

Un modelo matemático para el estudio de las recombinaciones genéticas en ciliados

M.G. Rodríguez Sánchez¹, L.A. García García².

^{1,2} Departamento de Ciencias Básicas, UAM Azcapotzalco, Ciudad de México., México.
rsmg@azc.uam.mx, stokes_1560@hotmail.com



RESUMEN

Texto / Text

El código genético almacenado en el ADN está compuesto por una larga cadena de combinaciones de las letras A, T, G, C, que corresponden a los elementos Adenina, Timina, Guanina y Citosina, respectivamente. Se conocen tres recombinaciones genéticas en el código genético, que inciden en la evolución y sobrevivencia de una especie. Se presenta un modelo matemático para la investigación sobre el ADN. El material aquí presentado, está centrado en el estudio de los ciliados, seres unicelulares cuyas células tienen dos núcleos. Los conceptos matemáticos que se utilizan forman parte de la teoría de delta-matroides.

INTRODUCCIÓN

En 1953 Watson y Crick presentaron el modelo de la doble hélice de la molécula del ADN

↓
Comprensión del código genético



Los **ciliados** son organismos unicelulares con la característica de poseer dos núcleos, un macronúcleo (MAC) y un micronúcleo (MIC). El MAC tiene muchas copias de cromosomas cortos, los cuales tienen uno o pocos genes. El código del MIC es usado para la reproducción y el MAC para la regulación general de la célula. Durante la reproducción sexual el recién formado MIC es transformado en un MAC. Este proceso es llamado **recombinación genética**. Para ejemplificar los procesos en el ADN, se utilizará la proteína Actina I en el ciliado *Sterkiella nova*.

MODELO GENÉTICO

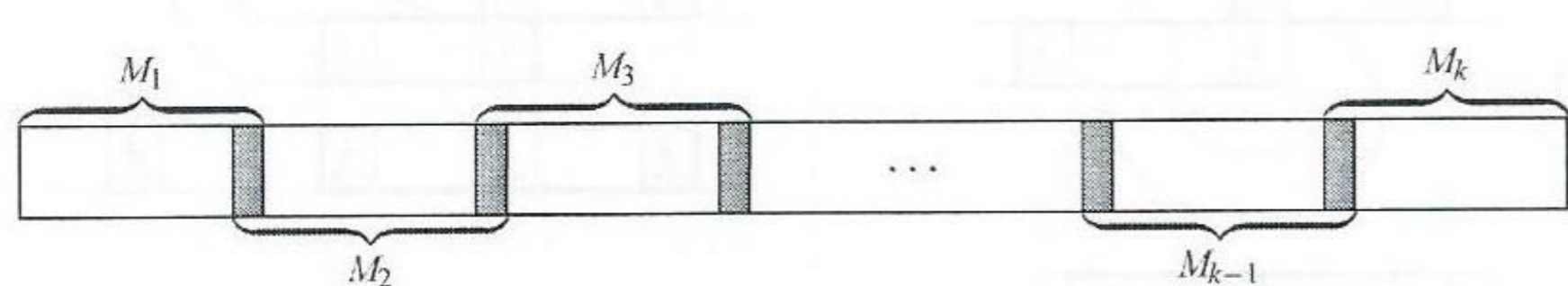
A. Codificación de los genes

Los genes del ADN del MIC están desordenados y posiblemente invertidos respecto al MAC correspondiente, están separados por segmentos cortos de ADN no genético, denominado IES's. Los IES's dividen al gen en fragmentos, llamado MDS (macromolecular destined segments). Así el MIC del gen de la Actina I del ciliado *Sterkiella nova* puede describirse con la siguiente cadena:

$$I_0 M_3 I_1 M_4 I_2 M_6 I_3 M_5 I_4 M_7 I_5 M_9 I_6 M_2 I_7 M_1 I_8 M_8 I_9 \quad (1)$$

$$M_1 M_2 M_3 M_4 M_5 M_6 M_7 M_8 \cdot$$

Los MDS son los M_i 's y los IES son los I_i 's de la cadena. M_2 denota la inversión de M_2 . Entre cada par de MDS's consecutivos (MAC) hay una intersección, estos segmentos se llaman **punteros**. En la figura siguiente, puede observarse la colocación de los punteros en un MAC con k MDS's.



