина нуклеотида полагается равной $2,4 \text{ \AA}$. (Все числовые значения взяты из [3]).

Длина рестрикта (в ангстремах) вычисляется как сумма длин нуклеотидов. Молекулярный вес однонитевой молекулы вычисляется как сумма весов нуклеотидов. Молекулярный вес рестрикта в двунитевой молекуле вычисляется как сумма весов основной и комплементарной нитей рестрикта.

При расщеплении двунитевой молекулы ДНК возможны 2 варианта: "тупые концы", если позиции расщепления на обеих нитях совпадают, и "липкие концы" в противном случае:



В данной версии при подсчете молекулярных весов и длин рестриктов всегда предполагается первая возможность ("тупые концы").

4.10. Поиск открытых рамок считывания

С последовательности ДНК могут считываться по определенным правилам белки. Известны необходимые (но не достаточные) условия для этого. Участок последовательности, откуда может быть считан белок, называется открытой рамкой считывания (ORF - Open Reading Frame). Некоторые свойства считываемых белков могут быть определены из последовательности ДНК. Процесс построения белковой последовательности по известной нуклеотидной называется трансляцией.

Программа поиска открытых рамок считывания (называемая также программой трансляции) ищет на последовательности все открытые рамки. Для каждой найденной ORF, с которой считывается белок длины не менее заданной, выводится:

- молекулярный вес белка;
- заряд белка при заданном pH (по запросу);
- изоэлектрическая точка белка (по запросу);
- последовательность аминокислот;
- аминокислотный состав (по запросу);
- таблица встречаемости кодонов (по запросу).

В программе предполагается следующий механизм считывания белка:

1. Если последовательность задана в виде РНК, она считается матричной РНК. Если она задана в виде ДНК, то она считается матричной РНК с заменой U на T. В любом случае последовательность считается расположенной от 5' - к 3' -концу. (Эти предположения соответствуют способу представления последовательности в банке).

2. Считывание белка может начаться со специального триплетта - кодона инициации - и обязано закончиться другим специальным триплетом - кодоном терминации.

Стандартные инициаторы - AUG и GUG, терминаторы - UAG, UAA, UGA. В некоторых особых случаях роль инициаторов и терминаторов играют другие триплеты. Пользователю предоставляется возможность работать со стандартными, либо сформировать нестандартные наборы инициаторов и терминаторов.

3. С иницирующего кодона считывается аминокислота, которая является первой в последовательности аминокислот белка. Каждый следующий триплет кодирует некоторую аминокислоту; существует таблица соответствия. Терминирующий кодон прерывает этот процесс; с него аминокислота не считывается. Если до конца последовательности терминирующий кодон не встретился, процесс заканчивается по концу последовательности.

С инициатора всегда считывается фермилметионин, независимо от того, какая аминокислота считывается с этого кодона в середине кодирующей последовательности.

4. Очередным триплетом в открытой рамке считывания может

оказаться инициатор. Участок от этого инициатора до ближайшего терминатора также является открытой рамкой считывания; она является вложенной по отношению к первой. Глубина вложенности не ограничена. Ближайший встретившийся терминатор закрывает все ORF.

5. Все триплеты последовательности разбиваются на 3 непересекающиеся группы в соответствии с номером позиции в последовательности их первого нуклеотида, взятым по модулю 3. Соответственно говорят о трех рамках считывания. Процессы считывания белка внутри каждой рамки считывания идут независимо друг от друга.

Изоэлектрическая точка и заряд белка вычисляются следующим образом:

Изоэлектрическая точка вычисляется как корень функции заряда белка, убывающей по pH, с точностью до 0,02. Заряд при заданном pH вычисляется как сумма зарядов аминокислот, его составляющих, и еще двух слагаемых, отвечающих началу и концу белка (то есть сумма заряда белка длиной L аминокислот состоит из L + 2 слагаемых).

Большинство аминокислот дают нулевой вклад в сумму. Ненулевые вклады дают следующие аминокислоты (в скобках приводится однобуквенный код):

а) Arg (P), Lys (K), His (H), начало (первая аминокислота);

б) Glu (E), Asp (D), Tyr (Y), Cys (C), конец (последняя аминокислота).

Слагаемое типа а) вычисляется по следующей формуле:

$$1 - \frac{10^{\text{pH}-F}}{1 + 10^{\text{pH}-F}}$$

типа б) - по формуле:

$$-\frac{10^{\text{pH}-F}}{1 + 10^{\text{pH}-F}}$$

Здесь pH - задаваемое пользователем значение водородного показателя; имеет смысл рассматривать его в пределах от 2 до 12; F - константа, своя для каждой аминокислоты:

Аминокислота	Arg (P)	Lys (K)	His (H)	Glu (E)	Asp (D)	Tyr (Y)	Cys (C)	Нач.	Кон.
F	12,5	10,5	6,8	4,5	4,5	8,5	10,0	7,7	3,6

(значения F взяты из /4/).

5. ИНСТРУКЦИЯ ПО ЭКСПЛУАТАЦИИ

5.1. Вид и восстановление передаваемой информации

Система HEID передается на магнитной ленте, содержащей копии двух дисков - HEIDOC и HEID2. На диске HEIDOC содержатся исходные тексты и документация системы. На диске HEID2 содержится некоторый вариант банка, задачи системы и файлы вспомогательных данных. Диск HEIDOC не является обязательным для эксплуатации системы.

Восстановление диска HEIDOC:

```
>BRU /INI/REW/VAC:HEIDOC MT:DKN:
```

Восстановление диска HEID2:

```
>BRU/INI/REW/VAC:HEID2 MT:DKN:
```

После восстановления на диске HEIDOC (UIC 150, 2) будут содержаться:

1) 7 файлов документации:

- KONTPRIM.DPR - контрольный пример;
- OPPROG.DPR - описание программы;
- TXTPROG.DPR - тексты программ;
- OPPRIM.DPP - описание применения;
- RUKPROG.DPR - руководство программиста;
- RUKSYSPR.DPR - руководство системного программиста;
- RUKOPER.DPR - руководство оператора;

2) 66 файлов исходных текстов (.FTN);

3) 2 файла описания оверлеев (.ODL);

4) 4 командных файла строителя задач (.CMD);

5) командный файл генерации системы (GENER.COM).

После восстановления на диске HEID2 (UIC 150, 1) будут содержаться:

1) 566 файлов с нуклеотидными последовательностями (.BNK);

2) 4 задачи системы HEID:

- IDENT.TSK;
- FROM.TSK;
- POISK.TSK;
- OBR.TSK;

3) 3 файла данных, необходимых для нормальной эксплуатации системы:

- IDENTALL.DAT - список всех идентификаторов;
- RESLST.DAT - список стандартных рестриктаз;
- LOCATION.DAT - положение банка (устройство, UIC);

4) 2 файла данных, необходимых для контрольного примера (.TST).

5.2. Как работать с последовательностями, содержащимися в банке

После восстановления информации с ленты на магнитный диск с системой можно работать как со своим кодом идентификации пользователя (UIC), так и с кодом 150, 1, используя устройство, на котором установлен диск HEID2. Работа со своим кодом является более предпочтительной, так как при работе создаются файлы с протоколами поиска и обработки, хранение или уничтожение которых должно определяться пользователем. Поэтому при работе с банком необходима информация об устройстве и коде идентификации. Эта информация должна быть указана в файле LOCATION.DAT, принадлежащем пользователю. Формат информации - 13 символов вида

DK3: [150,001].

Запуск программы поиска при работе со своим кодом:

> RUN DKN:[150, 1] POISK.

Запуск программ обработки при работе со своим кодом:

> RUN DKN:[150, 1] OBR.

Работая с кодом 150, 1 и используя устройство, на котором установлен диск HEID2, можно запускать эти программы

и > RUN POISK

> RUN OBR.

Файл LOCATION.DAT тем не менее является необходимым.

5.3. Как работать с собственными последовательностями

Чтобы работать с собственной последовательностью, нужно записать ее на диск в формате Гейдельбергского банка. При этом необходимы поля ID и SQ. В поле ID с 6-й позиции (через 3 пробела после D) должен начинаться 6- или 8-символьный идентификатор последовательности. В поле SQ с 16-й позиции (через 13 пробелов после Q) должна находиться 5-значная длина последовательности. Имя файла должно совпадать с идентификатором последовательности; расширение должно быть .BNK.

Чтобы последовательность наряду с другими просматривалась программами поиска, нужно включить ее идентификатор в список, хранящийся в файле IDENTALL.DAT.

5.4. Как перейти к новой версии банка

Если новая версия банка поступает в виде последовательного файла, его нужно "разрезать" на отдельные файлы по одной последовательности в файле; каждому файлу нужно присвоить имя, совпадающее с идентификатором последовательности, и расширение .BNK. Это можно сделать при помощи программы FROM. Входные данные для программы FROM:

- банк в виде последовательного файла;
- список всех идентификаторов.

Список всех идентификаторов, имеющихся в последовательном файле в формате Гейдельбергского банка, можно получить с помощью программы IDENT. Имена файлов, с которыми работают программы IDENT и FROM, запрашиваются в диалоговом режиме.

"Разрезание" банка на отдельные файлы - процедура долгая. При снятии программы (например, из-за сбоя на диске), чтобы не повторять уже сделанную часть работы, следует откорректировать список идентификаторов, удалив из него идентификаторы обработанных последовательностей.

ЛИТЕРАТУРА

1. Ройтберг М.А. Алгоритм определения гомологии первичных структур. Препринт ОНТИ НЦБИ АН СССР, 1984.
2. Roberts R.J. NUCLEASES. Cold Spring Harbor Laboratory. Appendix A. Restriction Endonucleases. 1982, p. 311-340.
3. Уотсон Дж. Молекулярная биология гена. М., Мир, 1978.
4. Мецлер Д. Биохимия. М., Мир, 1980.

ПРИЛОЖЕНИЕ 1

СЕАНС ПОИСКА

РАБОТАЕТ ПРОГРАММА ПОИСКА СИСТЕМЫ NEID
 ПОИСК ВЕДЕТСЯ ПО ВСЕМ БАНКЕ ИЛИ ЕГО ПОДМНОЖЕСТВЕ
 ПОДМНОЖЕСТВО ЗАДАЕТСЯ СПИСОКОМ ИДЕНТИФИКАТОРОВ ПОСЛЕДОВАТЕЛЬНОСТЕЙ,
 СРЕДИ КОТОРЫХ НУЖНО ВЕСТИ ПОИСК
 СПИСОК ДОЛЖЕН БЫТЬ ПОДГОТОВЛЕН В НЕКОТОРОМ ФАЙЛЕ НА ДИСКЕ
 ПОИСК БУДЕТ ВЕСТИСЬ ВО ВСЕМ БАНКЕ? (ДА/НЕТ)
 НЕТ
 УКАЖИТЕ ИМЯ ФАЙЛА СО СПИСОКОМ ИДЕНТИФИКАТОРОВ
 DK5 (15,1) ID1.TST
 УКАЖИТЕ ИМЯ ФАЙЛА ДЛЯ ВЫХОДНОГО СПИСКА
 DELETE.UAT
 ПОИСКОВЫЙ ЗАПРОС БУДЕТ ВВОДИТЬСЯ С ДИСПЛЕЯ? (ДА/НЕТ)
 ДА
 НУЖНА ИНСТРУКЦИЯ ПО ВВОДУ ЗАПРОСА? (ДА/НЕТ)
 ДА
 ПОИСКОВЫЙ ЗАПРОС ПРЕДСТАВЛЯЕТ СОБОЙ ЛОГИЧЕСКУЮ ФУНКЦИЮ ТРЕБОВАНИЙ
 К ОТДЕЛЬНЫМ ПОЛЯМ ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ.
 ТРЕБОВАНИЕ К ОТДЕЛЬНОМУ ПОЛЮ - ЭТО ВЫРАЖЕНИЕ ВИДА
 (DE#TRNA)
 или
 (DE#TRNA),
 где "##" УПОТРЕБЛЯЕТСЯ В СМЫСЛЕ "СОДЕРЖИТ",
 "##" "НЕ СОДЕРЖИТ"
 В КАЧЕСТВЕ ИМЕНИ ПОЛЯ ДОПУСКАЮТСЯ ВСЕ СУЩЕСТВУЮЩИЕ В БАНКЕ
 ПОЛЯ ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ:
 ID,UT,DE,US,OC,KA,KN,NA,CC,FT,SA
 В ЗАПРОСЕ МОГУТ ПРИСУТСТВОВАТЬ 2 ТРЕБОВАНИЯ К ОДНОМУ ПОЛЮ, НАПРИМЕР:
 (DE#KNA)или(DE#TRNA) -
 НАЙТИ ВСЕ ПОСЛЕДОВАТЕЛЬНОСТИ, В ПОЛЕ ОПИСАНИЯ ДЕ КОТОРЫХ
 СОДЕРЖИТСЯ ИМЯ ПОЛЯ TRNA 19
 ПРИ СОСТАВЛЕНИИ ЗАПРОСА, СПОЛЗУЮТСЯ КВАДРАТНЫЕ СКОБКИ ПО СЛЕДУЮЩИМ ПРАВИЛАМ:
 1. ЧАСТЬ ЗАПРОСА, СОДЕРЖАЩАЯ ЛОГИЧЕСКИЕ ФУНКЦИИ ОДНОГО ТИПА,
 ЗАКЛЮЧАЕТСЯ В КВАДРАТНЫЕ СКОБКИ.
 2. МЕЖДУ КВАДРАТНЫМИ СКОБКАМИ ДОПУСКАЮТСЯ ФУНКЦИИ ТОЛЬКО ОДНОГО
 ТИПА, ОНЫ МОГУТ КАК СОВПАДАТЬ С ФУНКЦИЕЙ ВНУТРИ СКОБКИ, ТАК И
 ОТЛИЧАТЬСЯ ОТ НЕЙ.
 3. ВНУТРИ РАЗНЫХ КВАДРАТНЫХ СКОБКИ МОГУТ УПОТРЕБИТЬСЯ РАЗНЫЕ
 ЛОГИЧЕСКИЕ ФУНКЦИИ.
 ПРИМЕРЫ ПРАВИЛЬНО СОСТАВЛЕННЫХ ЗАПРОСОВ: ((IUBAAA))
 ((DE#TTT)и(DE#RRR))
 ((K#BUNA)или(DE#D#A)), ((R#V#)или(ID#A))
 ((IUBA)или(ID#B)) или ((DE#A)-L (DE#B)) или ((R#FT) (OC#BVB))
 ОГРАНИЧЕНИЯ:
 1. ОБЩАЯ ДЛИНА ЗАПРОСА - НЕ БОЛЕЕ 4 СТРОК ЗАРАНА
 2. ДОПУСКАЕТСЯ НЕ БОЛЕЕ 5 КВАДРАТНЫХ СКОБКИ
 3. ВНУТРИ СКОБКИ ДОПУСКАЕТСЯ НЕ БОЛЕЕ 6 ТРЕБОВАНИЙ К ОТДЕЛЬНЫМ ПОЛЯМ
 ПРИ НАРУШЕНИИ П.1-3 ВЫДАЕТСЯ СООБЩЕНИЕ ОБ ОШИБКЕ, ЗАПРОС ИГНОРИРУЕТСЯ
 4. ТЕКСТ, КОТОРЫЙ ИЩЕТСЯ В УКАЗАННОМ ПОЛЕ, ДОЛЖЕН БЫТЬ НЕ ДЛИННЕЕ 20 СИМВОЛОВ
 ПРИ НАРУШЕНИИ П.4 СООБЩЕНИЕ ОБ ОШИБКЕ НЕ ВЫДАЕТСЯ,
 НО В ПОИСКЕ УЧАСТВУЮТ ТОЛЬКО 20 ПЕРВЫХ СИМВОЛОВ УКАЗАННОМУ ТЕКСТА
 ВВЕДИТЕ ОЧЕРЕДНУЮ СТРОКУ ЗАПРОСА
 ((DE#KNA))
 ПОСЛЕДОВАТЕЛЬНОСТЬ SCRNA5 УДОВЛЕТВОРЯЕТ ЗАПРОСУ
 ПОСЛЕДОВАТЕЛЬНОСТЬ SCRNA6 УДОВЛЕТВОРЯЕТ ЗАПРОСУ
 НАЙДЕНО 2 ПОСЛЕДОВАТЕЛЬНОСТИ, УДОВЛЕТВОРЯЮЩИХ ЗАПРОСУ.
 ВЫ ПРОДОЛЖИТЕ ПОИСК? (ДА/НЕТ)
 ДА
 ПОИСК БУДЕТ ВЕСТИСЬ ВО ВСЕМ БАНКЕ? (ДА/НЕТ)
 НЕТ
 УКАЖИТЕ ИМЯ ФАЙЛА СО СПИСОКОМ ИДЕНТИФИКАТОРОВ
 DK5 (15,1) ID1.TST
 УКАЖИТЕ ИМЯ ФАЙЛА ДЛЯ ВЫХОДНОГО СПИСКА
 DELETE.UAT
 ПОИСКОВЫЙ ЗАПРОС БУДЕТ ВВОДИТЬСЯ С ДИСПЛЕЯ? (ДА/НЕТ)

