**Genómica Computacional**

**Proyecto final**

**Programa para el diseño de primers para PCR**

**Paloma Abigail Ramírez Pizaña**

La replicación del DNA es el proceso por el cual se obtiene una molécula duplicada de DNA. Ocurre anterior a la duplicación celular, pues ésta debe duplicar su genoma para que cada célula hija contenga un juego completo e idéntico de material genético.

La replicación comienza siempre en puntos definidos del DNA llamados orígenes de replicación, que son secuencias concretas de DNA en las que se específicamente se inicia el proceso. En los tres dominios en los que dividimos las formas de vida, existen orígenes de replicación diferentes y el inicio de la replicación se da de diferente manera para los tres; por ejemplo, las bacterias presentan sólo un origen de replicación, mientras que en arqueas y en eucariontes suelen haber diferentes lugares de replicación n eucariotas, como es el caso de los humanos, encontraremos más de un origen de replicación.

Una de las moléculas involucradas en la replicación del DNA es la DNA polimerasa (DNA pol). Estas enzimas son las responsables de la síntesis del DNA, ya que son las que añaden nucleótido por nucleótido a la cadena naciente de forma complementaria al molde. Ya que las polimerasas son incapaces de catalizar la replicación de una cadena de DNA *de novo*, necesitan de una guía a partir de la cual puedan empezar a añadir nucleótidos sobre una hebra preexistente. Estas guías llamadas primer o cebador son cadena sencillas y cortas de DNA con un hidroxilo 3’ libre que sirve como punto de partida o iniciador en la replicación del DNA. Su estructura y disposición propicia la elongación de una nueva cadena de DNA nueva formando pares de bases complementarios a una hebra molde.

Cuando se inicia la replicación en el sitio de origen se forma una horquilla de replicación que separa las dos hebras que además se encuentran en antisentido una con respecto de la otra, entonces se da la síntesis de ambas cadenas, por lo tanto, también se necesitan dos primers; uno para polimerizar la hebra positiva o sentido (5’-3’), llamado primer forward y otro para la hebra negativa o antisentido (3’-5’), llamado primer reverse. Es importante destacar que, la cadena sintetizada a partir del molde antisentido, su síntesis también ocurre en sentido 5’-3’ a partir del origen, aunque lo hace en pequeños fragmentos que posteriormente serán unidos por una enzima ligasa, a diferencia de la cadena sentido que se elonga sin parar y sin cortes.

Actualmente, muchas técnicas de estudio de laboratorio en biología molecular utilizan las bases de la replicación; tales como la secuenciación de DNA y la reacción en cadena de la polimerasa (PCR). Esta última es una técnica de síntesis in vitro de secuencias específicas, y es por excelencia la forma más simple y rápida de multiplicar de forma masiva el DNA.

Al igual que en la replicación, la PCR necesita de los primers que son diseñados de manera que sean complementarios a cada uno de los extremos 3’ del fragmento de DNA que se desea amplificar.

Esta técnica además se basa en la repetición de un ciclo de tres etapas:

1. Desnaturalización: donde la doble hélice de DNA se separa en dos hebras y para ello se realiza una incubación de la muestra a altas temperaturas (93-97 °C). Finalmente, al disminuir la temperatura se dará la renaturalización.

2. Hibridación: en esta etapa, los cebadores se unen a las zonas 3’ complementarias que flanquean el fragmento que queremos amplificar. Se realiza gracias a la disminución de la temperatura (50-65 °C).

3. Extensión: se produce la síntesis de una cadena sencilla por cada una de las dos hebras molde.

El diseño y selección de los primers utilizados en PCR es uno de los pasos clave para el éxito de esta técnica, por ello, existen algunas consideraciones a tomar en cuenta para la construcción de dichos oligonucleótidos.

Para esta técnica, los primers deben ser moléculas de DNA cortas y sintetizadas de forma química, con un rango de longitud de 18-25 bases, ya que con oligonucleótidos de mayor longitud, disminuye su eficiencia y a menor longitud, la especificidad se ve comprometida.

Además se recomienda que ambos primers (forward y reverse) tengan temperaturas de fusión (Tm) parecidas y dentro de un rango de 55-65 °C para evitar perder eficiencia en la amplificación.

Los oligonucleótidos deberán contener secuencias con un contenido GC entre 45 a 55%, de manera contraria, al tener una secuencia de poliC o poliG, los primers se vuelven menos eficientes.

Por otra parte, se debe evitar que las secuencias contengan *hairpin loop* que son

estructuras que interfieren en el alineamiento formando cadenas dobles de forma parcial. Así como evitar la autocomplementariedad para evitar formar dímeros del primer o estructuras secundarias.

Finalmente, las secuencias de los primer deben ser complementarias en un 100% a la cadena molde que se quiere amplificar y se recomienda que se encuentren a 35-50 bases de distancia de la región a amplificar.

Aunque ya existen múltiples programas que calculan y diseñan los primers cómorimer3 y primer teniendo en cuenta las consideraciones anteriores y evitando el trabajo de hacer el diseño a mano, ahorrando tiempo y agilizando el proceso; el presente proyecto intenta abordar el tema para presentar una propuesta de programa bioinformático que sea capaz de diseñar los primers de una secuencia dada por el usuario.

Si bien el enfoque puede verse limitado en la parte de programación, se intenta aplicar los conocimientos adquiridos en la materia de Genómica Computacional, en conjunto con mi formación como bióloga, para intentar hacer un ejercicio en que ofrezca una solución práctica para diseñar primers que puedan ser utilizados para PCR.

Objetivo

* Crear un programa bioinformático que permita al usuario realizar de forma rápida y efectiva el diseño de primers para su aplicación en técnicas de PCR.

Objetivos particulares

* Realizar una investigación bibliográfica para conocer las características que debe cumplir un primer utilizado en PCR.
* Crear un script que permita al usuario diseñar los oligonucleótidos forward y reverse para usar como primer en PCR.

Metodología

Inicialmente se realizó una búsqueda en diferentes fuentes bibliográficas para conocer las características que deben cumplir los oligonucleótidos al ser diseñados para PCR, usando palabras en inglés como: “primer design”. Los resultados se filtraron por fecha de publicación descartando los artículos científicos publicados antes de 2018.

Posteriormente, para la realización del programa se utilizó el cuaderno virtual Google Colab (https://colab.research.google.com). El código para realizar el programa se escribió en el lenguaje de Python.

Tomando en cuenta los parámetros consultados en la literatura, para este ejercicio se establecieron los mostrados en la Tabla 1.

*Tabla 1.*  Parámetros utilizados para el diseño de primers forward y reverse de una cadena de DNA molde para la reacción en cadena de la polimerasa (PCR).

|  |  |
| --- | --- |
| Característica | Parámetros |
| Longitud (nucleótidos) | 21 |
| Extremo 3’ | Guanina o citocina (G o C) |
| Contenido GC | 40-60% |
| Temperatura de fusión | 55-65 °C |

Se utilizaron dos librerías: tabulate, para regresar al usuario la información sobre los primer diseñados en formato de tabla, y pydna.amplify para crear un amplicón de referencia para asegurar que los primers fueron diseñados de manera correcta.

Resultados