B. Operaciones de recombinación

Formalización: Sea $k \in \mathbb{Z}^+$. Considérese un alfabeto finito Γ cuyas letras son $\{1, 2, \dots, k\} \cup \{\bar{1}, \bar{2}, \dots, \bar{k}\}$ y w una palabra sobre Γ , donde cada letra aparece exactamente dos veces. Se dice que w es una **palabra de doble ocurrencia** sobre Γ [4]. Sean k MDS's en la cadena del MIC y sea $w = w_1 w_2 \dots w_k$; $w_i \in \Gamma$, las letras w_i representan a los punteros.

A M_i le corresponde $i(i+1)$, a su inversión M_i le corresponde $(i+1)\bar{i}$, para $i \in \{2, \dots, k-1\}$. A M_1 y M_k les corresponde solo un puntero.

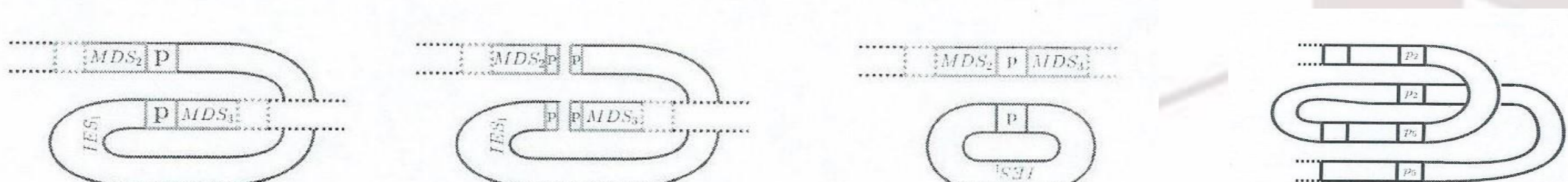
La forma de Actina I de *Sterkiella nova* tiene la cadena de punteros:

$$34 \ 45 \ 67 \ 56 \ 78 \ 9 \ 32 \ 2 \ 89.$$

Operaciones de recombinación para ejecutar la desfragmentación del gen en su forma MIC:

Escisión de lazo

Esta operación elimina un IES entre un par de punteros consecutivos que tienen la misma orientación. Une los fragmentos MDS_2 y MDS_3 y elimina el IES_2 , se produce una molécula circular.

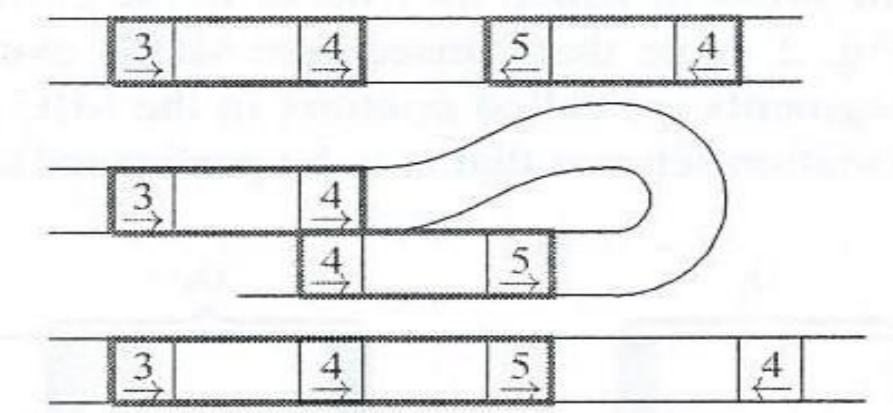


Recombinación de doble lazo

Esta operación de recombinación se aplica a una molécula de ADN con dos pares de punteros cuyos segmentos delimitados por cada par están superpuestos. La molécula se pliega en dos lazos alineando a los dos pares de punteros. Se produce un intercambio de los segmentos superpuestos. (Fig. derecha)

Recombinación de horquilla

Se aplica a una molécula en la que el par de secuencias repetidas tienen orientación opuesta. En este caso la molécula forma un pliegue en forma de horquilla, alineando el par de punteros invirtiendo el segmento entre ellos. Se ejemplifica, con un fragmento de una cadena con cuatro punteros: $p_3 p_4 \bar{p}_5 \bar{p}_4$. En la figura, se usan direcciones para denotar diferentes direcciones.