ДА
НУЖНА ИНСТРУКЦИИ ПО ВВОДУ ЗАПРОСА? (ДА/НЕТ)
НЕТ
ВВЕДИТЕ ОЧЕРЕДНУЮ СТРОКУ ЗАПРОСА
[(DE#R#NA)]
ПОСЛЕДОВАТЕЛЬНОСТЬ AD2895 УДОВЛЕТВОРЯЕТ ЗАПРОСУ
ПОСЛЕДОВАТЕЛЬНОСТЬ ЕС#ЕСА УДОВЛЕТВОРЯЕТ ЗАПРОСУ
ПОСЛЕДОВАТЕЛЬНОСТЬ #M#ЕСА УДОВЛЕТВОРЯЕТ ЗАПРОСУ
ПОСЛЕДОВАТЕЛЬНОСТЬ Х#Н#Е#А УДОВЛЕТВОРЯЕТ ЗАПРОСУ
НАЙДЕНО 4 ПОСЛЕДОВАТЕЛЬНОСТИ, УДОВЛЕТВОРЯЮЩИХ ЗАПРОСУ.
ВЫ ПРОДОЛЖИТЕ ПОИСК? (ДА/НЕТ)
ДА
ПОИСК БУДЕТ ВЕСТИТЬСЯ ВО ВСЕМ БАНКЕ? (ДА/НЕТ)
НЕТ
УКАЖИТЕ ИМЯ ФАЙЛА СО СПИСКОМ ИДЕНТИФИКАТОРОВ
DK5:(150,1)ID1.TST
УКАЖИТЕ ИМЯ ФАЙЛА ДЛЯ ВЫХОДНОГО СПИСКА
DELETE.DAT
ПОИСКОВЫЙ ЗАПРОС БУДЕТ ВВОДИТЬСЯ С ДИСПЛЕЯ? (ДА/НЕТ)
ДА
НУЖНА ИНСТРУКЦИИ ПО ВВОДУ ЗАПРОСА? (ДА/НЕТ)
НЕТ
ВВЕДИТЕ ОЧЕРЕДНУЮ СТРОКУ ЗАПРОСА
[(DE#R#NA)]ИЛИ[(DE#R#NA)]
ПОСЛЕДОВАТЕЛЬНОСТЬ AD2895 УДОВЛЕТВОРЯЕТ ЗАПРОСУ
ПОСЛЕДОВАТЕЛЬНОСТЬ ЕС#ЕСА УДОВЛЕТВОРЯЕТ ЗАПРОСУ
ПОСЛЕДОВАТЕЛЬНОСТЬ #M#ЕСА УДОВЛЕТВОРЯЕТ ЗАПРОСУ
ПОСЛЕДОВАТЕЛЬНОСТЬ SC#N#A#S УДОВЛЕТВОРЯЕТ ЗАПРОСУ
ПОСЛЕДОВАТЕЛЬНОСТЬ SC#R#N#A#6 УДОВЛЕТВОРЯЕТ ЗАПРОСУ
ПОСЛЕДОВАТЕЛЬНОСТЬ Х#Н#Е#А УДОВЛЕТВОРЯЕТ ЗАПРОСУ
НАЙДЕНО 6 ПОСЛЕДОВАТЕЛЬНОСТЕЙ, УДОВЛЕТВОРЯЮЩИХ ЗАПРОСУ.
ВЫ ПРОДОЛЖИТЕ ПОИСК? (ДА/НЕТ)
ДА
ПОИСК БУДЕТ ВЕСТИТЬСЯ ВО ВСЕМ БАНКЕ? (ДА/НЕТ)
НЕТ
УКАЖИТЕ ИМЯ ФАЙЛА СО СПИСКОМ ИДЕНТИФИКАТОРОВ
DK5:(150,1)ID1.TST
УКАЖИТЕ ИМЯ ФАЙЛА ДЛЯ ВЫХОДНОГО СПИСКА
DELETE.DAT
ПОИСКОВЫЙ ЗАПРОС БУДЕТ ВВОДИТЬСЯ С ДИСПЛЕЯ? (ДА/НЕТ)
ДА
НУЖНА ИНСТРУКЦИИ ПО ВВОДУ ЗАПРОСА? (ДА/НЕТ)
НЕТ
ВВЕДИТЕ ОЧЕРЕДНУЮ СТРОКУ ЗАПРОСА
[(DE#R#NA)]ИЛИ[(DE#R#NA)]
НАЙДЕНО 6 ПОСЛЕДОВАТЕЛЬНОСТЕЙ, УДОВЛЕТВОРЯЮЩИХ ЗАПРОСУ.
ВЫ ПРОДОЛЖИТЕ ПОИСК? (ДА/НЕТ)
НЕТ
ПРОТОКОЛ ПОИСКА НАПЕЧАТАТЬ? (ДА/НЕТ)
ДА
ПРОГРАММА ПОИСКА СИСТЕМЫ #E#I#D КОНЧИЛА РАБОТУ

ПРИЛОЖЕНИЕ 2

СЕАНС ОБРАБОТКИ

```

РАБОТАЮТ ПРОГРАММЫ ОБРАБОТКИ СИСТЕМЫ NEIO
ИДЕНТИФИКАТОРЫ ПОСЛЕДОВАТЕЛЬНОСТЕЙ БУДУТ ВВОДИТЬСЯ С ДИСПЛЕЯ? (ДА/НЕТ)
ДА
ПАРАМЕТРЫ ОБРАБОТКИ БУДУТ ВВОДИТЬСЯ С ДИСПЛЕЯ? (ДА/НЕТ)
ДА
НУЖНО ЛИ ЗАПИСЫВАТЬ ПАРАМЕТРЫ ОБРАБОТКИ В ФАЙЛ
  ДЛЯ ПОСЛЕДУЮЩЕГО ИСПОЛЬЗОВАНИЯ? (ДА/НЕТ)
НЕТ
УКАЖИТЕ ИДЕНТИФИКАТОР ПОСЛЕДОВАТЕЛЬНОСТИ ИЛИ ***
RNRESA
РАБОТАЕМ С ПОСЛЕДОВАТЕЛЬНОСТЬЮ
ID RNRESA P,CHIRABILIS'RESA BENE
ДЛИНА ПОСЛЕДОВАТЕЛЬНОСТИ 1329 В.Р.
ВЫ БУДЕТЕ РАБОТАТЬ СО ВСЕЙ ПОСЛЕДОВАТЕЛЬНОСТЬЮ? (ДА/НЕТ)
ДА
ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТИ:
ПЕЧАТЬ      - ПЕЧАТЬ ПОСЛЕДОВАТЕЛЬНОСТИ
ДИСПЛЕЙ    - ПРОСМОТР ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
ОПИСАНИЕ   - ПРОСМОТР ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ
ФОРМИРОВАНИЕ - ФОРМИРОВАНИЕ НОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗ ФРАГМЕНТОВ
ГОМОЛОГИЯ  - СРАВНЕНИЕ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
СТРУКТУРА  - ОБРАБОТКА ПЕРВИЧНОЙ СТРУКТУРЫ
УЧАСТКИ    - ПОИСК УЧАСТКОВ
РЕСТРИКЦИЯ - ПРОГРАММА РЕСТРИКЦИИ
ТРАНСЛЯЦИЯ - ПРОГРАММА ТРАНСЛЯЦИИ
КОНЕЦ      - КОНЧИТЬ ОБРАБОТКУ ПОСЛЕДОВАТЕЛЬНОСТИ
НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
П
ВЫ ВЫБРАЛИ ПЕЧАТЬ
РАБОТАЕТ ПРОГРАММА ПЕЧАТИ ПОСЛЕДОВАТЕЛЬНОСТИ
ПЕЧАТАТЬ ЛИ КОМПЛЕМЕНТАРНУЮ НИТЬ? (ДА/НЕТ)
ДА
ПРОГРАММА МОЖЕТ ПЕЧАТАТЬ ПО 100 ИЛИ ПО 80 СИМВОЛОВ В СТРОКЕ
ВАМ НУЖНО ПЕЧАТАТЬ ПО 100 СИМВОЛОВ? (ДА/НЕТ)
НЕТ
ПРОГРАММА ПЕЧАТИ ПОСЛЕДОВАТЕЛЬНОСТИ КОНЧИЛА РАБОТУ
ПРОТОКОЛ ОБРАБОТКИ СОХРАНИТЬ? (ДА/НЕТ)
ДА
ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТИ:
ПЕЧАТЬ      - ПЕЧАТЬ ПОСЛЕДОВАТЕЛЬНОСТИ
ДИСПЛЕЙ    - ПРОСМОТР ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
ОПИСАНИЕ   - ПРОСМОТР ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ
ФОРМИРОВАНИЕ - ФОРМИРОВАНИЕ НОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗ ФРАГМЕНТОВ
ГОМОЛОГИЯ  - СРАВНЕНИЕ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
СТРУКТУРА  - ОБРАБОТКА ПЕРВИЧНОЙ СТРУКТУРЫ
УЧАСТКИ    - ПОИСК УЧАСТКОВ
РЕСТРИКЦИЯ - ПРОГРАММА РЕСТРИКЦИИ
ТРАНСЛЯЦИЯ - ПРОГРАММА ТРАНСЛЯЦИИ
КОНЕЦ      - КОНЧИТЬ ОБРАБОТКУ ПОСЛЕДОВАТЕЛЬНОСТИ
НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
Д
ВЫ ВЫБРАЛИ ДИСПЛЕЙ
  1  CTTAAAGAAA  TTATTAATAA  TTAAC1TGAT  ACTGTATGAT  TATACAGTAT   50
 51  AATGAGTTTC  AACAAGCAAA  ATCATATAGC  TTTTAATGGT  AGTGACCCAT  100
101  CTTTATGGTT  CACTGCCSAG  AGGAGATAAA  CATGGCIATT  GATGAAAAAC  150
151  AACAATAAGC  ATTGCCGCA  GCACTTGGTC  AAATTGAAAA  GCAATTTGGT  200
201  AAAGGTTCTA  TCATGCTCT  GGGCGAAGAC  CGTTCCATGA  ACBTAGAAAAC  250
251  TATCTCTACA  GGATCTTAT  CATTAGACGT  TGCTTTAGGT  GCAAGTGGAT  300
301  TGCCACG1GG  CCGTAT1GT  GAAATCTATG  GCCCTGAATC  T1CTGGTAAA  350
351  ACAACCTTGA  CTTCTACAAT  TAT1GCTCT  GCTCAGCGTG  AAGGAAAAAT  400
401  TTGTGCA1TT  ATTGATGCTG  AACATGCA1T  AGACCCAATT  TATGCTCAAA  450
451  AGCTAGGTGT  CGATATCGAC  AACTTACTTT  GCTCTCAACC  TGACACAGGT  500

