



ESTRUCTURAS ALGEBRAICAS Y SÍNTESIS DE PROTEÍNAS : UNA
DESCRIPCIÓN DEL GEN CAPN10

ALGEBRAIC STRUCTURES AND SYNTHESIS PROTEIN: A DESCRIPTION OF
GENE CAPN10

OBIDIO RUBIO * AND RUTH NORIEGA **

Received, Jan. 10, 2015

Accepted, May. 15, 2016.

Resumen

En este artículo principalmente divulgativo, presentamos algunas formas de la modelación algebraica de las secuencias del Genoma Humano, con especial énfasis en describir mutaciones en los genes, que al modificar la síntesis de proteínas implican enfermedades genéticas como la Diabetes Mellitus. Se describe mutaciones como endomorfismos sobre un R -módulo, el cual está conformado por una suma directa de grupos de secuencias del gen $2q37.3$, sobre los anillos \mathbb{Z}_{64} y \mathbb{Z}_{125} , donde ocurre el haplotipo compuesto por los polimorfismos SNP 43, 19 y 63.

Palabras clave. Genoma humano, genética, Diabetes Mellitus.

Abstract

This article mainly informative, some ways of algebraic modeling of human genome sequences is presented, with special emphasis on describing mutations in the genes, which by modifying protein synthesis involve genetic diseases such as Diabetes Mellitus. Mutations as endomorphisms on an R -module, which consists of a direct sum of groups of sequences $2q37.3$ gene, on the rings \mathbb{Z}_{64} and \mathbb{Z}_{125} , where el haplotype compound for the polymorphisms SNP43, SNP19 and SNP63 occur.

Keywords. Human genome, genetics, Diabetes Mellitus.

1. Introducción. La presente trabajo contiene los aspectos fundamentales de la genética molecular, describiendo el funcionamiento del código genético y su preponderante función en la síntesis de proteínas, para ello se usa el secuenciamiento de del genoma humano; fundamental en el proceso es la estructura del ADN y el ARN, su transcripción y traducción hasta llegar a los aminoácidos. Este proceso es modelado matemáticamente, usando para ello los conceptos de la teoría de grupos abelianos finitos, cuyos endomorfismos representan las transcripciones de genes lo cual puede generar alguna mutación y esta en alguna enfermedad genética, tal es el caso de la Diabetes mellitus de tipo 2 que intentaremos describirlo como caso especial.

Las relaciones cuantitativas entre codones y entre genes dadas en los procesos de evolución molecular son un sorprendente desafío a la biología matemática[6].

Muchos intentos han sido hechos para introducir una caracterización formal del código genético (Jungck, 1978; Siemion, 1995; Jiménez-Montaña, 1999; Gillis, 2001), citados por Sanchez [17], sin embargo, existen trabajos que modelan el código genético buscando una relación mas allá de una relación cuantitativa entre las secuencias genómicas del ADN ([19]).

Aquí, presentamos una forma de construir un álgebra C_g sobre el anillo \mathbb{Z}_{64} , compuesto de 64 tripletas del Código Genético Estandar(CGE), donde cada tripleta(codon) está formada a partir de las 4 bases nitrogenadas: adenina(A), guanina(G), uracilo(U) y citosina(C), las que constituyen un grupo isomorfo a \mathbb{Z}_4 . Luego se construye el espacio S N -dimensional de secuencias de ADN de genes, compuesto por el conjunto de todas la 64^N secuencias con N codones[19],

* Departamento de Matemáticas, Universidad Nacional de Trujillo, Av. Juan Pablo II s/n., Ciudad Universitaria, Trujillo-Perú (orubio@unitru.edu.pe).

**Departamento de Matemáticas, Universidad Nacional de Trujillo, Av. Juan Pablo II s/n., Ciudad Universitaria, Trujillo-Perú (rnoriega@unitru.edu.pe)

La operación de suma definida en este conjunto es una manera de obtener todos los codones consecutivamente a partir del codón *AAC*, de tal manera que el código genético representará un código escala de energía de interacción de aminoácidos en proteínas. En esta interacción hay variaciones importantes, las cuales se pueden verificar mostrando que una mutación puede ser descrita por medio de automorfismos $f : (Z_{64})^N \rightarrow (Z_{64})^N$ sobre C_g .