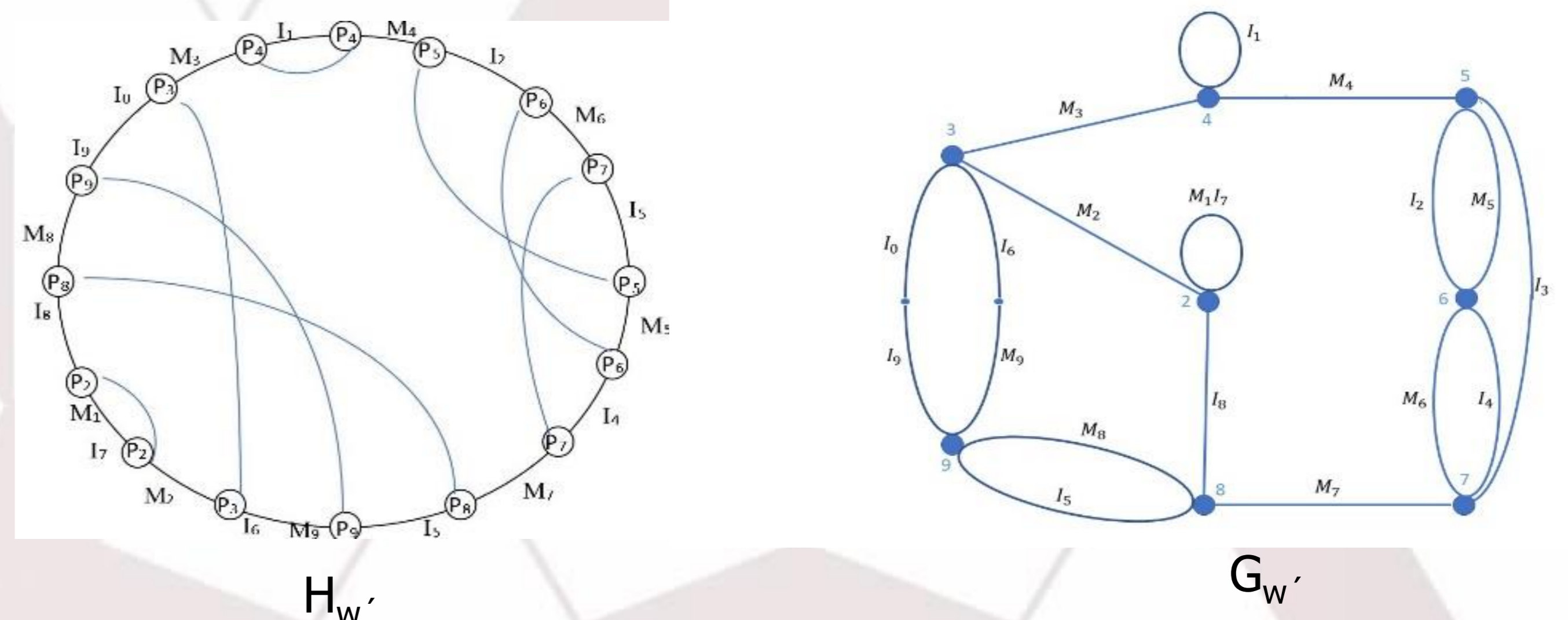
C. Gráficas asociadas al modelo

Se puede expandir la expresión (1) si se agregan los punteros, así la cadena para la *Sterkiella nova* quedaría como:

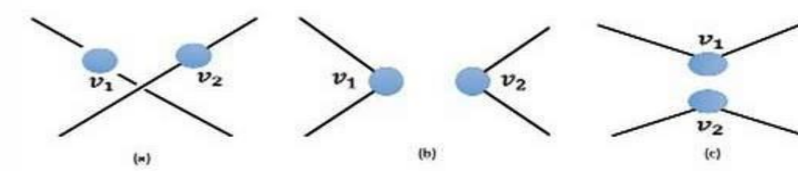
$$w' = I_0 p_3 M_3 p_4 I_1 p_4 M_4 p_5 I_2 p_6 M_6 p_7 I_3 p_5 M_5 p_6 I_4 p_7 M_7 p_8 I_5 p_9 M_9 I_6 p_3 M_2 p_2 I_7 M_1 p_2 I_8 p_8 M_8 p_9 I_9$$

La cadena representada por w' se puede colocar sobre una circunferencia siguiendo el orden de los elementos para construir una gráfica denominada $D_{w'}$, se identifican los punteros con los vértices de una gráfica y los MDS's y IES's como las aristas en la gráfica. Si a $D_{w'}$ se le da una orientación, se obtiene una digráfica, sin las aristas azules. Si se unen los punteros con el mismo índice, por una arista, se obtendrá una gráfica que se denotará como $H_{w'}$.