```

```

501 GAACAAGCTC TGGAAATTTG TGATGCATTA TCCTCTCTG GTCGGTTGA 550
551 TGTATTGTC GTGGACTCCG TGGCAGCATT AACACCAAAA GCTGAAATG 600
601 AAGGTGAAAT TGGTGAITCA CACTTGGTT TACCCGCACG TATGATGAGC 650
651 CAAGCTATGC GTAACATAGC GGGTAACSTT AAAAACTTA ATACACTGCT 700
701 GATTTTCATT AACCAAAATC GTATGAAAAT CGTGTGTAIG TTTGGTAACC 750
751 CAGAAACCCAC GACCGGTGGT AATGCBCTTA AATCTATGC TTCTGTCTGT 800
801 TTAGACATTC GTCCGATTCG CTCTGTCAAA AATGGTATG AAGTCATTGG 850
851 TAGTGAGACT CGCTTAAAG TTGTAAAAA TAAAGTGGCT GCACCGTTTA 900
901 AACAGCTGA ATTCCAAAT ATGTACGGTG AAGGTATTA TACCTATGGC 950
951 GAACGTATTG ATTTAGGTTT TAAACATAA TTAGTAGAGA AAGCAGGTGC 1000
1001 TTGTATAGC TACAATGCCG AAAAAATGG TCAAGGTAAG GCTAACGCAA 1050
1051 CCAATTACTT AAAAACAACAT CCTAAATGT ACAATGAGT AAACASTAAA 1100
1101 TTCCGTGAAA TGTGTTAAA TCATGCTGGT GAATTCACAA GTGCTGCGGA 1150
1151 TTTTGCAGGT GAAGAGTCAAG ACAGTATGC TGACACACAA AAAAGATTA 1200
1201 TAGCTGGTTC TCATGCTGTT TGTGTAATA TACACCTTA ATCATGGCT 1250
1251 ATTATCAGCA CAGCATCCCA TAGAATAACT TGT

```

ПРОТОКОЛ ОБРАБОТКИ. СОХРАНИТЬ? (ДА/НЕТ)

ДА

ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТИ

- ПЕЧАТЬ - ПЕЧАТЬ ПОСЛЕДОВАТЕЛЬНОСТИ
- ДИСПЛЕИ - ПРОСМОТР ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
- ОПИСАНИЕ - ПРОСМОТР ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ
- ФОРМИРОВАНИЕ - ФОРМИРОВАНИЕ НОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ИЗ ФРАГМЕНТОВ
- ГОМОЛОГИЯ - СРАВНЕНИЕ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
- СТРУКТУРА - ОБРАБОТКА ПЕРВИЧНОЙ СТРУКТУРЫ
- УЧАСТКИ - ПОИСК УЧАСТКОВ
- РЕСТРИКЦИЯ - ПРОГРАММА РЕСТРИКЦИИ
- ТРАНСЛЯЦИЯ - ПРОГРАММА ТРАНСЛЯЦИИ
- КОНЕЦ - КОНЧИТЬ ОБРАБОТКУ ПОСЛЕДОВАТЕЛЬНОСТИ

НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА

0

ВЫ ВЫБРАЛИ ОПИСАНИЕ

РАБОТАЕТ ПРОГРАММА ПРОСМОТРА ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ

ПРОСМОТРЕТЬ МОЖНО ВСЕ ОПИСАНИЕ ИЛИ УКАЗАВНЕ ПОЛЯ

ПРОГРАММА ЗАКАНЧИВАЕТ РАБОТУ, КУДА В КАЧЕСТВЕ ИМЕНИ ПОЛЯ УКАЗЫВАЕТСЯ ***

УКАЖИТЕ ИМЯ ПОЛЯ, СЛОВО "ВСЕ" ИЛИ ***

```

ВСЕ
ID  PNRCSA      P.MIRABILIS RECA GENE
DT  18-APR-84
DE  P.MIRABILIS RECA GENE
KW  DNA BINDING PROTEIN
SU  1329 B.P., 422 A; 394 T; 284 G; 229 C

```

УКАЖИТЕ ИМЯ ПОЛЯ, СЛОВО "ВСЕ" ИЛИ ***

```

DE
DE  P.MIRABILIS RECA GENE

```

УКАЖИТЕ ИМЯ ПОЛЯ, СЛОВО "ВСЕ" ИЛИ ***

```

PP

```

В ОПИСАНИИ ПОСЛЕДОВАТЕЛЬНОСТИ НЕ СОДЕРЖИТСЯ ПОЛЯ PP

УКАЖИТЕ ИМЯ ПОЛЯ, СЛОВО "ВСЕ" ИЛИ ***

ПРОГРАММА ПРОСМОТРА ОПИСАНИЯ КОНЧИЛА РАБОТУ

ПРОТОКОЛ ОБРАБОТКИ СОХРАНИТЬ? (ДА/НЕТ)

ДА

ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТИ

- ПЕЧАТЬ - ПЕЧАТЬ ПОСЛЕДОВАТЕЛЬНОСТИ
- ДИСПЛЕИ - ПРОСМОТР ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
- ОПИСАНИЕ - ПРОСМОТР ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ
- ФОРМИРОВАНИЕ - ФОРМИРОВАНИЕ НОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ИЗ ФРАГМЕНТОВ
- ГОМОЛОГИЯ - СРАВНЕНИЕ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
- СТРУКТУРА - ОБРАБОТКА ПЕРВИЧНОЙ СТРУКТУРЫ
- УЧАСТКИ - ПОИСК УЧАСТКОВ
- РЕСТРИКЦИЯ - ПРОГРАММА РЕСТРИКЦИИ
- ТРАНСЛЯЦИЯ - ПРОГРАММА ТРАНСЛЯЦИИ
- КОНЕЦ - КОНЧИТЬ ОБРАБОТКУ ПОСЛЕДОВАТЕЛЬНОСТИ

НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА

501- 550
 GAACAAGCTCTGAAAATTTGTGATGCATTATCTCGTTCTGGTGGGTGA
 * * * * *
 GAGCAGGCACTGAAAATCTGTGACGCCCTGGCGCGTCTGGCGCAGTAGA

551- 600
 TGYTATTGTCGTGGACTCCGTGGCAGCATTACACCCAAAAGCTGAAATTG
 * * * * *
 CGTTATCGTCGTGACTCCGTGGCGGCACTGACGCCGAAAAGCGGAAATCG

601- 650
 AAGGTGAAATTTGGTGATTCACACGTTGTTAGCCGCACGTATGATGAGC
 * * * * *
 AAGGCGAAATCGGCGACTCTCACATGGCCCTTGGCGCACGTATGATGAGC

651- 700
 CAAGCTATGCGTAAACTAGCGGGTAACCTTAAAAACTCTAATACACTGCT
 * * * * *
 CAGGCGATGCGTAAGCTGGCGGGTAACCTGAAGCAGTCCAACACGCTGCT

701- 750
 GATTTTCATTAAACCAATTCGTATGAAAATCGGTGTATGTTTGGTAACC
 * * * * *
 GATCTTCATCAACCAGATCCGTATGAAAATGGTGTGATGTTTGGTAACC

751- 800
 CAGAAACCACGACCGGTGGTAATGCGCTTAAATTCATGCTTCTGTTGCT
 * * * * *
 CGGAAACCACCTACCGGTGGTAACGCGCTGAAATTCACGCCICTGTTGCT

801- 850
 TTAGALATTCGTGCGCATTGGCTCTGTCAAAAATGGTGATGAAGTCATTGG
 * * * * *
 CTCGACATCCGTGATCGGCGCGGIGAAAGAGGGCGAAAACGTTGGTGGG

851- 900
 TAGTGAGACTCGCGTTAAATTTGTTAAAAATAAAGTGGCTGCACCGTTTA
 * * * * *
 TAGCGAAACCCCGCTGAAALTTGGTGAAGAACAATACTCTGCGCGCTTTA

901- 950
 AACAACTGAATTCCAAATATGTACGTTGAAGGTATTAATACCTATGGC
 * * * * *
 AACAGGCTGAATTCAGATCTCTACGGCGAAGGTATCAACTTCTACGGC

951- 1000
 GAACTGATTGATTTAGGTGTAAACATAAGTTAGTAGAGAAAGCAAGTGC
 * * * * *
 GAACTGGTTGACCTGGGCGTAAAAGAGAAGCTGATCGAGAAAGCAAGCGC

1001- 1050
 TTGGTATAGCTACAATGGCAGAAAAAATGGTCAAGGTAAAGCTAACGCAA
 * * * * *
 GTGGTACAGCTACAAGGTGAGAAGATCGGTCAGGGTAAAGCGAATGCCA

1051- 1100
 CCAATTACTIAAAGAACAATCCTGAAATGTACAATGAGTTAAACACTAAA
 * * * * *
 CTGCTGGCTGAAAGATAACCCGGAAACCGGAAAGAGATCGAGAAGAAA

КОНЕЦ = КОНЧИТЬ ОБРАБОТКУ ПЕРВИЧНОЙ СТРУКТУРЫ
 НАЗВАНИИ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 П
 РАБОТАЕТ ПРОГРАММА ВЫЧИСЛЕНИЯ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ НУКЛЕОТИДОВ И ПАР
 ПРОГРАММА СЧИТАЕТ ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ А,Т,С И ПАР А,Т И С,С.
 ЗАДАЙТЕ ДЛИНУ УЧАСТКОВ РАЗБИЕНИЯ
 200
 ДЛИНА УЧАСТКОВ РАЗБИЕНИЯ 200

НОМЕРА УЧАСТКОВ 19

 * 1 * 2 * 3 * 4 * 5 * 6 *

 А * 75117,5% * 53120,5% * 53126,5% * 50129,0% * 65132,5% * 75137,5%
 Т * 54129,5% * 54129,5% * 63131,5% * 60130,0% * 59129,5% * 50125,0%
 С * 35117,5% * 40123,0% * 44122,0% * 43121,5% * 49124,5% * 45122,5%
 С * 31115,5% * 42121,0% * 40120,0% * 39119,5% * 27113,5% * 30115,0%

 АТ * 134107,0% * 112150,0% * 116158,0% * 110159,0% * 124162,0% * 125162,5%
 СС * 66133,0% * 86144,0% * 84142,0% * 82141,0% * 76138,0% * 75137,5%

НОМЕРА УЧАСТКОВ

 * 7 *

 А * 43133,3%
 Т * 44134,1%
 С * 20117,1%
 С * 20115,5%

 АТ * 87167,4%
 СС * 42132,6%

 ПРОГРАММА ВЫЧИСЛЕНИЯ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ НУКЛЕОТИДОВ И ПАР КОНЧИЛА РАБОТУ
 ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПЕРВИЧНОЙ СТРУКТУРЫ :
 ПЛОТНОСТЬ = ВЫЧИСЛЕНИЕ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ НУКЛЕОТИДОВ И ПАР
 ТАБЛИЦА = ПОЛУЧЕНИЕ ТАБЛИЦЫ БЛИЖАЙШИХ СОСЕДЕЙ
 СОЧЕТАНИЯ = ПОДСЧЕТ СОЧЕТАНИЙ ПО ДВА И ПО ТРИ
 КОНЕЦ = КОНЧИТЬ ОБРАБОТКУ ПЕРВИЧНОЙ СТРУКТУРЫ
 НАЗВАНИИ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 Т
 ВЪ ВЪВРАТИ ТАБЛИЦА
 РАБОТАЕТ ПРОГРАММА ПОСТРОЕНИЯ ТАБЛИЦЫ БЛИЖАЙШИХ СОСЕДЕЙ

ТАБЛИЦА БЛИЖАЙШИХ СОСЕДЕЙ

А					Т					
* А *	* Т *	* С *	* А *	* Т *	* С *	* А *	* Т *	* С *	* А *	
-4*	136 *	146 *	98 *	81 *		-4*	150 *	120 *	82 *	60 *
-3*	127 *	131 *	95 *	69 *		-3*	126 *	124 *	71 *	71 *
-2*	137 *	140 *	80 *	65 *		-2*	117 *	99 *	105 *	72 *
-1*	168 *	97 *	76 *	81 *		-1*	113 *	119 *	89 *	73 *
0*	422 *	0 *	0 *	0 *		0*	0 *	394 *	0 *	0 *
1*	168 *	113 *	70 *	71 *		1*	97 *	119 *	121 *	50 *
2*	137 *	117 *	87 *	81 *		2*	140 *	99 *	92 *	61 *
3*	127 *	126 *	82 *	87 *		3*	131 *	124 *	74 *	63 *
4*	136 *	130 *	87 *	69 *		4*	106 *	120 *	93 *	73 *

Б					С					
* А *	* Т *	* С *	* А *	* Т *	* С *	* А *	* Т *	* С *	* А *	
-4*	87 *	93 *	51 *	53 *		-4*	69 *	73 *	52 *	34 *
-3*	82 *	74 *	77 *	51 *		-3*	87 *	63 *	41 *	37 *
-2*	87 *	92 *	53 *	52 *		-2*	81 *	61 *	46 *	40 *
-1*	70 *	121 *	53 *	40 *		-1*	71 *	50 *	66 *	35 *
0*	0 *	0 *	204 *	0 *		0*	0 *	0 *	0 *	229 *
1*	76 *	89 *	53 *	60 *		1*	81 *	73 *	40 *	35 *
2*	80 *	105 *	53 *	46 *		2*	65 *	72 *	52 *	40 *
3*	95 *	71 *	77 *	41 *		3*	69 *	71 *	51 *	37 *
4*	98 *	82 *	51 *	52 *		4*	81 *	60 *	53 *	34 *