Sin embargo, este modelo es limitado, puesto que hay cromosomas completos y genes que no pueden ser descritos por medio de la Z_{64} -álgebra C_g . Además de esto, mutaciones que consisten inserciones y deleciones en secuencias de ADN, como el SNP19 en el haplotipo SNP 43, SNP 63 sobre el gen 2q37.3 que producen la diabetes mellitus de tipo 2(DM2), no pueden ser descritas por esta álgebra.

Por tanto, es necesario extender el grupo $(C_g, +)$ a un grupo de código genético extendido, lo cual conducirá que todas las secuencias de ADN de longitud N pueden ser descritos de alguna forma por grupos abelianos finitos que se pueden descomponer en suma directa de 2-grupos y 5-grupos homocíclicos.

2. Conceptos básicos de genética. A continuación presentamos algunos conceptos básicos de la genética¹².

2.1. Genética. Rama de las ciencias biológicas que se dedica a el estudio de la naturaleza, organización, función, expresión, transmisión de la información hereditaria de una generación a la siguiente, su objeto de estudio son los genes, los cuales pueden abordarse desde distintas perspectivas.

2.2. Anatomía celular básica. Según la descripción de Valero [21], cualquier célula de nuestro organismo, a excepción de los eritrocitos, contiene un *núcleo* y un *citoplasma*. El núcleo contiene la mayoría de la información genética de la célula, tiene un aspecto granular debido a la *cromatina*, que justo antes de la división celular, ésta se condensa formando unas estructuras llamadas *cromosomas*. Los **genes** residen en los cromosomas; se transmiten de padres a hijos y por ello se consideran las unidades básicas y funcionales de la herencia. Los genes están compuestos por ácido desoxirribonucleico-ADN y proteínas auxiliares. Al total de la información contenida en los genes se le denomina **genoma**, y sirve para dirigir y regular el desarrollo y funcionamiento de los seres vivos.

Todas las células, menos los *gametos*(células reproductoras, óvulos y espermatozoides) contienen 23 pares de cromosomas. Uno de los pares de cromosomas determina el sexo del individuo y por ello se les denomina *cromosomas sexuales*, los 22 pares restantes se denominan *autosomas* numerados en función del tamaño y clasificados en grupos en función de la posición del *centrómetro*.

2.3. La herencia a nivel molecular: ADN, ARN y proteínas. La **Genética Molecular** estudia la información contenida en el ADN y cómo se utiliza dicha información para dar lugar a las proteínas. En los **eucariotas**(organismos cuyas células tienen núcleo definido), el contenido de ADN no mantiene una relación lineal con la complejidad biológica.

2.3.1. Ácido desoxirribonucleico-ADN. Físicamente, es el componente básico de los cromosomas. La molécula de ADN tiene tres componentes: un azúcar (**desoxi-D-ribosa**), un grupo fosfato y cuatro bases nitrogenadas, nucleótidos (letras): dos de las bases son **pirimidinas**: **Citosina**(C) y **Timina**(T) y las otras dos son **purinas**: **Adenina**(A) y **Guanina**(G).

El ADN consta de dos cadenas entrelazadas *antiparalelas*(tienen direcciones opuestas) de nucleótidos. La estructura que forman estas cadenas es una doble hélice, que mantiene la estabilidad mediante enlaces débiles (puentes de hidrógeno), de tal forma que la timina se une con la adenina y la citosina con la guanina. En el núcleo, la doble hélice de ADN se condensa y empaqueta formando una estructura superenrollada.

Del total del ADN de un organismo, hasta ahora se cree que sólo un 20% es funcional, es decir está involucrado en generar proteínas o cumplir una función en la célula.

2.3.2. De genes a proteínas. La información que contiene el ADN debe ser transportada al citoplasma para que allí tenga lugar la síntesis de proteínas. El transvase de información es universal:



¹C. Valero, C. Hernandez-chico *comprendiendo una enfermedad hereditaria: conceptos básicos de genética* Asociación española de neurofibromatosis, Madrid, 2011.