Si en $H_{w'}$ se contraen las aristas que unen los vértices con punteros con el mismo número, se obtendrá una multigráfica 4-regular $G_{w'}$. En los vértices tienen el índice del puntero que le corresponde.

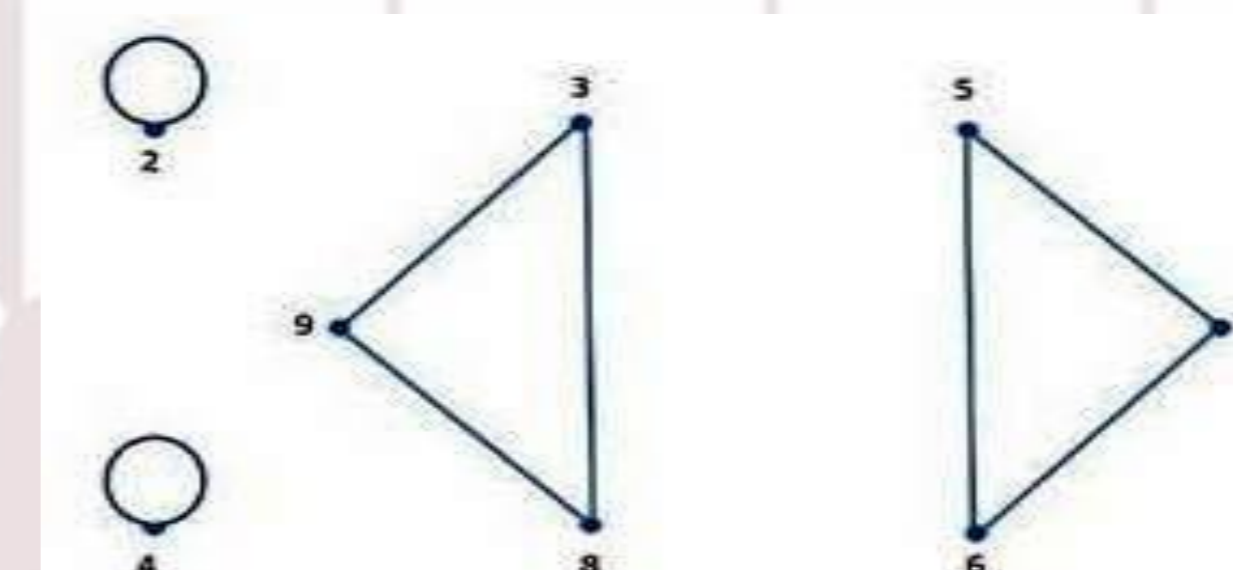


Cada vértice de $G_{w'}$ se divide en dos vértices de grado 2, en una de las formas:



Equivalentemente, se pueden seguir las aristas y vértices de la cadena w' , en la multigráfica $G_{w'}$, se obtiene un circuito Euleriano que corresponde al MIC del ciliado en estudio. Si se eliminan en $G_{w'}$ las aristas etiquetadas con M_i se obtiene una gráfica con tres componentes conexas, que producen una partición de las aristas I_j , estas son $\{I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8, I_9, I_0\}$.

Considérese la gráfica $H_{w'}$, se puede construir su gráfica $G_{w'}^c$ de intersección: por cada arista de $H_{w'}$ se pone un vértice de $G_{w'}^c$ y dos vértices de $G_{w'}^c$ son adyacentes si las aristas correspondientes en $H_{w'}$ se intersectan. Para el caso de estudio, la gráfica $G_{w'}^c$:



CONCLUSIONES

El entendimiento del código genético que se encuentra en el ADN de toda célula es fundamental para conocer el fenómeno de la vida y de su evolución en la Tierra. El modelo matemático basado en tópicos de la teoría de delta-matroides proveen de una herramienta profunda para su comprensión y estudio.