ПРОГРАММА ПОСТРОЕНИЯ ТАБЛИЦЫ БЛИЖАЙШИХ СОСЕДЕЙ КОНЧИЛА РАБОТУ
 ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПЕРВИЧНОЙ СТРУКТУРЫ:
 ПЛОТНОСТЬ = ВЫЧИСЛЕНИЕ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ НУКЛЕОТИДОВ И ИХ ПАР
 ТАБЛИЦА = ПОЛУЧЕНИЕ ТАБЛИЦЫ БЛИЖАЙШИХ СОСЕДЕЙ
 СОЧЕТАНИЯ = ПОДСЧЕТ СОЧЕТАНИЙ ПО ДВА И ПО ТРИ
 КОНЕЦ = КОНЧИТЬ ОБРАБОТКУ ПЕРВИЧНОЙ СТРУКТУРЫ
 НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 С
 ВЫ ВЫБРАЛИ СОЧЕТАНИЯ
 РАБОТАЕТ ПРОГРАММА ОБРАБОТКИ СОЧЕТАНИЙ НУКЛЕОТИДОВ
 В ВЕРХНЕЙ СТРОКЕ=ЧИСЛО РЕАЛЬНО ВСТРЕТИВШИХСЯ СОЧЕТАНИЙ, В НИЖНЕЙ=МАТ.ОЖИДАНИЕ

СОЧЕТАНИЯ ПО ДВА

```

*****
AA * AT * AG * AC * TA * TT * TG * TC * GA * GT * GG * GC * CA * CT * CG * CC
*****
100* 113* 70* 71* 97* 118* 121* 56* 76* 89* 53* 66* 81* 73* 40* 35
*****
80* 70* 92* 80* 70* 56* 73* 63* 92* 73* 90* 83* 80* 63* 83* 72
*****

```

СОЧЕТАНИЯ ПО ТРИ

```

*****
AAA* AAT* AAB* AAC* ATA* ATT* ATG* ATC* AGA* AGT* AGG* AGC* ACA* ACT* ACC*
*****
72* 39* 29* 28* 18* 46* 32* 17* 19* 15* 15* 21* 20* 17* 11* 15
*****
22* 18* 23* 26* 18* 14* 18* 16* 23* 18* 24* 21* 20* 16* 21* 18
*****
TAA* TAT* TAB* TAC* TTA* TTT* TTG* TTC* TGA* TGT* TGG* TGC* TSA* TST* TCG* TCC
*****
36* 32* 17* 12* 41* 25* 36* 16* 41* 23* 29* 28* 22* 19* 10* 5
*****
10* 14* 18* 16* 14* 11* 15* 13* 16* 14* 19* 17* 16* 12* 16* 14
*****
BAA* BAT* BAB* BAS* BTA* BTT* BTG* BTC* BGA* BGT* BGG* BGC* BSA* BST* BCB* BCC
*****
32* 19* 9* 16* 22* 28* 29* 18* 8* 31* 3* 11* 10* 27* 12* 9
*****
23* 18* 24* 21* 18* 14* 19* 17* 24* 19* 25* 27* 21* 17* 22* 19
*****
CAA* CAT* CAB* CAS* CTA* CTT* CTG* CTC* CGA* CGT* CGG* CGC* CSA* CST* CCB* CSC
*****
28* 23* 15* 15* 16* 20* 24* 13* 8* 29* 6* 6* 13* 9* 7* 6
*****
26* 16* 21* 18* 10* 12* 16* 14* 21* 17* 22* 19* 16* 14* 19* 16
*****

```

ПРОГРАММА ОБРАБОТКИ СОЧЕТАНИЙ НУКЛЕОТИДОВ КОНЧИЛА РАБОТУ
 ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПЕРВИЧНОЙ СТРУКТУРЫ:
 ПЛОТНОСТЬ = ВЫЧИСЛЕНИЕ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ НУКЛЕОТИДОВ И ИХ ПАР
 ТАБЛИЦА = ПОЛУЧЕНИЕ ТАБЛИЦЫ БЛИЖАЙШИХ СОСЕДЕЙ
 СОЧЕТАНИЯ = ПОДСЧЕТ СОЧЕТАНИЙ ПО ДВА И ПО ТРИ
 КОНЕЦ = КОНЧИТЬ ОБРАБОТКУ ПЕРВИЧНОЙ СТРУКТУРЫ
 НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 К
 ВЫ ВЫБРАЛИ КОНЕЦ
 ПРОТОКОЛ ОБРАБОТКИ СОХРАНИТЬ? (ДА/НЕТ)
 ДА
 ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТИ
 ПЕЧАТЬ = ПЕЧАТЬ ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
 ДИСПЛЕЙ = ПРОСМОТР ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
 ОПИСАНИЕ = ПРОСМОТР ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ
 ФОРМИРОВАНИЕ = ФОРМИРОВАНИЕ НОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗ ФРАГМЕНТОВ
 ГОМОЛОГИЯ = СРАВНЕНИЕ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
 СТРУКТУРА = ОБРАБОТКА ПЕРВИЧНОЙ СТРУКТУРЫ
 УЧАСТКИ = ПОИСК УЧАСТКОВ
 РЕСТРИКЦИЯ = ПРОГРАММА РЕСТРИКЦИИ

ТРАНСЛЯЦИЯ = ПРОГРАММА ТРАНСЛЯЦИИ
КОНЕЦ = КОНЧИТЬ ОБРАБОТКУ ПОСЛЕДОВАТЕЛЬНОСТИ
НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
У
ВЫ ВЫБРАЛИ УЧАСТКИ
ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ПОИСКА УЧАСТКОВ:
ДАННЫМ = ПОИСК УЧАСТКОВ ДАННОГО ВИДА
 (ДОПУСКАЮТСЯ НЕСОВПАДЕНИЯ ТОЛЬКО ТИПА ЗАМЕНЫ)
ГОМОЛОГИЯ = ПОИСК УЧАСТКОВ СО СЛАБОЙ ГОМОЛОГИЕЙ
 (ДОПУСКАЮТСЯ НЕСОВПАДЕНИЯ ТИПА ЗАМЕНЫ, ВСТАВКИ, ВЫПАДЕНИЯ)
ПРЯМЫЕ = ПОИСК ПРЯМЫХ ПОВТОРОВ
ИНВЕРТИРОВАННЫЕ = ПОИСК ИНВЕРТИРОВАННЫХ ПОВТОРОВ (ПАЛИНДРОМОВ)
ШИПЛЬКИ = ПОИСК УЧАСТКОВ ДВОЙНОЙ СИММЕТРИИ ("ШИПЛЕК")
КОНЕЦ = КОНЧИТЬ ПОИСК УЧАСТКОВ
НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
Д
ВЫ ВЫБРАЛИ ДАННЫМ
РАБОТАЕТ ПРОГРАММА ПОИСКА УЧАСТКОВ ДАННОГО ВИДА
ЕСЛИ ПОСЛЕДОВАТЕЛЬНОСТЬ ЯВЛЯЕТСЯ ДНК, ПОИСК ВЕДЕТСЯ И НА КОМПЛЕМЕНТАРНОЙ НИТИ.
ПРИ ЗАДАНИИ УЧАСТКОВ ДОПУСТИМЫМИ СИМВОЛАМИ ЯВЛЯЮТСЯ А,Т,С,У,Н,(,/,).
Н ОЗНАЧАЕТ, ЧТО НА ЭТОМ МЕСТЕ МОЖЕТ СТОЯТЬ ЛЮБОЙ СИМВОЛ
ВОЗМОЖНОСТЬ ВЫБОРА ОДНОГО ИЗ ДВУХ ИЛИ БОЛЕЕ НУКЛЕОТИДОВ
ЗАДАЕТСЯ В ВИДЕ (А/Т) ИЛИ (А/Т/С)
ОДИН ИЗ ПАРАМЕТРОВ ПРОГРАММЫ = ДОПУСТИМОЕ КОЛИЧЕСТВО НЕСОВПАДЕНИЯ ТИПА ЗАМЕНЫ.
ПРИ НУЛЕВОМ ЗНАЧЕНИИ ИЩУТСЯ В ТОЧНОСТИ УКАЗАННЫЕ УЧАСТКИ
ПРОГРАММА ЗАКАНЧИВАЕТ РАБОТУ, КОГДА УКАЗЫВАЕТСЯ УЧАСТОК "*"**

ПОСЛЕ НАХОЖДЕНИЯ 50 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ 19
ВАС УСТРАИВАЕТ ЧИСЛО 50? (ДА/НЕТ)
НЕТ
ВВЕДИТЕ ДРУГОЕ ОГРАНИЧЕНИЕ
20
ПОСЛЕ НАХОЖДЕНИЯ 20 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ
БУДЕТ ЛИ КОЛИЧЕСТВО НЕСОВПАДЕНИИ ОДИНАКОВЫМ ДЛЯ ВСЕХ УЧАСТКОВ (ДА/НЕТ)?
ДА
ЗАДАЙТЕ КОЛИЧЕСТВО ДОПУСТИМЫХ НЕСОВПАДЕНИИ ТИПА ЗАМЕНЫ
0
КОЛИЧЕСТВО ДОПУСТИМЫХ НЕСОВПАДЕНИИ ТИПА ЗАМЕНЫ РАВНО 0
ВВЕДИТЕ УЧАСТОК ИЛИ "*"**
 АААТТТ
ВВЕДЕН УЧАСТОК
 АААТТТ
ВЫ ПОДТВЕРЖДАЕТЕ ЭТО? (ДА/НЕТ)
ДА

***** ОСНОВНАЯ НИТЬ *****
НАЙДЕННЫЕ УЧАСТКИ ЗАЧИНАЮТСЯ С ПОЗИЦИИ
 397 514 1290
ВСЕГО НАЙДЕНО 3 УЧАСТКОВ
ВВЕДИТЕ УЧАСТОК ИЛИ ***

ПРОГРАММА ПОИСКА УЧАСТКОВ КОНЧИЛА РАБОТУ
ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ПОИСКА УЧАСТКОВ:
ДАННЫМ = ПОИСК УЧАСТКОВ ДАННОГО ВИДА
 (ДОПУСКАЮТСЯ НЕСОВПАДЕНИЯ ТОЛЬКО ТИПА ЗАМЕНЫ)
ГОМОЛОГИЯ = ПОИСК УЧАСТКОВ СО СЛАБОЙ ГОМОЛОГИЕЙ
 (ДОПУСКАЮТСЯ НЕСОВПАДЕНИЯ ТИПА ЗАМЕНЫ, ВСТАВКИ, ВЫПАДЕНИЯ)
ПРЯМЫЕ = ПОИСК ПРЯМЫХ ПОВТОРОВ
ИНВЕРТИРОВАННЫЕ = ПОИСК ИНВЕРТИРОВАННЫХ ПОВТОРОВ (ПАЛИНДРОМОВ)
ШИПЛЬКИ = ПОИСК УЧАСТКОВ ДВОЙНОЙ СИММЕТРИИ ("ШИПЛЕК")
КОНЕЦ = КОНЧИТЬ ПОИСК УЧАСТКОВ
НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
Г
ВЫ ВЫБРАЛИ ГОМОЛОГИЯ
РАБОТАЕТ ПРОГРАММА ПОИСКА УЧАСТКОВ СО СЛАБОЙ ГОМОЛОГИЕЙ
ВВЕДИТЕ УЧАСТОК ИЛИ ***
 ААТТТГССССААА
ВВЕДЕН УЧАСТОК
 АААТТТГССССААА
ВЫ ПОДТВЕРЖДАЕТЕ ЭТО? (ДА/НЕТ)
ДА
ЗАДАЙТЕ КОЛИЧЕСТВО ДОПУСТИМЫХ НЕСОВПАДЕНИИ
0
КОЛИЧЕСТВО ДОПУСТИМЫХ НЕСОВПАДЕНИИ = 0
ЗАДАЙТЕ КОЛИЧЕСТВО НЕСОВПАДЕНИИ ТИПА ЗАМЕНЫ,
КОТОРОЕ ЭКВИВАЛЕНТНО ОДНОЙ ВСТАВКЕ ИЛИ ВЫПАДЕНИЮ

3
 3 ЗАМЕН ЭКВИВАЛЕНТНО ОДНОЙ ВСТАВКЕ ИЛИ ВЫПАДЕНИЮ
 ПОСЛЕ НАХОЖДЕНИЯ 50 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ
 ВАС УСТРАИВАЕТ ЧИСЛО 50?
 НЕТ
 ВВЕДИТЕ ДРУГОЕ ОГРАНИЧЕНИЕ

4
 ПОСЛЕ НАХОЖДЕНИЯ 4 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ

С ПОЗИЦИИ 323 НАЧИНАЕТСЯ УЧАСТОК N 1, НЕСОВПАДЕНИИ = 6, ИЗ НИХ ЗАМЕН = 68

```

TC AT TG
AA T GGCC A
AA T GGCC A
AT TG AA

```

С ПОЗИЦИИ 324 НАЧИНАЕТСЯ УЧАСТОК N 2, НЕСОВПАДЕНИИ = 6, ИЗ НИХ ЗАМЕН = 68

```

TC A C TG
A T TGG CC AA
A T TGG CC AA
AA T G CA

```

С ПОЗИЦИИ 397 НАЧИНАЕТСЯ УЧАСТОК N 3, НЕСОВПАДЕНИИ = 5, ИЗ НИХ ЗАМЕН = 31

```

T TT
AAATTTG GC A
AAATTTG GC A
G CC AA

```

С ПОЗИЦИИ 433 НАЧИНАЕТСЯ УЧАСТОК N 4, НЕСОВПАДЕНИИ = 6, ИЗ НИХ ЗАМЕН = 31

```

CCG AT T
A AATTT GC SAAA
A AATTT GC SAAA
GG C

```

НАЙДЕНО СВЫШЕ 4 УЧАСТКОВ. ПРОДОЛЖАТЬ ПОИСК?
 НЕТ
 ВВЕДИТЕ УЧАСТОК ИЛИ ***

 ПРОГРАММА ПОИСКА ГОМОЛОГИЧНЫХ УЧАСТКОВ КОНЧИЛА РАБОТУ
 ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ПОИСКА УЧАСТКОВ:

ДАННЫЕ = ПОИСК УЧАСТКОВ ДАННОГО ВИДА
 (ДОПУСКАЮТСЯ НЕСОВПАДЕНИЯ ТОЛЬКО ТИПА ЗАМЕН)

ГОМОЛОГИЯ = ПОИСК УЧАСТКОВ СО СЛАБОЙ ГОМОЛОГИЕЙ
 (ДОПУСКАЮТСЯ НЕСОВПАДЕНИЯ ТИПА ЗАМЕН, ВСТАВКИ, ВЫПАДЕНИИ)