²Edith del R. Garcia *Manual de autoaprendizaje de genética* [accesado el 02/06/15] <http://www.chapingo.mx/prepa/pdf/gen.pdf>

El primer proceso, de ADN a ARN(ácido Ribonucléico), ocurre en el núcleo y se llama **Transcripción**. La composición del ARN es similar a la del ADN pero contiene como azúcar la **ribosa** en vez de desoxi-D-ribosa y **Uracilo(U)** en vez de Timina. El ARN transcrito utilizando el ADN como molde se llama **ARN mensajero(ARNm)**.

La transcripción está muy regulada; en un primer paso se forma una molécula de ARN totalmente complementaria al ADN molde. Dicha molécula, *transcrito primario*, sufre una serie de modificaciones que le confieren estabilidad. Posteriormente, antes de abandonar el núcleo para ir al citoplasma, se procesa; descartándose determinados fragmentos no codificantes para proteínas, *intrones*, y uniéndose los codificantes, *exones*. El ARN resultante se llama *ARN maduro* y es el que sale al citoplasma para ser *traducido* a proteína.

Las proteínas o polipéptidos están compuestos por *aminoácidos*. Dado que existen 20 aminoácidos diferentes y tan solo 4 bases diferentes en el ARN. El código utilizado está basado en **tripletes o codones**(combinación de tres bases). De este modo, los 20 aminoácidos se codificarían con 64 combinaciones posibles(en realidad son 61 porque 3 de los codones son de parada de síntesis), algunos aminoácidos son codificados por varios codones pero siempre un codón sólo codifica un aminoácido.

El proceso de **traducción**, paso de ARNm a proteína, es complejo. El ARNm no puede unir directamente los aminoácidos y necesita un intermediario, **ARN transferente(ARNt)**. Dicho ARNt tiene forma de trébol que en una de sus hojas tiene una secuencia llamada **anticodón** que es complementaria al codón. La maquinaria celular donde se produce la síntesis de proteínas se encuentra en el citoplasma y es el **ribosoma**, quien primero se une al sitio de iniciación de las síntesis en el ARNm, permitiendo la interacción entre el ARNm y el ARNt. Una vez incorporado el aminoácido correspondiente se desplaza al siguiente codón y así sucesivamente hasta que el ribosoma llega a uno de los codones de parada en el ARNm y la traducción cesa.

2.4. Mutación. Una mutación es un cambio en la secuencia del ADN. Las mutaciones pueden ser el resultado de errores en la copia del ADN durante la división celular, la exposición a radiaciones ionizantes o a sustancias químicas denominadas mutágenos, o infección por virus. Las mutaciones de la línea germinal se producen en los óvulos y el esperma y puede transmitirse a la descendencia, mientras que las mutaciones somáticas se producen en las células del cuerpo y no se pasan a los hijos.

2.4.1. Mutación puntual. Una mutación puntual se produce cuando se altera un solo par de bases. Las mutaciones puntuales pueden tener uno de los tres efectos siguientes. En primer lugar, la sustitución de una base puede ser una mutación silenciosa o sea, el codon alterado produce el mismo aminoácido. En segundo lugar, la sustitución de base puede ser una mutación sin sentido en que el codon alterado da lugar a un aminoácido diferente. En tercer lugar, la sustitución de una base puede producir una mutación sin sentido y el codon alterado puede corresponder a una señal de terminación.

2.5. Alelo. Un alelo es cada una de las dos o más versiones de un gen. Un individuo hereda dos alelos para cada gen, uno del padre y el otro de la madre. Los alelos se encuentran en la misma posición dentro de los cromosomas homólogos. Si los dos alelos son idénticos, el individuo es homocigoto para este gen. En cambio, si los alelos son diferentes, el individuo es heterocigoto para este gen. Aunque el término alelo fue usado originariamente para describir variaciones entre los genes, ahora también se refiere a las variaciones en secuencias de ADN no codificante.

2.6. Haplotipo. Un haplotipo es un conjunto de variaciones del ADN, o polimorfismos, que tienden a ser heredados juntos. Haplotipo se puede referir a una combinación de alelos o a un conjunto de polimorfismos de nucleótido sencillo (SNPs) que se encuentran en el mismo cromosoma.

