Programación de la técnica CRISPR-Cas9 en Python

Ma Paula Jiménez Torres -C04095

CRISPR-Cas9 es una técnica de modificación genética que consiste en la programación de la proteína Cas9 para editar a conveniencia el ADN de una célula. El ADN es la molécula del código genético, donde se encuentran las instrucciones para todo el funcionamiento del ser que las porta. El ADN está codificado en 4 nucleótidos, en los que para escribir las secuencias se utilizan solo sus iniciales que son A, T, G y C. Cada instrucción es un gen y cada gen es una secuencia de combinaciones de estas 4 letras. Entonces al utilizar CRISPR-Cas9 para modificar el ADN, se modifican las instrucciones y con ello al individuo. El objetivo del proyecto es programar en Python el proceso de la edición genética con el método CRISPR-Cas9.

Primero para poner en contexto se explica cómo funciona esta técnica. El método CRISPR es un mecanismo de defensa contra los virus que tienen ciertos tipos de bacterias. Cuando un virus ataca a una bacteria este introduce su material genético en ella, si la bacteria sobrevive al ataque guarda un trozo de la secuencia de ADN del virus para hacer una copia en ARN y dársela a la proteína Cas9. El ARN dado va a servir como guía para Cas9 que si encuentra una secuencia complementaria esta cortará el ADN del virus así inutilizando.

El método CRISPR-Cas9 consiste en programar a la proteína Cas9 para cortar en la secuencia de ADN que se quiera. Al darle un ARN guía a la proteína esta buscará una coincidencia y cortará el ADN cuando lo encuentre. Así se puede modificar a conveniencia el genoma de una célula. Este método se usa sobre todo para inhabilitar genes o para agregar nuevas secuencias de ADN en un lugar específico del genoma. Cuando se corta el genoma y se le introduce una secuencia “plantilla” de ADN a la célula, la plantilla debe tener sus extremos iguales a los extremos que resultaron al corte del genoma, para que la célula cuando intente repararse tomara de ejemplo la plantilla y así agregando la secuencia que se quería.

**Explicación del código**

Lo que se quiere conseguir es poder modificar una secuencia de ADN agregándole una nueva combinación de letras deseadas en el lugar deseado de la secuencia, pero siempre tratando de imitar lo más posible el proceso real de CRISPR.

Utilizando la programación orientada a objetos se crearon 2 clases con los siguientes atributos y métodos.

|  |  |
| --- | --- |
| Cas9 | Celula |
| * aminoacidos * arn\_guia | * genoma * adn\_extra * proteína\_extra |
| * cortar | * adn\_extra.setter * proteina\_extra.setter |

Los atributos arn\_guia, adn\_extra y proteína\_extra no son necesarios para la creación de los objetos entonces vienen con valor nulos y se solo se cambia con el setter con el fin de imitar la realidad. Todos los atributos de las clases tienen los métodos get y set, pero en el diagrama se quiso enfatizar en los setter de adn\_extra y proteína\_extra ya que en estos se agregaron ciertos procesos con el fin de imitar que apenas se le agrega la cas9 o la plantilla la célula va a reaccionar a estos cambios.

La Cas9 tiene una secuencia específica de aminoacidos y solo se puede crear un objeto Cas9 si tiene esa secuencia exacta. Se descargó la secuencia de ADN que forma a la proteína del banco de genes NIH en formato FASTA (formato especializado para las secuencias de ADN, utilizando por genetistas, microbiólogos, etc). Para abrir en Python un archivo tipo FASTA se utilizó la librería Biopython y también se utilizó para pasar de la secuencia de ADN a la de aminoacidos usando las funciones transcribe() y translate().

El proceso para modificar una secuencia de ADN utilizando el código es el siguiente:

* Se crea un objeto Cas9 utilizando la secuencia especifica de aminoacidos
* Se crea un objeto Celula utilizando el genoma que se quiera editar
* Se introduce la secuencia arn\_guia al objeto Cas9, para que se haga un corte en el lugar deseado
* Se introduce el objeto Cas9 a el objeto Celula mediante el setter de proteína\_extra y en la que automáticamente si se detecta que es una Cas9 entonces Cas9 cortará el genoma de la célula
* Se introduce la secuencia de ADN plantilla a la Celula y esta con el setter de adn\_extra automáticamente reparará su genoma utilizando de ejemplo la plantilla

Con el código creado se puede editar una secuencia de ADN, pero cabe aclarar que el código se centra en la imitación del proceso natural que pasa con CRISPR-Cas9, por lo que hay maneras más optimas y sencillas de editar una secuencia en Python. Una limitación del código es que, por temas de falta de recursos tecnológicos al momento de hacer el proyecto, por una falta de potencia no se pudo hacer pruebas con secuencias muy largas, pero se considera que el proyecto cumple con los resultados objetivo que se querían obtener.

Se recomienda la aplicación de la programación orientada a objetos para proyectos relacionados a procesos biológicos porque al momento de querer imitar fielmente el proceso se considera que es la manera más análoga. También se recomienda la librería Biopython para proyectos similares por su optimización en las secuencias genéticas.

**Bibliografía**

NCBI. (2024). cas9 type II CRISPR RNA-guided endonuclease Cas9 [Streptococcus pyogenes] <https://www.ncbi.nlm.nih.gov/gene/69900935>

Wellcome Genome Campus. (n.d.). What is CRISPR-Cas9?. YourGenome. <https://www.yourgenome.org/theme/what-is-crispr-cas9/>

Innovative Genomics Institute. (n.d.). What is CRISPR?. <https://innovativegenomics.org/what-is-crispr/>