ПРЯМЫЕ = ПОИСК ПРЯМЫХ ПОВТОРОВ

ИНВЕРТИРОВАННЫЕ = ПОИСК ИНВЕРТИРОВАННЫХ ПОВТОРОВ (ПАЛИНДРОМОВ)

ШИПЛЬКИ = ПОИСК УЧАСТКОВ ДВОЙНОЙ СИММЕТРИИ ("ШИПЛЕК")

КОНЕЦ = КОНЧИТЬ ПОИСК УЧАСТКОВ

НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СЛОВА

П
 ВЫ ВЫБРАЛИ ПРЯМЫЕ ПОВТОРЫ
 РАБОТАЕТ ПРОГРАММА ПОИСКА ПРЯМЫХ ПОВТОРОВ
 ВВЕДИТЕ МИНИМАЛЬНУЮ ДЛИНУ ПОВТОРА

5
 МИНИМАЛЬНАЯ ДЛИНА ПОВТОРА = 5
 ВВЕДИТЕ МИНИМАЛЬНУЮ ДЛИНУ "ГОЛОВКИ"

3
 МИНИМАЛЬНАЯ ДЛИНА "ГОЛОВКИ" = 3
 ВВЕДИТЕ МАКСИМАЛЬНУЮ ДЛИНУ "ГОЛОВКИ"

10
 МАКСИМАЛЬНАЯ ДЛИНА "ГОЛОВКИ" = 10
 ВВЕДИТЕ ДОПУСТИМОЕ ЧИСЛО НЕСОВПАДЕНИИ ТИПА ЗАМЕН

0
 КОЛИЧЕСТВО ДОПУСТИМЫХ НЕСОВПАДЕНИИ ТИПА ЗАМЕН = 0
 ПОСЛЕ НАХОЖДЕНИЯ 50 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ
 ВАС УСТРАИВАЕТ ЧИСЛО 50? (ДА/НЕТ)
 НЕТ
 ВВЕДИТЕ ДРУГОЕ ОГРАНИЧЕНИЕ

20
 ПОСЛЕ НАХОЖДЕНИЯ 20 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ

С ПОЗ.	9	19	ПОВТОР ДЛ. 5	С ГОЛОВКОЙ ДЛ. 5	ААТТА ААТТА
С ПОЗ.	323	337	ПОВТОР ДЛ. 5	С ГОЛОВКОЙ ДЛ. 9	ААТСТ ААТСТ
С ПОЗ.	566	569	ПОВТОР ДЛ. 5	С ГОЛОВКОЙ ДЛ. 4	СГТGG СГТGG
С ПОЗ.	593	605	ПОВТОР ДЛ. 8	С ГОЛОВКОЙ ДЛ. 7	ТыАААТТG ТГАААТТG

С ПОЗ. 594 606 ПОВТОР ДЛ. 7 С ГОЛОВКОЙ ДЛ. 7 GAAATTG GAAATTG
 С ПОЗ. 595 607 ПОВТОР ДЛ. 6 С ГОЛОВКОЙ ДЛ. 7 AAATTG AAATTG
 С ПОЗ. 596 608 ПОВТОР ДЛ. 5 С ГОЛОВКОЙ ДЛ. 7 AATTG AATTG
 С ПОЗ. 603 612 ПОВТОР ДЛ. 5 С ГОЛОВКОЙ ДЛ. 4 GGTGA GGTGA
 С ПОЗ. 610 625 ПОВТОР ДЛ. 5 С ГОЛОВКОЙ ДЛ. 10 TTGGT TTGGT
 С ПОЗ. 796 808 ПОВТОР ДЛ. 5 С ГОЛОВКОЙ ДЛ. 7 TTGGT TTGGT
 С ПОЗ. 864 873 ПОВТОР ДЛ. 6 С ГОЛОВКОЙ ДЛ. 4 GTTAAA GTTAAA
 С ПОЗ. 865 874 ПОВТОР ДЛ. 5 С ГОЛОВКОЙ ДЛ. 4 TTTAAA TTTAAA
 С ПОЗ. 866 881 ПОВТОР ДЛ. 6 С ГОЛОВКОЙ ДЛ. 10 TAAAGT TAAAGT
 С ПОЗ. 867 882 ПОВТОР ДЛ. 5 С ГОЛОВКОЙ ДЛ. 10 AAAGT AAAGT
 НАЙДЕНО 14 ПРЯМЫХ ПОВТОРОВ. МАКСИМАЛЬНАЯ ДЛИНА ПОВТОРА = 8
 ПРОГРАММА ПОИСКА ПРЯМЫХ ПОВТОРОВ КОНЧИЛА РАБОТУ
 ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ПОИСКА УЧАСТКОВ:
 ДАННЫЙ = ПОИСК УЧАСТКОВ ДАННОГО ВИДА
 ГОМОЛОГИЯ (ДОПУСКАЕТСЯ НЕСОВПАДЕНИЯ ТОЛЬКО ТИПА ЗАМЕНЫ)
 = ПОИСК УЧАСТКОВ СО СЛАБОЙ ГОМОЛОГИЕЙ
 (ДОПУСКАЕТСЯ НЕСОВПАДЕНИЯ ТИПА ЗАМЕНЫ, ВСТАВКИ, ВЫПАДЕНИЯ)
 ПРЯМЫЕ = ПОИСК ПРЯМЫХ ПОВТОРОВ
 ИНВЕРТИРОВАННЫЕ = ПОИСК ИНВЕРТИРОВАННЫХ ПОВТОРОВ (ПАЛИНДРОМОВ)
 ШПИЛЬКИ = ПОИСК УЧАСТКОВ ДВОЙНОЙ СИММЕТРИИ ("ШПИЛЕК")
 КОНЕЦ = КОНЧИТЬ ПОИСК УЧАСТКОВ
 НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 И
 ВЫ ВЫБРАЛИ ИНВЕРТИРОВАННЫЕ
 РАБОТАЕТ ПРОГРАММА ПОИСКА ИНВЕРТИРОВАННЫХ ПОВТОРОВ
 ВВЕДИТЕ МИНИМАЛЬНУЮ ДЛИНУ ИНВЕРТИРОВАННЫХ ПОВТОРОВ, КОТОРАЯ ВАС ИНТЕРЕСУЕТ
 5
 МИНИМАЛЬНАЯ ДЛИНА ИНВЕРТИРОВАННЫХ ПОВТОРОВ = 5
 ВВЕДИТЕ МИНИМАЛЬНУЮ ДЛИНУ "ГОЛОВКИ", КОТОРАЯ ВАС ИНТЕРЕСУЕТ
 3
 МИНИМАЛЬНАЯ ДЛИНА "ГОЛОВКИ" = 3
 ВВЕДИТЕ МАКСИМАЛЬНУЮ ДЛИНУ "ГОЛОВКИ", КОТОРАЯ ВАС ИНТЕРЕСУЕТ
 10
 МАКСИМАЛЬНАЯ ДЛИНА "ГОЛОВКИ" = 10
 ВВЕДИТЕ ДОПУСТИМОЕ ЧИСЛО НЕСОВПАДЕНИЙ
 9
 ДОПУСТИМОЕ ЧИСЛО НЕСОВПАДЕНИЙ = 9
 ПОСЛЕ НАХОЖДЕНИЯ 50 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ
 ВАС УСТРАИВАЕТ ЧИСЛО 50? (ДА/НЕТ)
 НЕТ
 ВВЕДИТЕ ДРУГОЕ ОГРАНИЧЕНИЕ
 20
 ПОСЛЕ НАХОЖДЕНИЯ 20 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ
 С ПОЗ. 9 20 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 6 AATTA ATTA
 С ПОЗ. 35 46 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 6 TATGA AGTAT
 С ПОЗ. 144 155 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 6 GAAAA AAAAG
 С ПОЗ. 145 153 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 3 AAAAC SAAAA
 С ПОЗ. 276 290 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 9 GACGT TGCAC
 С ПОЗ. 285 297 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 7 TTAGG GTATT
 С ПОЗ. 285 296 АНВ.ПОВТОР ДЛ. 6 С "ГОЛОВКОЙ" ДЛ. 5 TTAGGT TGGATT
 С ПОЗ. 285 295 АНВ.ПОВТОР ДЛ. 7 С "ГОЛОВКОЙ" ДЛ. 3 TTAGGTG GTGGATT
 С ПОЗ. 451 465 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 9 AGSTA ATCSA
 С ПОЗ. 470 485 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 10 SAACT TCAAC
 С ПОЗ. 788 796 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 3 TGCTT TTGGT
 С ПОЗ. 835 850 АНВ.ПОВТОР ДЛ. 6 С "ГОЛОВКОЙ" ДЛ. 9 GTGATG GTAGTG
 С ПОЗ. 917 934 АНВ.ПОВТОР ДЛ. 7 С "ГОЛОВКОЙ" ДЛ. 10 AATTATG GTATTAA
 С ПОЗ. 933 945 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 7 GGAT TATGG
 С ПОЗ. 1066 1080 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 9 AACAT TACAA
 С ПОЗ. 1068 1099 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 6 GTTAA AATTG
 С ПОЗ. 1068 1098 АНВ.ПОВТОР ДЛ. 6 С "ГОЛОВКОЙ" ДЛ. 4 GTTAAA AAATTG
 С ПОЗ. 1129 1139 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 5 GTGAA AAGTG
 С ПОЗ. 1268 1216 АНВ.ПОВТОР ДЛ. 5 С "ГОЛОВКОЙ" ДЛ. 3 TTGCT CTGTT
 НАЙДЕНО 14 ИНВЕРТИРОВАННЫХ ПОВТОРОВ. МАКСИМАЛЬНАЯ ДЛИНА = 7
 ПРОГРАММА ПОИСКА ИНВЕРТИРОВАННЫХ ПОВТОРОВ КОНЧИЛА РАБОТУ
 ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ПОИСКА УЧАСТКОВ:
 ДАННЫЙ = ПОИСК УЧАСТКОВ ДАННОГО ВИДА
 ГОМОЛОГИЯ (ДОПУСКАЕТСЯ НЕСОВПАДЕНИЯ ТОЛЬКО ТИПА ЗАМЕНЫ)
 = ПОИСК УЧАСТКОВ СО СЛАБОЙ ГОМОЛОГИЕЙ
 (ДОПУСКАЕТСЯ НЕСОВПАДЕНИЯ ТИПА ЗАМЕНЫ, ВСТАВКИ, ВЫПАДЕНИЯ)
 ПРЯМЫЕ = ПОИСК ПРЯМЫХ ПОВТОРОВ
 ИНВЕРТИРОВАННЫЕ = ПОИСК ИНВЕРТИРОВАННЫХ ПОВТОРОВ (ПАЛИНДРОМОВ)
 ШПИЛЬКИ = ПОИСК УЧАСТКОВ ДВОЙНОЙ СИММЕТРИИ ("ШПИЛЕК")
 КОНЕЦ = КОНЧИТЬ ПОИСК УЧАСТКОВ
 НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 И

Вы выбрали шпильки
 Работает программа поиска участков двойной симметрии ("шпильки")
 Введите минимальную длину "шпильки", которая вас интересует
 5
 Минимальная длина "шпильки" = 5
 Введите минимальную длину "головки", которая вас интересует
 3
 Минимальная длина "головки" = 3
 Введите максимальную длину "головки", которая вас интересует
 10
 Максимальная длина "головки" = 10
 Введите допустимое число несовпадений
 0
 Допустимое число несовпадений = 0
 После нахождения 50 участков нужного вида поиск можно будет прервать
 вас устраивает число 50? (да/нет)
 нет
 Введите другое ограничение
 20
 После нахождения 20 участков нужного вида поиск можно будет прервать
 С ПОЗ. 9 18 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 4 AATTA TAATT
 С ПОЗ. 29 45 "шпилька" дл. 6 С "ГОЛОВКОЙ" дл. 10 ATACTG CAGTAT
 С ПОЗ. 29 44 "шпилька" дл. 7 С "ГОЛОВКОЙ" дл. 8 ATACTGT ACAGTAT
 С ПОЗ. 29 43 "шпилька" дл. 8 С "ГОЛОВКОЙ" дл. 6 ATACTGTA ACAGTAT
 С ПОЗ. 29 42 "шпилька" дл. 9 С "ГОЛОВКОЙ" дл. 4 ATACTGTAT ATACAGTAT
 С ПОЗ. 39 49 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 5 ATTAI ATAAT
 С ПОЗ. 39 48 "шпилька" дл. 6 С "ГОЛОВКОЙ" дл. 3 ATTATA TATAAT
 С ПОЗ. 180 194 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 4 CAAAT ATTTG
 С ПОЗ. 180 193 "шпилька" дл. 6 С "ГОЛОВКОЙ" дл. 7 CAAAT ATTTG
 С ПОЗ. 180 192 "шпилька" дл. 7 С "ГОЛОВКОЙ" дл. 5 CAAATG CAATTTG
 С ПОЗ. 231 239 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 3 CGTTC GAACG
 С ПОЗ. 243 255 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 7 CTACA TCTAC
 С ПОЗ. 454 467 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 4 TGTCG CGACA
 С ПОЗ. 484 496 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 3 ACCTG CAGGT
 С ПОЗ. 664 617 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 8 GTYAA TTCAL
 С ПОЗ. 756 766 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 5 ACCAC GTTGT
 С ПОЗ. 932 941 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 4 AGGTA TACST
 С ПОЗ. 1100 1104 "шпилька" дл. 5 С "ГОЛОВКОЙ" дл. 9 GTCAW CTGAC
 найдем 10 "шпильки". Максимальная длина = 4
 программа поиска "шпильки" кончила работу
 предлагает вам следующие виды поиска участков
 данные = поиск участков данного вида
 (допускается несовпадения только типа зачехи)
 гомологии = поиск участков со слабой гомологией
 (допускается несовпадения типа зачехи, впадения)
 прямые = поиск прямых повторов
 инвертированные = поиск инвертированных повторов (палиндромов)
 шпильки = поиск участков двойной симметрии ("шпильки")
 конец = кончат поиск участков
 названия режимов можно сокращать до 1-го символа
 к
 вы выбрали конец
 протокол обработки лохранто? (да/нет)
 да
 предлагаем вам следующие виды обработки последовательности:
 печать = печать последовательности
 дисплей = просмотр последовательности на дисплее
 описание = просмотр описания последовательности
 формирование = формирование новой последовательности из фрагментов
 гомология = сравнение двух последовательностей
 структура = обработка первичной структуры
 участки = поиск участков
 рестрикции = программа рестрикции
 трансляция = программа трансляции
 конец = кончить обработку последовательности
 названия режимов можно сокращать до 1-го символа
 р
 вы выбрали рестрикции
 работает программа рестрикции
 у вас есть вопросы к символу стандартных рестриктаз? (да/нет)
 да
 работаем со списком стандартных рестриктаз
 предлагаем следующие режимы работы:
 весь = просмотреть весь список