2.7. Polimorfismo de un solo nucleótido - SNP. (Single Nucleotide Polymorphism, pronunciado snip) es una variación en la secuencia de ADN que afecta a una sola base (adenina (A), timina (T), citosina (C) o guanina (G)) de una secuencia del genoma. Sin embargo, algunos autores consideran que cambios de unos pocos nucleótidos, como también pequeñas inserciones y deleciones (indels) pueden ser consideradas como SNP, siendo entonces más adecuado el término Polimorfismo de nucleótido simple.

2.8. El Código Genético. Es un diccionario para traducir una secuencia de nucleótidos en el ARN a una secuencia de aminoácidos en una proteína. La síntesis de proteína es la traducción del lenguaje de secuencia de nucleótidos en el ARNm al lenguaje de secuencia de aminoácidos. Se dilucida en el año 1961 por Crick, Brenner y colaboradores[6]

Un mismo aminoácido es codificado por varios codones, salvo Triptófano y Metionina que están codificados por un único codon. Los codones que codifican un mismo aminoácido en muchos casos comparten los dos primeros nucleótidos con lo que se minimiza el efecto de las mutaciones. En estos casos una mutación en la tercera posición del codon no cambia el aminoácido codificado, mutación silenciosa.

El codon AUG que codifica la metionina es el codon de inicio y hay tres codones que establecen la señal de terminación de la traducción (UAA, UAG, UGA). Las mutaciones que ocurren en estos codones dan lugar a la síntesis de proteínas anómalas.

3. Modelación. Muchos problemas en biología crean la necesidad de descubrir nuevas teorías matemáticas o de potenciar y desarrollar las ya existente. Recíprocamente, el tremendo desarrollo de la investigación en matemática hace posible el estudio de problemas cada vez mas complejos de la biología, especialmente en el ámbito de la genética.

Algunos autores han realizados trabajos extensos, en lo relacionado con la genética, entre los cuales se encuentran Sánchez, Morgado y Grau[19]. Sus aportes consisten en construir una herramienta que permita estudiar las relación entre las estructuras algebraicas con las propiedades de los codones y sus procesos que se llevan a cabo en la síntesis de proteínas. En este trabajo se buscara exponer las anteriores ideas identificando una aplicación en las mutaciones en los genes que desarrollan la diabetes millitus Tipo 2.

Investigadores como Jennings(1917), Serebrovsky(1934) y Gilvenko(1936)(citados por [17]), aportaron a su trabajo, al darle un carácter mas matemático incluyendo casos diversos. La formulación algebraica de la herencia se presentó en 1939, cuando Etherington planteó una estructura mas compleja que en los anteriores casos. Etherington es considerado el padre de las aplicaciones del álgebra a la Genética.

3.1. Algebra del código genético. El código genético ha sido representado, generalmente, en una tabla de 4 columnas donde los codones son localizados según la segunda base, ampliamente discutidos por Crick [4], [6]. Según Sánchez [19] el orden de los codones debe reflejar sus propiedades físicas y químicas de las cuatro bases. La importancia de la posición de base es sugerida por la frecuencia de errores encontradas en los codones. Los errores en la tercera base son mas frecuentes que en la primera base y, a la vez estos son mas frecuentes que los errores en la segunda base[17]. Estas posiciones, sin embargo, también son conservativas con respecto a los cambios en polaridad de los aminoácidos coodificados.

La base de este trabajo son estructuras algebraicas bien conocidas[11, 7]: grupos, anillos, módulos y álgebras que construimos a partir de las cuatro bases del ADN que forman el código genético.

En realidad las álgebras genéticas aquí son secuencias de ADN. El término “genético ” aparece donde generalmente deberíamos usar una secuencia de ADN.

3.2. Grupos de cuatro bases. Permitamos que las bases de ADN se pueden organizar de acuerdo a sus propiedades físico-químicas. Los tipos químicos, (purina y pirimidina) y los enlaces del número de hidrógenos son elementos de las interacciones del codon-anticodon identificando un orden en el conjunto base $\{A, C, G, U\}$.

Asumimos la estructura del grupo B_4 formado por las bases nitrogenadas en el ARN $\{A, C, G, U\}$ junto con la operación \bullet , según la tabla del a figura 3.1 siendo isomorfo a \mathbb{Z}_4 . Elegimos \mathbb{Z}_4 por su carácter cíclico, que permite estabilidad a la molécula de ADN.

