

Mapa de Restricción

En 1970 Hamilton Smith descubrió que la enzima de restricción HindII que parte las moléculas del ADN para cada ocurrencia, sitio o secuencia GTGCAC ó GTTAAC, partiendo una molécula larga en un conjunto de fragmentos de restricción.

Si se conoce la secuencia de ADN genómico de un organismo, entonces la construcción de un mapa de restricción para HindII equivale a encontrar todas las ocurrencias de GTGCAC y GTTAAC que se encuentran en el genoma. El primer genoma bacteriano fue secuenciado veinticinco años después del descubrimiento de las enzimas de restricción, por varios años biólogos se forzaron a construir mapas de restricción para los genomas sin conocimientos previos de la secuencia de ADN de los genomas.

Las enzimas de restricción son conocidas como endonucleasas, son enzimas que cortan los enlaces fosfodiéster del material genético a partir de una secuencia que reconocen. Las mismas permiten cortar ADN hebra doble, en donde se reconocen secuencias palindrómicas (secuencias que se leen de igual forma en ambas direcciones).

Son extraídas de organismos procarióticos (bacterias), donde actúan como un mecanismo de defensa, para degradar material genético extraño que entra en la célula. Las bacterias tienen la capacidad de metilar su ADN, lo cual sirve para distinguir entre el ADN extraño y el ADN propio. Las enzimas de restricción no pueden cortar ADN metilado, de este modo solo afectan el ADN extranjero y no el ADN bacteriano.

La enzima de restricción tipo II mejor conocido como hindII posee las siguientes características:

- Sólo tienen actividad de restricción.
- Cortan de manera consistente y predecible dentro de la secuencia que reconocen.
- Sólo requieren Mg^{++} como cofactor.
- No necesitan ATP

Algunas Aplicaciones de las enzimas de restricción:

1. Hacer mapa de restricción de un plásmido o bacteriófago.
2. Fragmentar DNA genómico para separación de electroforesis y "Southern Blot".
3. Generación de fragmentos para ser usados como sondas marcadas en "Southern" y "Northern" blotting.
4. Generación de fragmentos para ser subclonados en los vectores apropiados, creación de ADN recombinante.

Existe una variedad de enfoques experimentales para el mapeo de restricción, cada uno con ventajas y desventajas. La distancia entre dos restricciones individuales, sitios de codificación corresponde a la longitud del fragmento de restricción entre los dos sitios y puede medirse mediante la técnica de electroforesis en gel. Esto no requiere conocimiento de la secuencia de ADN. Los Biólogos pueden variar las condiciones experimentales para producir un resumen completo o resumen parcial de ADN. El problema del mapeo de restricción puede formularse en términos de recuperación de posiciones de puntos cuando solo las distancias de emparejamiento entre esos puntos son conocidos.

Para formular el problema del mapeo de restricción, se debe introducir algunas notaciones. Un multiconjunto que permite elementos duplicados ({2, 2, 2, 3, 3, 4, 5}, es un conjunto con elementos duplicados 2 y 3). Si $X = \{x_1 = 0, x_2, \dots, x_n\}$ es un conjunto de n puntos en un segmento de línea en orden creciente, entonces ΔX denota el conjunto de todos $\binom{n}{2}$ distancias emparejadas entre los puntos en X .

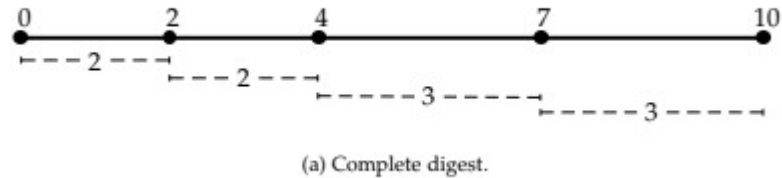
$$\Delta X = \{x_j - x_i: 1 \leq i < j \leq n\}$$

Por ejemplo, si $X = \{0, 2, 4, 7, 10\}$, entonces $\Delta X = \{2, 2, 3, 3, 4, 5, 6, 7, 8, 10\}$, que son las diez distancias por emparejamiento entre estos puntos. En el mapeo de restricción, se nos da ΔX , los datos experimentales sobre la longitud de los fragmentos. El problema es reconstruir X desde ΔX .