19

САИТ - ПУСМОТРЕТЬ САИТ УКАЗАНОИ РЕСТРИКТАЖИ
РЕСТРИКТАЖИ - ПУСМОТРЕТЬ ВСЕ РЕСТРИКТАЖИ С УКАЗАНЫМ САИТОМ
КОДЕИ - КОМЧАТЬ РАБОТУ С ПИСПОКОМ
НАЗВАНИИ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СЛИБУЛА

B

ВНИ БИДРАЛИ БЕСЬ

ACCI	GT(A/C)(G/T)AC
ACCI	CG#CG
ATUI	#CC(A/T)GG
ATUI	#CC(A/T)GG
ATUCI	T#GATCA
ACVI	G(PU)#CG(PY)C
AUSI	TG#GCLA
ANSII	G(PU)#CG(PY)C
ASUI	G#GNC
AVAI	C#(PY)CG(PU)G
AVAI	G#G(A/T)CC
AVAI	ATGCAI
AVAI	C#(PY)CG(PU)G
AVRI	CCTAGG
AVRI	AG#CT
ALUI	CC#(A/T)GG
APVI	CCG#GG
BALI	G#GATLC
BAMFI	G#GATLC
BAMHI	G#GATLC
BAMKI	G#GATLC
BAMNI	G#GATLC
BAMNX	G#G(A/T)CC
BBVSI	GC(A/T)GC
BBVI	GC(A/T)GC
BCLI	T#GATCA
BCE170	CTGLAAG
BCEH	CG#CG
BGLI	GCCNNNN#NGGC
BGLII	A#GATCT
BSPRI	GG#CC
BSTI	G#GATCC
BSTII	G#GTNACC
BSTPI	G#GTNACC
BSTNI	CC#(A/T)GG
BSURI	GG#CC
BSU1070	GG#CC
BSU1114	GG#CC
BSU1247	CTGCCAG
BBRI	A#AGCTT
BALI	TGG#CLA
BLUI	C#TCGAG
BLUII	GG#CC
BLUI	AT#CGAT
CLAI	GG#CC
CLTI	G#G(A/T)CC
CAUL	GLG#C
CFUI	A#AGCTT
CHUL	GT(PY)#(PU)AC
CHUII	T#GATCA
CPEI	C#TNAG
DDEI	G#ATC
DPNI	#GATC
DPNI	#LC(A/T)GG
ECLII	G#GTNACC
ECAI	CCG#GG
ECCI	G#AATTC
ECURI	(PU)(PU)A#T(PY)(PY)
ECURI	A#AGCTT
HSUI	G(A/T)GC(A/T)#C
HGIAI	G#G(A/T)CC
HGIBL	G#G(A/T)CC
HGICII	G#G(A/T)CC
HGICII	G#TGCAC

MGIEI	GG(A/T)CC
KPNI	GLTACGL
MSTI	TLGGCA
MBUI	#GATC
MBUII	GAAGA(N)6#
MNUI	C#CGG
MNUIII	#GATC
ECURII	#CC(A/T)GG
ECOB	TGA(N)8TGCT
ECUK	AAC(N)6GTGC
ECUPL	AGACC
ECUPI5	CAGCAL
FSPAII	G#GTNACC
FNUAI	G#ANTL
FNUAII	#GATC
FNUCL	#GATC
FNUDI	GG#CC
FNUDII	C#CGG
FNUDIII	GCG#C
FNUEI	#GATC
FNUAH I	GC#NGL
MAEI	(A/T)GG#CC(A/T)
MAEII	(PU)GLGC#(PY)
MAEIII	GG#CC
MAPII	C#CGG
HGAI	GACGC(N)5#
HMGII	GG#CC
HMAI	GLG#C
HMAII	G#ANTL
HINL95bI	C#FLG
HINBIII	A#AGCTT
HINCII	GT(PY)#(PU)AC
HINDI	CAC
HINDII	GT(PY)#(PU)AC
HINDIII	A#AGCTT
HINDIV	GAC
HINFII	G#ANTC
HINFIII	A#AGCTT
HINHII	(PU)GLGC#(PY)
HPMI	GLTGA(N)8#
HPAI	GT#AAC
HPAII	C#CGG
MNLI	CCTL
MNLII	GT(PY)#(PU)AC
MNNII	GL#CC
MNNIV	GCG#C
MUBI	#GATC
MSPII	C#CGG
MSIII	C#TCGAG
NGUI	(PU)GLGC#(PY)
NGUII	GG#CC
OXAI	AG#CT
PVUI	CGATCG
PVUII	CAG#CTG
PALI	GG#CC
PSYI	CTGCAGG
PFAI	#GATC
PMAI	CTGCAGG
RSPI	CGATCG
RSHI	C#AT#CG
RSAI	GT#AC
RSHII	G#AATTC
SMAI	CCC#GCG
SAU3A	#GATC
SAU9bI	G#GNCL
SFAI	GL#CC
SFANI	GATGC
SACI	GAGCT#C

SACII	CCGC#GG
SALPI	CTGLA#G
SPAI	C#TCGAG
SALI	G#TLGAC
SBUI	CCGC#GG
SBNI	CCGC#GG
SCUI	C#TLGAG
SEXI	C#TLGAG
SGUI	C#TLGAG
SHYI	CCGC#GG
SLAI	C#TLGAG
SLUI	C#TLGAG
SPMI	GCATG#C
SSII	GAGCT#C
SSTII	CCGL#GG
THAL	CG#LG
TGLI	CCGC#GG
TAGI	T#CLA
XAMI	G#TLGAC
XBAI	T#CTAGA
XNUI	C#TCGAG
XNUII	(PU)#BATIC(PY)
XMAI	C#CCGGG
XMAII	CTGLA#G
XMAIII	C#GGCLG
XNII	CGATC#
XONI	CTGLA#G
XONII	CGATC#G
XPAI	C#TCGAG

ПРЕДЛАГАЕМ СЛЕДУЮЩИЕ РЕЖИМЫ РАБОТЫ:
 ВЕСЬ = ПРОСМОТРЕТЬ ВСЕ СПИСОК
 САИТ = ПОСМОТРЕТЬ САИТ УКАЗАННОЙ РЕСТРИКТАЗЫ
 РЕСТРИКТАЗЫ = ПОСМОТРЕТЬ ВСЕ РЕСТРИКТАЗЫ С УКАЗАННЫМ САИТОМ
 КОНЕЦ = КОНЧИТЬ РАБОТУ СО СПИСОКОМ
 НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА

C
 ВЫ ВЫБРАЛИ САИТ
 УКАЖИТЕ ИМЯ РЕСТРИКТАЗЫ

ALUI	
ALUI	AG#ST

ПРЕДЛАГАЕМ СЛЕДУЮЩИЕ РЕЖИМЫ РАБОТЫ:
 ВЕСЬ = ПРОСМОТРЕТЬ ВСЕ СПИСОК
 САИТ = ПОСМОТРЕТЬ САИТ УКАЗАННОЙ РЕСТРИКТАЗЫ
 РЕСТРИКТАЗЫ = ПОСМОТРЕТЬ ВСЕ РЕСТРИКТАЗЫ С УКАЗАННЫМ САИТОМ
 КОНЕЦ = КОНЧИТЬ РАБОТУ СО СПИСОКОМ
 НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 ВЫ ВЫБРАЛИ РЕСТРИКТАЗЫ

УКАЖИТЕ САИТ РЕСТРИКЦИИ

GG#CC	
BSP#I	GG#CC
BSUN#I	GG#CC
BSU1#76	GG#CC
BSU1114	GG#CC
BLUI#I	GG#CC
CLY#I	GG#CC
FNU#I	GG#CC
MAE#I#I	GG#CC
MNG#I	GG#CC
MNN#I#I	GG#CC
NGO#I#I	GG#CC
PAL#I	GG#CC
SFA#I	GG#CC

ПРЕДЛАГАЕМ СЛЕДУЮЩИЕ РЕЖИМЫ РАБОТЫ:
 ВЕСЬ = ПРОСМОТРЕТЬ ВСЕ СПИСОК
 САИТ = ПОСМОТРЕТЬ САИТ УКАЗАННОЙ РЕСТРИКТАЗЫ
 РЕСТРИКТАЗЫ = ПОСМОТРЕТЬ ВСЕ РЕСТРИКТАЗЫ С УКАЗАННЫМ САИТОМ

КОНЕЦ - КОНЧАТЬ РАБОТУ СО СПИСКОМ
 НАЗВАНИИ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 К
 ВЫ ВЫБРАЛИ КОНЕЦ
 ЗАКОНЧЕНА РАБОТА СО СПИСКОМ СТАНДАРТНЫХ РЕСТРИКТАЗ

СУЩЕСТВУЕТ ЛИ У ВАС ЛИЧНЫЙ НАБОР РЕСТРИКТАЗ?(ДА/НЕТ)
 НЕТ

НАЧАЛОСЬ ФОРМИРОВАНИЕ НАБОРА РЕСТРИКТАЗ
 НУЖНА ЛИ ИНСТРУКЦИЯ ПО ВВОДУ РЕСТРИКТАЗ (ДА/НЕТ)?

ДА
 ВВОДИТЕ НАЗВАНИЕ РЕСТРИКТАЗ - ПОЛЕ СХУМЕ В СПИСКЕ РОБЕРТСА
 КАЖДОЕ НАЗВАНИЕ ВВОДИТЕ СО СЛЕДУЮЩЕЙ СТОКИ
 ВМЕСТО РИМСКОЙ ЦИФРЫ "1" В НАЗВАНИЯХ ВВОДИТЕ
 ЛАТИНСКУЮ БУКВУ "I" БЕЗ ПРОБЕЛА ОТ ПРЕДЫДУЩИХ БУКВ

ДЛЯ РЕСТРИКТАЗ, ОТСУТСТВУЮЩИХ В СПИСКЕ РОБЕРТСА,
 ЧЕРЕЗ ПРОБЕЛ ОТ НАЗВАНИЯ ВВОДИТЕ САЙТ РЕСТРИКЦИИ
 ПРИ УКАЗАННОМ САЙТЕ РЕСТРИКЦИИ И НАЛИЧИЕМ ЕГО В СПИСКЕ РОБЕРТСА
 РАССМАТРИВАЕТСЯ УКАЗАННЫМ ПОЛЬЗОВАТЕЛЕМ САЙТ
 ПРИЗНАКОМ КОНЦА ВВОДА ЯВЛЯЕТСЯ СТРОКА ***

НУЖНА ЛИ ИНСТРУКЦИЯ ПО ВВОДУ САЙТОВ РЕСТРИКЦИИ?(ДА/НЕТ)

ДА
 ДОПУСТИМЫМИ СИМВОЛАМИ ПРИ ВВОДЕ САЙТОВ РЕСТРИКЦИИ ЯВЛЯЮТСЯ
 A,T,G,C,U,R,Y,(,/,),N,P,1,2,3,4,5,6,7,8,9
 СИМВОЛ P РАЗРЕШАЕТСЯ ТОЛЬКО В СЧЕТАНИЯХ PU ДЛЯ ПУРИНА И PY ДЛЯ ПИРИМИДИНА
 ВОЗМОЖНОСТЬ ВЫБОРА ОДНОГО ИЗ ДВУХ НУКЛЕОТИДОВ
 ЗАПИСЫВАЕТСЯ В СКОБКАХ ЧЕРЕЗ НАКЛОННУЮ ЧЕРТУ, НАПРИМЕР, (A/T)
 МЕСТО РАЗРЕЗАНИЯ УКАЗЫВАЕТСЯ СИМВОЛОМ #
 ДОПУСКАЕТСЯ СОКРАЩЕНИЕ ТИПА N(4) ДЛЯ NNNN
 ЧИСЛО ПОПТОРЕНИИ ДОЛЖНО БЫТЬ ОДНОЗНАЧНЫМ

ВВЕДИТЕ ИМЯ РЕСТРИКТАЗА ИЛИ ***

ALUI

ВВЕДИТЕ ИМЯ РЕСТРИКТАЗА ИЛИ ***

MAEIII

ВВЕДИТЕ ИМЯ РЕСТРИКТАЗА ИЛИ ***

СФОРМИРОВАН СЛЕДУЮЩИЙ НАБОР РЕСТРИКТАЗ:

ALUI

ALGCT

MAEIII

GGCC

ПОДТВЕРДИТЕ ПРАВИЛЬНОСТЬ (ДА/НЕТ)

ДА

НАБОР РЕСТРИКТАЗ СФОРМИРОВАН

СОХРАНИТЬ СФОРМИРОВАННЫЙ НАБОР? (ДА/НЕТ)

НЕТ

НАБОР РЕСТРИКТАЗ НЕ СОХРАНЕН

ПРИ НАХОЖДЕНИИ САЙТА НА КОМПЛЕМЕНТАРНОЙ НИТИ В НИЖНЕЙ СТРОКЕ -

НОМЕР СООТВЕТСТВУЮЩЕЙ ПОЗИЦИИ ОСНОВНОЙ НИТИ

РЕСТРИКТАЗА ALUI САЙТ ALGCT

*** ОСНОВНАЯ НИТЬ ***

451 506 590 653 905 1008 1040 1022 1325

ВСЕГО НАЙДЕНО 9 УЧАСТКОВ

РЕСТРИКТАЗА MAEIII САЙТ GGCC

*** ОСНОВНАЯ НИТЬ ***

164 309 339

ВСЕГО НАЙДЕНО 3 УЧАСТКОВ

ЯВЛЯЕТСЯ ЛИ МОЛЕКУЛА ДНК ДВУМНУКЛЕОТИДНОЙ? (ДА/НЕТ)