$A \leftrightarrow 0$	\bullet	A	C	G	U	$+$	0	1	2	3
$C \leftrightarrow 1$	A	A	C	G	U	0	0	1	2	3
$G \leftrightarrow 2$	C	C	G	U	A	1	1	2	3	0
$U \leftrightarrow 3$	G	G	U	A	C	2	2	3	0	1
	U	U	A	C	G	3	3	0	1	2

FIGURA 3.1. Grupo de bases nitrogenadas, y su isomorfismo

3.2.1. Construcción del Grupo de Codones. Recordemos que un Codon es un triplete de nucleótidos, obtenidos del producto de $B_4 \times B_4 \times B_4$, es decir existen sólo 64 codones en el ARN.

Denotemos este conjunto por

$$C_g = \{XYZ, \quad \text{tal que,} \quad X, Y, Z \in B_4\}$$

Se puede asociar cada codon a un número, de 0 hasta 63, y luego veremos que los codones actúan isomórficamente con la estructura algebraica de cíclica $(\mathbb{Z}_{64}, +)$.

Introducimos la operación suma de los codones XYZ y X',Y',Z' siguiendo a Sanchez[19], que respete las distinciones entre las posiciones de las bases del codon, el orden del conjunto de las 4 bases y la operación suma de las bases dada en la tabla. La operación suma entre dos codones es obtenida desde la base menos importante biológicamente(tercera posición del codon: Z y Z') a la de más importancia (segunda posición del codon: Y y Y'):

1. Las bases correspondientes en la tercera posición son añadidas de acuerdo a la tabla de la suma 3.1.
2. Si la base resultante de la operación suma es previa en el orden a la s abases sumadas(los órdenes en el conjunto de bases), entonces el nuevo valor es escrito y la base C (o G si el grupo de bases dual es usado) es sumada a la siguiente posición .
3. Las otras bases son sumadas de acuerdo a la tabla de la suma, paso dos, yendo desde la primera base a la segunda base.

Por ejemplo, la suma de codones $AGC \oplus UGU$, usando el grupo primal de bases tenemos:

$C+U = A$, las terceras bases son sumadas y la base C es sumada a la siguiente posición, porque la base A precede a la base C y U en el conjunto de bases $\{A, C, G, U\}$.

$A+U+C=U+ C=A$, las bases primeras y la base C obtenida en el primer paso son añadidas.

Otra vez la base C es añadida a las bases segundas t tenemos $G+G+C= A+C=C$.

Finalmente tenemos el resultado

$$AGC \oplus UGU = ACA$$

La operación de suma definida satisface los axiomas de grupo, obteniendo el grupo denotado por (C_g, \oplus) , ya que todos los grupos cíclicos finitos con el mismo número de elementos son isomorfos, entonces este grupo es isomorfo a $(\mathbb{Z}_{64}, +)$, según se puede ver en la Tabla 3.1.

TABLA 3.1
Isomorfismo del código genético C_g con \mathbb{Z}_{64} propuesto por Sánchez 2005(a)

	No	A Codón	*	No	C Codón		No	G Codón		No	U Codón		
A	0	AAA	K	16	ACA	T	32	AGA	R	48	AUA	I	A
	1	AAC	N	17	ACC	T	33	AGC	S	49	AUC	I	C
	2	AAG	K	18	ACG	T	34	AGG	R	50	AUG	M	G
A	3	AAU	N	19	ACU	T	35	AGU	S	51	AUU	I	U
C	4	CAA	K	20	CCA	P	36	CGA	R	52	CUA	L	A
	5	CAC	H	21	CCC	P	37	CGC	R	53	CUC	L	C
	6	CAG	Q	22	CCG	P	38	CGG	R	54	CUG	L	G
C	7	CAU	H	23	CCU	P	39	CGU	R	55	CUU	L	U
G	8	GAA	E	24	GCA	A	40	GGA	G	56	GUA	V	A
	9	GAC	D	25	GCC	A	41	GGC	G	57	GUC	V	C
	10	GAG	E	26	GCG	A	42	GCG	G	58	GUG	V	G
G	11	GAU	D	27	GCU	A	43	GGU	G	59	GUU	V	U
U	12	UAA	Y	28	UCA	S	44	UGA	-	60	UUA	L	A
	13	UAC	-	29	UCC	S	45	UGC	C	61	UUC	F	C
	14	UAG	Y	30	UCG	S	46	UGG	W	62	UUG	L	G
U	15	UAU	-	31	UCU	S	47	UGU	C	63	UUU	F	U

Funciona de la siguiente manera: el codon de inicio AUG,(metionina) es inverso al codon de parada UAG.

$$AUG \oplus UAG = AAA, \quad \equiv \quad 50 + 14 = 0$$

TEOREMA 1. El conjunto C_g formado por los codones y la operación \oplus definida anteriormente forman un grupo.