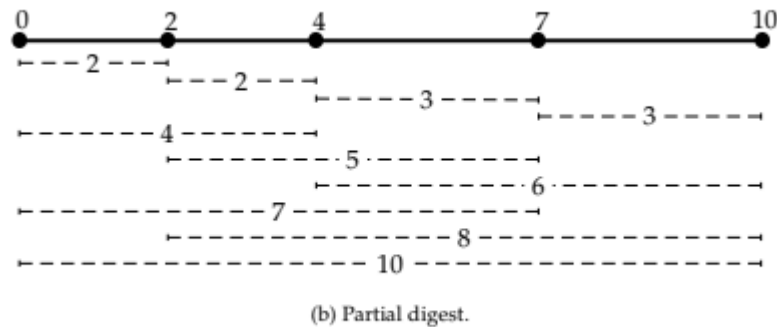
La notación $\binom{n}{k}$, se lee "n combinatoria k", lo cual significa "El número de subconjuntos distintos de k elementos tomado de un conjunto (mayor) de n elementos", y está dado por la expresión $\frac{n!}{(n-k)!k!}$. En particular $\binom{n}{2} = \frac{n(n-1)}{2}$ es el número de diferentes pares de elementos de un conjunto de n elementos.

Por ejemplo, si $n = 5$, el conjunto $\{1, 2, 3, 4, 5\}$ tiene $\binom{5}{2} = 10$ subconjuntos formados por dos elementos: $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{1, 5\}$, $\{2, 3\}$, $\{2, 4\}$, $\{2, 5\}$, $\{3, 4\}$, $\{3, 5\}$ y $\{4, 5\}$. Estos diez subconjuntos dan lugar a la elementos $x_2 - x_1$, $x_3 - x_1$, $x_4 - x_1$, $x_5 - x_1$, $x_3 - x_2$, $x_4 - x_2$, $x_5 - x_2$, $x_4 - x_3$, $x_5 - x_3$, y $x_5 - x_4$.

Digestión Completa produce solo fragmentos entre restricción consecutiva de sitios.



Digestión Parcial produce fragmentos entre cualquiera de los dos sitios de restricción.



El conjunto $\Delta X = \{2, 2, 3, 3, 4, 5, 6, 7, 8, 10\}$ se deriva de $\{0, 2, 4, 7, 10\}$

Aunque la electroforesis en gel permite determinar las longitudes de fragmentos de ADN fácilmente, a menudo es difícil juzgar su multiplicidad. La cantidad de fragmentos diferentes de una determinada longitud puede ser difícil de determinar.

Representación de $\Delta X = \{2, 2, 3, 3, 4, 5, 6, 7, 8, 10\}$ como tabla bidimensional, con los elementos de $X = \{0, 2, 4, 7, 10\}$ a lo largo del lado superior y derecho. los elemento en (i, j) en la tabla es el valor $x_j - x_i$ para $1 \leq i < j \leq n$.

	0	2	4	7	10
0		2	4	7	10
2			2	5	8
4				3	6
7					3
10					

Partial Digest Problem:

A partir de los resultados obtenidos podemos resolver el problema de la digestión parcial, dado todos los emparejamientos entre los puntos en una línea, podemos reconstruir las posiciones de estos puntos.

INPUT: El multiconjunto de emparejamiento de distancias L , contenido $\binom{n}{2}$ de enteros.

OUTPUT: Un conjunto X , de n enteros, tal que $\Delta X = L$

Este problema de resumen parcial, o PDP, a veces se denomina problema de Turnpike en ciencias de la computación, supongamos que conoce el conjunto de distancias entre cada (no necesariamente consecutiva) par de salidas en una carretera que conduce de una pueblo a otro.

¿Se podría reconstruir la geografía de la carretera desde esta información?

Es decir, ¿podría encontrar la distancia desde la primera ciudad a cada salida? Aquí, las "salidas de la autopista" son los sitios de restricciones en el ADN; las longitudes de los fragmentos de restricción de ADN resultantes corresponden a las distancias entre las salidas de la autopista. Computacionalmente, la única diferencia entre el El problema de Turnpike y el PDP es que se dan las distancias entre las salidas en millas en el problema Turnpike, mientras que las distancias entre la restricción los sitios se dan en nucleótidos en el PDP.