ДА

* НАЧ. ПОЗ.	КОМ. ПОЗ.	ДЛ. В ПУЗ. ПУЗ. ПУЗ. ПУЗ.	БИ ОБРАЗОВАНЫ РЕСТРИКТАЗАМИ	И	ПОЛ. ВЕС	И	ДЛ. В АНГСТРЕМ.	И	% GC-ПАР
*	1	163	163	НАЧАЛО MAEIII	100294.		554.20		30.06
*	164	308	145	MAEIII MAEIII	89241.		493.00		45.52
*	309	329	21	MAEIII MAEIII	12923.		71.40		38.10
*	330	450	121	MAEIII ALUI	74403.		411.40		39.67
*	451	505	55	ALUI ALUI	33050.		187.00		45.45
*	506	589	84	ALUI ALUI	51690.		285.60		45.24
*	590	652	63	ALUI ALUI	38773.		214.20		44.44
*	653	904	252	ALUI ALUI	155079.		850.00		39.29
*	905	1007	103	ALUI ALUI	63382.		350.20		35.92
*	1008	1034	32	ALUI ALUI	19692.		100.00		37.50
*	1040	1201	162	ALUI ALUI	99691.		554.00		37.65

* 1202 1324 123 ALUI ALUI 75085. 410.00 30.52
 * 1325 1329 5 MAEIII KONEC 3077. 17.00 40.00

ПРОГРАММА РЕСТРИКЦИИ КОНЧИЛА РАБОТУ
 ПРОТОКОЛ ОБРАБОТКИ СОХРАНИТЬ? (ДА/НЕТ)

ДА

НУЖНА ЛИ ИНСТРУКЦИЯ ПО ВВОДУ САИТОВ РЕСТРИКЦИИ? (ДА/НЕТ)

ДА

ДОПУСТИМЫМИ СИМВОЛАМИ ПРИ ВВОДЕ САИТОВ РЕСТРИКЦИИ ЯВЛЯЮТСЯ

A,T,G,C,U,P,Y,(,/,),N,#,1,2,3,4,5,6,7,8,9

СИМВОЛ P РАЗРЕШАЕТСЯ ТОЛЬКО В СЧЕТАНИЯХ PU ДЛЯ ПУРИНА И PY ДЛЯ ПИРИМИДИНА

ВОЗМОЖНОСТЬ ВЫБОРА ОДНОГО ИЗ ДВУХ НАКЛЕТИДОВ

ЗАИМЕСНЯЕТСЯ В СЛОВКАХ ЧЕРЕЗ НАКЛОМННУЮ ЧЕРТУ, НАПРИМЕР, (A/T)

МЕСТО РАЗРЕЗАНИЯ УКАЗЫВАЕТСЯ СИМВОЛОМ #

ДОПУСКАЮТСЯ СОКРАЩЕНИЯ ТИПА N(4) ДЛЯ NNNN

ЧИСЛО ПОВТОРЕНИИ ДОЛЖНО БЫТЬ ОДНОЗНАЧНЫМ

ВВЕДИТЕ ИМЯ РЕСТРИКТАЗА ИЛИ ***

ALUI

ВВЕДИТЕ ИМЯ РЕСТРИКТАЗА ИЛИ ***

MAEIII

ВВЕДИТЕ ИМЯ РЕСТРИКТАЗА ИЛИ ***

СФОРМИРОВАН СЛЕДУЮЩИЙ НАБОР РЕСТРИКТАЗ:

ALUI A6ACT

MAEIII 66WCC

ПОДТВЕРДИТЕ ПРАВИЛЬНОСТЬ (ДА/НЕТ)

ДА

НАБОР РЕСТРИКТАЗ СФОРМИРОВАН

СОХРАНИТЬ СФОРМИРОВАННЫЙ НАБОР? (ДА/НЕТ)

НЕТ

НАБОР РЕСТРИКТАЗ НЕ СОХРАНЕН

ПРИ НАИЖДЕИИ САИТА НА КОМПЛЕМЕНТАРНОЙ НИТИ В НИЖНЕЙ СТРОКЕ

НОМЕР СООТВЕТСТВУЮЩЕЙ ПОЗИЦИИ ОСНОВНОЙ НИТИ

РЕСТРИКТАЗА ALUI САИТ A6ACT

*** ОСНОВНАЯ НИТЬ ***

451 506 590 653 905 1007 1007 1202 1325

ВСЕГО НАЙДЕНО 9 УЧАСТКОВ

РЕСТРИКТАЗА MAEIII САИТ 66WCC

*** ОСНОВНАЯ НИТЬ ***

164 309 339

ВСЕГО НАЙДЕНО 3 УЧАСТКОВ

ЯВЛЯЕТСЯ ЛИ МОЛЕКУЛА ДНК ДВУНИТЕВОЙ? (ДА/НЕТ)

DA	* НАЧ.:	КОМ.ИДЛ. В1	ОБРАЗОВАН	:	МОЛ.	:	ДЛ. В	:	X
* ПОЗ.:	ПОЗ.ИМУЛ.1	РЕСТРИКТАЗАМИ	:	ВЕС	:	ТАНГСТРЕН	:	СС-ПАР	
*	1	163	163 НАЧАЛО MAEIII	:	100294.	:	554.00	:	30.06
*	164	308	145 MAEIII MAEIII	:	09241.	:	493.00	:	45.52
*	309	329	21 MAEIII MAEIII	:	12923.	:	71.40	:	38.10
*	330	450	121 MAEIII ALUI	:	74403.	:	411.40	:	39.67
*	451	505	55 ALUI ALUI	:	33050.	:	187.00	:	45.45
*	506	589	84 ALUI ALUI	:	51690.	:	285.00	:	45.24
*	590	652	63 ALUI ALUI	:	38773.	:	214.20	:	44.44
*	653	904	252 ALUI ALUI	:	155079.	:	850.00	:	39.29
*	905	1007	103 ALUI ALUI	:	63380.	:	350.20	:	35.92
*	1008	1034	32 ALUI ALUI	:	19690.	:	100.00	:	37.50
*	1040	1201	162 ALUI ALUI	:	99691.	:	55.00	:	37.65
*	1202	1324	123 ALUI ALUI	:	75085.	:	410.00	:	30.52
*	1325	1329	5 MAEIII KONEC	:	3077.	:	17.00	:	40.00

ПРОГРАММА РЕСТРИКЦИИ КОНЧИЛА РАБОТУ
 ПРОТОКОЛ ОБРАБОТКИ СОХРАНИТЬ? (ДА/НЕТ)

ДА

ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- ПЕЧАТЬ - ПЕЧАТЬ ПОСЛЕДОВАТЕЛЬНОСТИ
 - ДИСПЛЕИ - ПРОСМОТР ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
 - ОПИСАНИЕ - ПРОСМОТР ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ
 - ФОРМИРОВАНИЕ - ФОРМИРОВАНИЕ НОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗ ФРАГМЕНТОВ
 - ГОМОЛОГИЯ - СРАВНЕНИЕ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
 - СТРУКТУРА - ОБРАБОТКА ПЕРВИЧНОЙ СТРУКТУРЫ
 - УЧАСТКИ - ПОИСК УЧАСТКОВ
 - РЕСТРИКЦИЯ - ПРОГРАММА РЕСТРИКЦИИ
 - ТРАНСЛЯЦИЯ - ПРОГРАММА ТРАНСЛЯЦИИ
 - КОНЕЦ - КОНЧИТЬ ОБРАБОТКУ ПОСЛЕДОВАТЕЛЬНОСТИ
- НАЗВАНИЯ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА

У
 ВЫ ВЫБРАЛИ ТРАНСЛЯЦИЯ

ПОСЛЕДОВАТЕЛЬНОСТЬ ЗАДАНА В ВИДЕ ДНК -
 СЧИТАЕМ ЕЕ МАТРИЧНОЙ РНК С ЗАМЕНОЙ U НА T
 СТАНДАРТНЫЕ ИНИЦИАТОРЫ = AUG, GUG
 ВЫ БУДЕТЕ РАБОТАТЬ СО СТАНДАРТНЫМИ ИНИЦИАТОРАМИ? (ДА/НЕТ)
 ДА

СТАНДАРТНЫЕ ТЕРМИНАТОРЫ = UAA, UAG, UGA
 ВЫ БУДЕТЕ РАБОТАТЬ СО СТАНДАРТНЫМИ ТЕРМИНАТОРАМИ? (ДА/НЕТ)
 ДА

РАБОТАЕМ С ИНИЦИАТОРАМИ 3ATG GTG
 РАБОТАЕМ С ТЕРМИНАТОРАМИ ТАА ТАБ TGA
 ЗАДАЙТЕ МИНИМАЛЬНУЮ ДЛИНУ БЕЛКА, КОТОРЫЙ ВАС ИНТЕРЕСУЕТ
 350
 ПРОГРАММА РАССМАТРИВАЕТ БЕЛКИ ДЛИНОЙ НЕ МЕНЬШЕ 350
 ПОСЛЕ НАХОЖДЕНИЯ 50 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ
 ВАС УСТРАНЯЕТ ЧИСЛО 50? (ДА/НЕТ)
 НЕТ

ВВЕДИТЕ ДРУГОЕ ОГРАНИЧЕНИЕ

5
 ПОСЛЕ НАХОЖДЕНИЯ 5 УЧАСТКОВ НУЖНОГО ВИДА ПОИСК МОЖНО БУДЕТ ПРЕРВАТЬ
 НУЖНО ЛИ ВЫЧИСЛИТЬ ЗАРЯД БЕЛКА ПРИ ЗАДАННОМ PH (ДА/НЕТ)?
 ДА

ЗАДАЙТЕ ЗНАЧЕНИЕ PH (OT 2 ДО 12)

7.5

ЗАРЯД БЕЛКА ВЫЧИСЛЯЕТСЯ ПРИ PH= 7.50
 НУЖНО ЛИ ВЫЧИСЛИТЬ ЭЛЕКТРИЧЕСКУЮ ТОЧКУ БЕЛКОВ? (ДА/НЕТ)

ДА

НУЖНО ЛИ ПЕЧАТАТЬ АМИНОКЛОТНЫЙ СОСТАВ БЕЛКОВ? (ДА/НЕТ)

ДА

НУЖНО ЛИ ПЕЧАТАТЬ ТАБЛИЦУ ВСТРЕЧАЕМОСТИ КОДОНОВ? (ДА/НЕТ)

ДА

*** ОСНОВНАЯ НИТЬ ***

** 1 ** С ПОЗИЦИИ 105- 1196
 МОЖЕТ БЫТЬ СЧИТАН БЕЛОК МОЛЕКУЛЯРНОГО ВЕСА 39176,16
 С ЗАРЯДОМ 6,74 ПРИ PH= 7,50

ИЗОЭЛЕКТРИЧЕСКАЯ ТОЧКА БЕЛКА = 8,718
 V L H C P Q G S N D M A I N Q D K E K A L
 A A A L G E I Q K E F F G K G S I M R L G
 Q N R S M D V Q T I S T G S L S L N V A
 L G A G G L P R G R I V Q I Y G P Q S S
 Q K T L T L E V I A B A E R Q C K I C
 A F I N A G H A L N P I Y A E K L G V N
 I N D L L C S E P V T G Q E A L Q I C N
 A L S R S G A V N V I V V N S V A A L T
 P K A Q I Q G Q I G N S H V G L A A R M
 M S E A M R K L A G D L K D S D T L L I
 F I D E I R M K I G V M F G D P Q T T T
 G G D A L K F Y A S V R L N I R I G S
 V K D G N Q V I G S Q T R V K V V K D K
 V A A P F K F A Q F E I M Y G Q G I D T
 Y G Q L I N L G V K H K L V Q K A G A N
 Y S Y D G Q K I G E G K A D A T O Y L K
 Q H P Q M Y D Q L D T K L R Q N L L R H
 A G Q F T S A A F A G Q Q S N S N A
 N T K Q

АМИНОКЛОТНЫЙ СОСТАВ

 * G=ГЛИЦИН = 37 *
 * A=АЛАНИН = 37 *
 * V=ВАЛИН = 21 *
 * L=ЛЕЙЦИН = 32 *
 * I=ИЗОЛЕУЦИН = 26 *
 * F=ФЕНИЛАЛАНИН = 4 *

```

.....
* P=ПРОЛИН                = 9 *
.....
* W=ТРИПТОФАН             = 1 *
.....
* S=СЕРИН                  = 22 *
.....
* T=ТРЕОНИН                = 17 *
.....
* M=МЕТИОНИН              = 12 *
.....
* N=АСПАРГИН              = 21 *
.....
* C=ЦИСТЕИН                = 4 *
.....
* D=АСПАРГИНОВАЯ К-ТА     = 19 *
.....
* Q=ГЛУТАМИН              = 30 *
.....
* E=ГЛУТАМИНОВАЯ К-ТА    = 13 *
.....
* Y=ТИРОЗИН                = 9 *
.....
* H=ГИСТИДИН               = 6 *
.....
* K=ЛИЗИН                  = 25 *
.....
* R=АРГИНИН                = 14 *
.....
*                               ВСЕГО 364 *
*****