- Es asociativo.
- El elemento neutro es AAA.
- Para cada $XYZ \in C_g$ existe el inverso MNO , tal que $XYZ \oplus MNO = AAA$
- También se prueba que es conmutativo.

- (C_g, \oplus) es grupo cíclico

Observe que los codones con bases A y G en la tercera posición son codones pares y los codones con bases C y U son codones impares.

3.2.2. Álgebra de genes. Como el grupo (C_g, \oplus) es cíclico entonces existe algún generador XYZ tal que para diferentes naturales $nXYZ$ genera el grupo, ver [7], los generadores de este grupo son codones impares. De esta forma para cualquier codon u existe un natural k tal que $u = kX_1X_1Z_1$, mas aún este k se puede elegir en \mathbb{Z}_{64} , y el $X_1Y_1Z_1 = AAC$.

A continuación se puede definir una operación producto en \odot en C_g , tal que $AAC \odot AAC = AAC$ de modo que para todo $u, v \in C_g$

$$u \odot v = kAAC \odot k' AAC = k.k' AAC \odot AAC = k.k' AAC$$

donde \cdot denota la operación producto en $(\mathbb{Z}_{64}, +, \cdot)$, el cual es a su vez isomorfo a (C_g, \oplus, \odot) , Además también se puede definir la ley externa $f : \mathbb{Z}_{64} \times (C_g, \oplus)$ por $f(k, u) = ku$. Se puede interpretar al grupo (C_g, \oplus) como un \mathbb{Z}_{64} módulo y un \mathbb{Z}_{64} -álgebra sobre el anillo $(\mathbb{Z}_{64}, +, \cdot)$.

Ahora sea el espacio de sucesiones N -dimensional, compuesto por la suma directa de N veces (C_g, \oplus) , esto es

$$(P, +) = (C_g, \oplus)^N = (C_g, \oplus) \times \dots \times (C_g, \oplus), \quad N \text{ veces}$$

el cual a su vez es un \mathbb{Z}_{64} módulo y un \mathbb{Z}_{64} -álgebra sobre el anillo $(\mathbb{Z}_{64}, +, \cdot)$.

Si denotamos al anillo por \mathbb{Z}_{64} y $(P, +)$ por $(C_g)^N$: Podemos encontrar una base para $(C_g)^N$ compuesto por elementos de la forma $e_i = (0, \dots, AAC, \dots, 0)$ en la posición i -ésima, Cualquier elemento $u \in (C_g)^N$ se descompone en una combinación lineal de su base y cualquier automorfismo $f : (C_g)^N \rightarrow (C_g)^N$ tiene una representación matricial $A = (a_{ij})$ de orden $N \times N$. Finalmente como \mathbb{Z}_{64} es un 2-grupo abeliano, el endomorfismo es un automorfismo si y solo si $det(A)$ no es divisible por 2.

De acuerdo al teorema de Lagrange los órdenes de los elementos de \mathbb{Z}_{64} son los divisores de 64, los cuales tienen forma $2^m (m = 0, 1, 2, 3, 4, 5, 6)$. Los elementos con orden 2^m tienen la forma $2^{6-m}x$ donde x es un entero impar.

3.3. Conexiones con la mutación de genes: Automorfismos. Se ha encontrado conexiones muy interesantes entre las propiedades biológicas y algebraicas en el \mathbb{Z}_{64} -módulo $(C_g)^N$.