D. Gráficas circulares

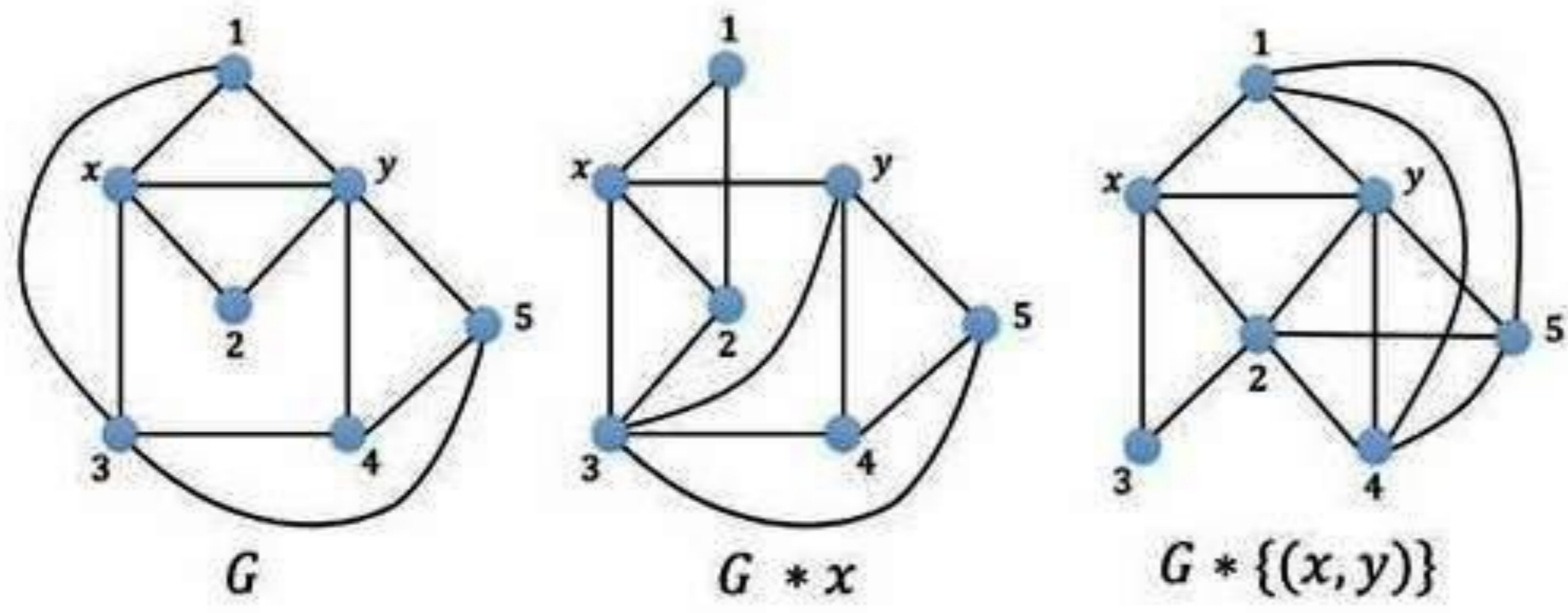
Las **gráficas circulares** son estructuras combinatorias con una teoría propia, desarrollada por Bouchet [6]. Sea G una gráfica $G=(V,E)$, donde V es un conjunto finito de vértices y $E=\{(u,v) \mid u, v \in V\}$ un conjunto de aristas, en G se permiten lazos, pero no aristas múltiples, así $\{v\}$ es un vértice lazo. Para $X \subset V$, $G[X]$ denota a la subgráfica de G inducida por X .

Operaciones en las gráficas circulares:

Para $v \in V$, se considera $N(v)$ la vecindad de v .

1. Complementación local. Se denota por $G * v$ y consiste en reemplazar la subgráfica inducida $G[N(v)]$ por su complemento, las aristas en $V \setminus N(v)$ son las mismas en G y en $G * v$.