```

ТАБЛИЦА ВСТРЕЧАЕМОСТИ КОДОНОВ

```

*****
*   *   T   *   C   *   A   *   G   *
*****
* T   * T   5 * T   12 * T   5 * T   2 *
*   *   *   *   *   *   *   *   *   *
*   *   * C   4 * C   7 * C   4 * C   2 *
*   *   *   *   *   *   *   *   *   *
*   *   * A   14 * A   3 * A   0 * A   0 *
*   *   *   *   *   *   *   *   *   *
*   *   * G   5 * G   0 * G   7 * G   1 *
*****
* C   * T   5 * T   3 * T   4 * T   12 *
*   *   *   *   *   *   *   *   *   *
*   *   * C   0 * C   0 * C   2 * C   2 *
*   *   *   *   *   *   *   *   *   *
*   *   * A   3 * A   5 * A   12 * A   1 *
*   *   *   *   *   *   *   *   *   *
*   *   * G   5 * G   1 * G   1 * G   0 *
*****
* A   * T   21 * T   4 * T   9 * T   5 *
*   *   *   *   *   *   *   *   *   *
*   *   * C   5 * C   5 * C   10 * C   7 *
*   *   *   *   *   *   *   *   *   *
*   *   * A   6 * A   7 * A   22 * A   0 *
*   *   *   *   *   *   *   *   *   *
*   *   * G   12 * G   1 * G   3 * G   0 *
*****
* G   * T   12 * T   16 * T   11 * T   27 *
*   *   *   *   *   *   *   *   *   *
*   *   * C   4 * C   3 * C   14 * C   6 *
*   *   *   *   *   *   *   *   *   *

```

```

*      * A      2 * A      14 * A      24 * A      4 *
*      *.....*
*      * G      3 * G      4 * G      6 * G      1 *
*****

```

```

** 2 ** С ПОЗИЦИИ 132- 1196
МОЖЕТ БЫТЬ СЧИТАН БЕЛОК МОЛЕКУЛЯРНОГО ВЕСА 38161,64
С ЗАРЯДОМ 7,62 ПРИ РН= 7,50
ИЗОЭЛЕКТРИЧЕСКАЯ ТОЧКА БЕЛКА = 8,871

```

```

V A I N Q D K E K A L A A A L G E I Q K
E F G K G S I M H L G Q N R S M D V Q T
I S T G S L S L N V A L G A G G L P R G
R I V Q I Y G P Q S S G K T Y L T L E V
I A S A E H Q G K I C A F I N A Q H A L
N P I Y A E K L G V N I N D L L C S E P
N T G Q E A L Q I C N A L S R S G A V N
V I V V N S V A A L T P K A Q I Q G Q I
G N S H V G L A A R M M S E A M R K L A
G D L K D S D T L L I F I D F I R M K I
G V M F G D P Q T T T G G O A L K F Y A
S V R L N I R R I G S V K D G N Q V I G
S Q T H V K V V K Q K V A A P F K F A Q
F E I M Y G Q G I O T Y G Q L I N L G V
K H K L V Q K A G A W Y S Y D G Q K I G
E G K A D A T D Y L K Q H P Q M Y D Q L
D T K L R Q M L L D H A G Q F T S A A N
F A G Q Q S N S N A N N T K Q

```

АМИНОКИСЛОТНЫЙ СОСТАВ:

```

*****
* G=ГЛИЦИН = 36 *
* A=АЛАНИН = 37 *
* V=ВАЛИН = 21 *
* L=ЛЕИЦИН = 31 *
* I=ИЗОЛЕИЦИН = 20 *
* F=ФЕНИЛАЛАНИН = 9 *
* P=ПРОЛАН = 8 *
* K=ТРИПТОФАН = 1 *
* S=СЕРИН = 22 *
* T=ТРЕОНИН = 17 *
* M=МЕТИОНИН = 11 *
* N=АСПАРГИН = 20 *
* C=ЦИСТЕИН = 3 *
* D=АСПАРГИНОВАЯ К-ТА = 16 *
* Q=ГЛУТАМИН = 29 *
* E=ГЛУТАМИНОВАЯ К-ТА = 13 *
* Y=ТИРОЗИН = 9 *
* H=ГИСТАДИН = 5 *

```

```

*****
* КЛИЗИМ = 25 *
*****
* РАФГИНИ = 14 *
*****
* ВСЕГО 355 *
*****

```

ТАБЛИЦА ВСТРЕЧАЕМОСТИ КОЛОНН

```

*****
* Т С А У *
*****
* Т * Т 5 * Т 12 * Т 5 * Т 2 *
* * * * *
* С * С 4 * С 2 * С 4 * С 1 *
* * * * *
* А * А 14 * А 3 * А 8 * А 0 *
* * * * *
* G * G 5 * G 1 * G 0 * G 1 *
*****
* С * Т 4 * Т 3 * Т 4 * Т 12 *
* * * * *
* С * С 0 * С 0 * С 1 * С 2 *
* * * * *
* А * А 3 * А 4 * А 12 * А 0 *
* * * * *
* G * G 5 * G 1 * G 1 * G 0 *
*****
* А * Т 21 * Т 4 * Т 9 * Т 3 *
* * * * *
* С * С 5 * С 5 * С 9 * С 2 *
* * * * *
* А * А 0 * А 7 * А 22 * А 0 *
* * * * *
* G * G 11 * G 1 * G 3 * G 0 *
*****
* G * Т 12 * Т 16 * Т 16 * Т 27 *
* * * * *
* С * С 4 * С 3 * С 10 * С 6 *
* * * * *
* А * А 2 * А 14 * А 24 * А 3 *
* * * * *
* G * G 3 * G 4 * G 5 * G 0 *
*****

```

НАЙДЕНО 2 ОТКРЫТЫХ РАМОК СЧИТЫВАНИЯ

*** КОМПЛЕМЕНТАРНАЯ НИТЬ ***

НАЙДЕНО 0 ОТКРЫТЫХ РАМОК СЧИТЫВАНИЯ
 ПРОГРАММА ТРАНСЛЯЦИИ КОНЧИЛА РАБОТУ
 ПРОТОКОЛ ОБРАБОТКИ СОХРАНИТЬ? (ДА/НЕТ)

- ДА
- ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТИ
- ПЕЧАТЬ = ПЕЧАТЬ ПОСЛЕДОВАТЕЛЬНОСТИ
 - ДИСПЛЕЙ = ПРОСМОТР ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
 - ОПИСАНИЕ = ПРОСМОТР ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ
 - ФОРМИРОВАНИЕ = ФОРМИРОВАНИЕ НОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗ ФРАГМЕНТОВ
 - ГОМОЛОГИЯ = СРАВНЕНИЕ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
 - СТРУКТУРА = ОБРАБОТКА ПЕРВИЧНОЙ СТРУКТУРЫ
 - УЧАСТКИ = ПОИСК УЧАСТКОВ
 - РЕСТРИКЦИИ = ПРОГРАММА РЕСТРИКЦИИ
 - ТРАНСЛЯЦИЯ = ПРОГРАММА ТРАНСЛЯЦИИ
 - КОНЕЦ = КОНЧИТЬ ОБРАБОТКУ ПОСЛЕДОВАТЕЛЬНОСТИ
- НАЗВАНИЕ РЕЖИМОВ МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
- 0

ВЫ ВЫБРАЛИ ФОРМИРОВАНИЕ РАБОТАЕТ ПРОГРАММА ФОРМИРОВАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗ ФРАГМЕНТОВ
 ОНА ПОЗВОЛИТ ВАМ СФОРМИРОВАТЬ ПОСЛЕДОВАТЕЛЬНОСТЬ ИЗ УКАЗАННЫХ ВАМИ ФРАГМЕНТОВ
 ПОСЛЕДОВАТЕЛЬНОСТЕЙ, СОДЕРЖАЩИХСЯ В БАНКЕ
 ПРОГРАММА ЗАКАНЧИВАЕТ РАБОТУ, КОГДА В КАЧЕСТВЕ СЛЕДУЮЩЕГО ИДЕНТИФИКАТОРА
 УКАЗЫВАЕТСЯ ***
 УКАЖИТЕ ИДЕНТИФИКАТОР ПОСЛЕДОВАТЕЛЬНОСТИ ИЛИ ***
 ЕСРБСА
 ДЛИНА ПОСЛЕДОВАТЕЛЬНОСТИ 1398 В.Р.
 ВЫ БУДЕТЕ РАБОТАТЬ СО ВСЕЙ ПОСЛЕДОВАТЕЛЬНОСТЬЮ? (ДА/НЕТ)
 НЕТ
 ВВЕДИТЕ НАЧАЛО УЧАСТКА
 1
 ВВЕДИТЕ КОНЕЦ УЧАСТКА
 100
 УКАЖИТЕ ИДЕНТИФИКАТОР ПОСЛЕДОВАТЕЛЬНОСТИ ИЛИ ***
 ХИМРА
 ДЛИНА ПОСЛЕДОВАТЕЛЬНОСТИ 3182 В.Р.
 ВЫ БУДЕТЕ РАБОТАТЬ СО ВСЕЙ ПОСЛЕДОВАТЕЛЬНОСТЬЮ? (ДА/НЕТ)
 НЕТ
 ВВЕДИТЕ НАЧАЛО УЧАСТКА
 1
 ВВЕДИТЕ КОНЕЦ УЧАСТКА
 100
 УКАЖИТЕ ИДЕНТИФИКАТОР ПОСЛЕДОВАТЕЛЬНОСТИ ИЛИ ***

 ПРОГРАММА ФОРМИРОВАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ КОНЧИЛА РАБОТУ
 ПРОТОКОЛ ОБРАБОТКИ СОХРАНИТЬ? (ДА/НЕТ)
 ДА
 ПРЕДЛАГАЕМ ВАМ СЛЕДУЮЩИЕ ВИДЫ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТИ:
 ПЕЧАТЬ • ПЕЧАТЬ ПОСЛЕДОВАТЕЛЬНОСТИ
 ДИСПЛЕЙ • ПРОСМОТР ПОСЛЕДОВАТЕЛЬНОСТИ НА ДИСПЛЕЕ
 ОПИСАНИЕ • ПРОСМОТР ОПИСАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ
 ФОРМИРОВАНИЕ • ФОРМИРОВАНИЕ НОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ИЗ ФРАГМЕНТОВ
 ЛОГИКА • СРАВНЕНИЕ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
 СТРУКТУРА • ОБРАБОТКА ПЕРИОДИЧЕСКОЙ СТРУКТУРЫ
 УЧАСТКИ • ИД.СН. УЧАСТКОВ
 РЕСТРИКЦИИ • ПРОГРАММА РЕСТРИКЦИЙ
 ТРАНСЛЯЦИИ • ПРОГРАММА ТРАНСЛЯЦИИ
 КОНЕЦ • КОНЧИТЬ ОБРАБОТКУ ПОСЛЕДОВАТЕЛЬНОСТИ
 НАЗВАНИЕ РЕЖИМА МОЖНО СОКРАЩАТЬ ДО 1-ГО СИМВОЛА
 К
 ВЫ ВЫБРАЛИ КОНЕЦ
 УКАЖИТЕ ИДЕНТИФИКАТОР ПОСЛЕДОВАТЕЛЬНОСТИ ИЛИ ***

 НУЖНО ЛИ НАПЕЧАТАТЬ ПРОТОКОЛ ОБРАБОТКИ? (ДА/НЕТ)
 ДА
 НУЖНО ЛИ НАПЕЧАТАТЬ ГРАФИЧЕСКИЙ ПРОТОКОЛ? (ДА/НЕТ)
 ДА
 ПРОГРАММА ОБРАБОТКИ СИСТЕМЫ МЕ10 КОНЧИЛА РАБОТУ

СОДЕРЖАНИЕ

1. ВВЕДЕНИЕ	3
2. КОНЦЕПЦИИ И ВОЗМОЖНОСТИ	4
2.1. Общее описание	4
2.2. Организация и преобразование данных в системе.	5
2.3. Возможности диалога	6
3. ПОИСК	7
3.1. Структура последовательности в Гейдельбергском банке	7
3.2. Возможности поиска	8
4. ОБРАБОТКА	9
4.1. Общее описание	9
4.2. Печать последовательности	10
4.3. Просмотр последовательности на дисплее	10
4.4. Просмотр описания последовательности	11
4.5. Формирование новой последовательности из фрагментов	11
4.6. Сравнение двух последовательностей	11
4.7. Обработка первичной структуры	11
4.7.1. Общее описание	11
4.7.2. Подсчет плотности распределения нуклеотидов и пар нуклеотидов	12
4.7.3. Построение таблицы ближайших соседей	12
4.7.4. Подсчет сочетаний нуклеотидов по два и по три.	13
4.8. Поиск участков на последовательности	13
4.8.1. Общее описание	13
4.8.2. Поиск участков данного вида	13
4.8.3. Поиск участков со слабой гомологией	14
4.8.4. Поиск прямых повторов	14
4.8.5. Поиск инвертированных повторов	14
4.8.6. Поиск участков двойной симметрии ("шпилек")	15
4.9. Поиски рестриктов при совместной рестрикции	15

4.10. Поиск открытых рамок считывания	17
5. ИНСТРУКЦИЯ ПО ЭКСПЛУАТАЦИИ	19
5.1. Вид и восстановление передаваемой информации . .	19
5.2. Как работать с последовательностями, содержа- щимися в банке	20
5.3. Как работать с собственными последовательно- стями	20
5.4. Как перейти к новой версии банка	20
ЛИТЕРАТУРА	21
ПРИЛОЖЕНИЕ 1. Сеанс поиска	22
ПРИЛОЖЕНИЕ 2. Сеанс обработки	24

Наталья Леонидовна Лунина

СИСТЕМА ОБРАБОТКИ НУКЛЕОТИДНЫХ
ПОСЛЕДОВАТЕЛЬНОСТЕЙ НЕИД. МАТЕРИАЛЫ ПО
МАТЕМАТИЧЕСКОМУ ОБЕСПЕЧЕНИЮ ЭВМ. Выпуск 9

Отпечатано и подготовлено к печати в Отделе
научно-технической информации НЦБИ

Редактор С.Я.Гудкова
Технический редактор С.М.Ткачук
Корректоры Т.К.Крейцер, Л.М.Орлова

Подписано в печать 11.11.84 г. Т18092
Уч.-изд.л. 3,0. Формат 60х90/16. Тираж 420 экз.
Бумага офсетная. Заказ 5738Р. Цена 40 к. Изд. № 272

Отпечатано на ротапринте в Отделе научно-технической
информации Научного центра биологических исследований
АН СССР в Пушкине