

Реализация цепочки ДНК на языке программирования Python

Кизянов Антон Олегович

Приамурский государственный университет имени Шолом-Алейхема

Студент

Аннотация

В данной статье описан процесс создания анимированной цепочки построения ДНК. Для создания используется язык программирования Python и его стандартная библиотека. Созданное приложение служит наглядным пояснением как связаны элементы цепочки ДНК.

Ключевые слова: Python, ДНК

Implementing a DNA strand in the Python programming language

Kizyanov Anton Olegovich

Sholom-Aleichem Priamursky State University

Student

Abstract

This article describes the process of creating an animated DNA building chain. For creation, the Python programming language and its standard library are used. The created application serves as a visual explanation of how the elements of the DNA chain are connected.

Keywords: Python, DNA

Дезоксирибонуклеиновая кислота — это молекула, которая существует в каждой клетке тела человека и содержит схему его роста. Это похоже на двойную спираль (своего рода витую лестницу) пар нуклеотидных молекул: гуанина, цитозина, аденина и тимина. Они представлены буквами G, C, A и T. ДНК — это длинная молекула; она микроскопическая, но, если бы его растянули, его 3 миллиарда пар оснований составили бы 2 метра! Эта программа представляет собой простую анимацию ДНК.

Цель исследования – написать приложение для анимации построения цепочки ДНК на языке программирования Python.

Ранее этим вопросом интересовались И.В. Степанян, М.Ю. Леднев развивали тему «Алгоритмы визуализации молекулярно-генетических последовательностей в пространствах двоично-ортогональных функций уолша» [1] в которой представили систему алгоритмических методов визуализации ДНК и РНК. Показано, что генетические нуклеотидные последовательности и тексты на естественном языке имеют определенные конфигурации своих визуальных представлений в пространствах двоично-ортогональных функций Уолша. Общим свойством этих представлений

является квазифрактальность. Монография представляет интерес для студентов, аспирантов, тех кто работает в рамках получения научной степени доктора наук, для специалистов в области молекулярного генетического кодирования, а также всех интересующихся такими дисциплинами, как математическое и физическое моделирование, молекулярная генетика, дискретная математика, теоретическая биология, параллельные вычисления, теория автоматов, когнитивная эргономика, психофизиология и оптимизация умственного труда, математическая генетика, алгоритмы визуализации данных. И.В. Степанян, М.Ю. Леднев с темой «Обзор некоторых возможностей генометрических информационных систем» [2], а подробнее про актуальность темы исследования обусловлена развитием методов матричной генетики и возникшего из нее генометрического подхода. В статье представлены результаты объединения этих разработок в едином программном комплексе в виде генометрической информационной системы для целей практического использования. Разработанная информационная система реализована на языке C++14 с использованием библиотеки Qt. В программном комплексе реализована параметризация следующих переменных: параметр масштаба N, смещение относительно начала последовательности, шаг перекрытия. Реализованы следующие системы суб-алфавитов: нуклеотидная, динуклеотидная. Реализованы типы визуализации: структурная, частотная, интегральная. Реализованы пространства визуализации: одномерные, двумерные, трехмерные. В реализациях алгоритмов были применены методы параллельных вычислений. Представлено введение нового параметра - шага перекрытия, позволяющего получать более четкие графические отображения. А.С. Полякова опубликовала статью «Будущее визуализации научных данных с использованием языка программирования python» [3] в статье демонстрируется применение высокоуровневого языка программирования Python часто применяющегося для повышения производительности научных приложений и широко используется в научных кругах, т.к. синтаксис Python минималистичен, а стандартная библиотека включает большой объем полезных функций. Библиотека Matplotlib позволяет получать компьютерную графику профессионального качества, которая впоследствии может более достоверно передать различные процессы и явления, облегчая восприятие визуальной информации.

Программа создает анимацию прокрутки путем печати строк из списка ROWS. Пары AT и CG вставляются в каждую строку с помощью строкового метода format().

```
import random
import sys
import time
```

```
PAUSE = 0.3
```

```
rows = [
```

```

'          ##',
'          #{}-{}#',
'          #{}---{}#',
'          #{}-----{}#',
'          #{}-----{}#',
'          #{}-----{}#',
'          #{}-----{}#',
'          #{}---{}#',
'          #{}-{}#',
'          ##',
'          #{}-{}#',
'          #{}---{}#',
'          #{}-----{}#',
'          #{}-----{}#',
'          #{}-----{}#',
'          #{}-----{}#',
'          #{}---{}#',
'          #{}-{}#']
try:
    i = 0

    while True:
        i = i + 1
        if i == len(rows):
            i = 0

        if i == 0 or i == 9:
            print(rows[i])
            continue

        Select = random.randint(1, 4)
        if Select == 1:
            left, right = 'A', 'T'
        elif Select == 2:
            left, right = 'T', 'A'
        elif Select == 3:
            left, right = 'C', 'G'
        elif Select == 4:
            left, right = 'G', 'C'

        print(rows[i].format(left, right))
        time.sleep(PAUSE)
except KeyboardInterrupt:
    sys.exit()

```

Результат работы можно увидеть на рисунке 1.

```

#C-----G#
#T---A#
#A-T#
##
#A-T#
#A---T#
#G-----C#
#C-----G#
#T-----A#
#C-----G#
#T---A#
#A-T#
##
#T-A#
#C---G#
#T-----A#
#A-----T#
#G-----C#
#A-----T#
#G---C#
#T-A#
##

```

Рис. 1 Бесконечная цепочка построения ДНК

Вывод

В этой статье было приведено описание принципа построения цепочки ДНК и реализовано приложения с его анимированным строением. Приложение позволяет лучше разобраться в строении ДНК благодаря простоте и анимированной составляющей.

Библиографический список

1. Степанян И.В., Леднев М.Ю. Алгоритмы визуализации молекулярно-генетических последовательностей в пространствах двоично-ортогональных функций Уолша. М.: КДУ, Добросвет, 2020. URL: <https://www.elibrary.ru/item.asp?id=44848139> (Дата обращения: 05.01.2022)
2. Степанян И.В., Леднев М.Ю. Обзор некоторых возможностей генометрических информационных систем // В сборнике: XXXII Международная инновационная конференция молодых ученых и студентов по проблемам машиноведения. Сборник трудов конференции. 2021. С. 393-400. URL: <https://www.elibrary.ru/item.asp?id=44746209> (Дата обращения: 05.01.2022)
3. Полякова А.С. Будущее визуализации научных данных с использованием языка программирование python // В сборнике: современные технологии: актуальные вопросы, достижения и инновации. сборник статей XXVII Международной научно-практической конференции. 2019. С. 47-49. URL: <https://www.elibrary.ru/item.asp?id=37315684> (Дата обращения: 05.01.2022)