2. Complementación de aristas. Se realiza sobre una arista $e=(u,v)$ de G y se denota por $G * \{(u,v)\}$. Para un vértice u se define su vecindad cerrada $N(u)$ como $N(u) \cup \{u\}$. Las vecindades de u y v se pueden partir en tres conjuntos $N(u) \setminus N(v)$, $N(v) \setminus N(u)$ y $N(u) \cap N(v)$. La gráfica $G * \{(u,v)\}$ se obtiene complementando todos los pares (x,y) tales que x y y son cada uno vecinos de u o v , pero que no están en la misma clase de la partición. Así, $G * \{(u,v)\}$ no cambiará las adyacencias entre vértices no adyacentes a u y v , lazos.



En la gráfica circular se consideran vértices con lazos, para modelar la siguiente situación: en una cadena formada por los punteros, se codifica su orientación relativa, para cada par de punteros que tienen diferente orientación p y \bar{p} , se agrega un lazo en el vértice que corresponde a p , en la gráfica circular. A una cadena de código genético le corresponde una palabra de doble ocurrencia.

La operaciones de recombinación vistas en la gráfica G consisten en:

Gráfica circular	Cadena genética
Escisión de lazo	\rightarrow remoción de un vértice aislado sin lazo
$G * v \setminus v$	\rightarrow recombinación de horquilla (v , puntero)
$G * \{(u,v)\} \setminus \{u,v\}$	\rightarrow recombinación de doble lazo

(\setminus operación de borrado; v :puntero p ; u,v punteros p_1 y p_2)

Este trabajo se basa en una serie de investigaciones plasmadas en [2][3][13].

Las cadenas de información del ADN, pueden verse como combinaciones de las cuatro letras de las bases presentes en el ADN,

modelo matemático



para manejar las cadenas y los cambios en los procesos de recombinación genética

La teoría de gráficas circulares, que es parte de la teoría de delta-matroides, desarrollada por Bouchet en múltiples artículos, algunos de ellos [5][6][11], permite construir un modelo fiel para ciliados, pues las operaciones de recombinación que ocurren en su micronúcleo, durante los procesos de replicación, coinciden con las operaciones sobre la gráfica circular que se construye mediante el proceso explicado.

Se puede ampliar el modelo, construyendo la matriz de adyacencia de la gráfica circular G_w^c , e introduciendo el concepto de delta-matride para interpretar las operaciones de recombinación en este contexto.

Ramas de matemáticas involucradas en este modelo:

las palabras de doble ocurrencia, las multigráficas 4-regulares y las gráficas circulares,

Estas forman parte de una teoría más extensa, la teoría de *delta-matroides*.

REFERENCIAS PRINCIPALES

- [3] R. Brijder. Recombination faults in gene assembly in ciliates modeled using multimatroids. Theoretical Computer Science, 608. (2015). pp 27-35.
- [4] A. Kotzig. Eulerian lines in finite 4-valent graphs and their transformations. Theory of graphs, Proceedings of the Colloquium, Tihany, Hungary. (1966) pp 219-230. Academic Press, New York (1968)
- [5] A. Bouchet. Compatible Euler tours and supplementary Eulerian vectors. European journal of combinatorics, vol. 14, no 6, (1993). pp 513-520.
- [6] A. Bouchet. Circle graph obstructions. Journal of Combinatorial Theory, Series B, 60(1), (1994). pp 107-144.
- [7] L. Traldi. Binary matroids and local complementation. European Journal of Combinatorics, 45, (2015) pp 21-40 .
- [9] L. A. García García. Un modelo combinatorio en ensamblamiento genético. Tesis Maestría en Optimización. UAM-A, (2018)
- [10] L. A. García García, M. G. Rodríguez. Cálculo de la distancia de inversión cromosómica usando estructuras matroidales. Memorias de la Sociedad Matemática Mexicana, Vol. 15, (2019). ISSN 1870-2112.
- [11] A. Bouchet. A. Bouchet. k-transformations, local complementations and switching. In Cycles and Rays Springer Netherlands. (1990). pp. 41-50.
- [12] A. Bouchet. Graphical presentations of isotropic systems. Journal of Combinatorial Theory, Series B, 45(1), (1988). pp 58-76.
- [13] S. Hannenhalli, P. A. Pevzner. Transforming cabbage into turnip: polynomial algorithm for sorting signed permutations by reversals. Journal of the ACM (JACM), 46(1), (1999) pp 1-27.