El conjunto de genes mutantes con respecto al tipo natural puede representarse a través de endomorfismo y automorfismos en $(C_g)^N$, un endomorfismo se llama local si existe un $k \in \{1, 2, \dots, N\}$ y $a_{ik} \in \mathbb{Z}_{64}, i \in \{1, 2, \dots, N\}$ tal que $f(e_i) = a_{ik}e_i$,

Por tanto un endomorfismo es un automorfismo si y solo si su matriz es diagonal y el elemento a_{kk} en la representación matricial es un número impar, del endomorfismo que cumple $f(e_k) = a_{kk}e_k$ y los demás funcionan $f(e_i) = e_i$ para $i \neq k$.

Se sabe que las mutaciones de un solo nucleótido que preservan el tipo químico (transiciones) son frecuentemente menos peligrosas que las que alteran sus propiedades(transversiones). Consecuentemente las transiciones son mutaciones en su mayoría simples que se encuentran en la naturaleza.

En la Tabla del código genético 3.1 se observa que las transiciones y transversiones son conectadas con cambios en la paridad y el orden de los codones según:

1. Sin alterar la paridad del codon, las transiciones en la segunda base pueden mantener el orden del codon.
2. La transición, ya sea en la primera o en la tercera base mantienen la paridad del codon pero el orden del codon puede cambiar. Estas mutaciones generalmente no presentan cambios extremos en las propiedades físico-químicas de los aminoácidos.
3. Las transversiones ya sea en la primera o en la segunda base mantienen la paridad del codon pero pueden variar el orden del codon. Las transversiones en la primera base generalmente mantienen las propiedades hidrofóbicas, sin embargo las transversiones de la segunda base son generalmente peligrosas debido a las proteínas.
4. Las transversiones de la tercera base cambian la paridad del codon y el orden, sin embargo esta clase de mutaciones generalmente no produce cambios hidrofóbicos extremos.

Si se considera el orden de los elementos (de los genes o secuencias de ADN), las mutaciones(transiciones, transversiones) funcionan como se puede ver en el siguiente teorema y su corolario, siguiendo a Sánchez [19].

TEOREMA 2. *Para cualquier par (α, β) de elementos de $(C_g)^N$, existe un endomorfismo $f : (C_g)^N \rightarrow (C_g)^N$, tal que $f(\alpha) = \beta$, si y solo si el orden de β es divisor del orden de α .*

Prueba: Puesto que $(C_g)^N$ es un 2-group homocíclico Abeliano isomorfo al 2-group homocíclico Abeliano $(Z_{64})^N$, se cumple el teorema para este caso particular. Veamos, es bien conocido que cualquier homomorfismo de grupos, que cualquier elemento α es transformado en un elemento cuyo orden divide al orden de α , por tanto la condición reversa es probada.

Para probar la directa, sea $\alpha, \beta \in (Z_{64})^N$ de órdenes 2^{m_α} y 2^{m_β} respectivamente, esto es $\alpha = 2^{6-m_\alpha}x_\alpha$ y $\beta = 2^{6-m_\beta}x_\beta$, donde x_α y x_β son generadores de Z_{64} ,

Podemos elegir una base (x_1, \dots, x_N) de $(Z_{64})^N$ con $x_1 = x_\alpha$ y otra (y_1, \dots, y_N) de $(Z_{64})^N$ con $y_1 = x_\beta$. Estas bases definen un único automorfismo g en $(Z_{64})^N$ tal que $g(x_i) = y_i$ para $i = 1, \dots, N$. Ahora, suponiendo que $2^{m_\alpha} \geq 2^{m_\beta}$, esto es $m - \alpha \geq m_\beta$, tomando el automorfismo $f = 2^{m_\alpha - m_\beta}g$ se cumple $f(\alpha) = 2^{m_\alpha - m_\beta}g(\alpha) = 2^{m_\alpha - m_\beta}g(2^{6-m_\alpha}x_\alpha) = 2^{6-m_\beta}g(x_\alpha) = 2^{6-m_\beta}x_\beta = \beta$, que es lo que deseamos probar.

□

Este Teorema dice lo siguiente: Para cualquier par de genes (α, β) en el espacio de secuencias de dimensión N , existe al menos un automorfismo f en $(C_g)^N$ tal que $f(\alpha) = \beta$, además se cumple:

$$f(a \cdot (\mu + \nu)) = a \cdot f(\mu) + a \cdot f(\nu), \quad \forall \mu, \nu \in (C_g)^N, \quad a \in Z_{64}.$$

Estos automorfismos interpretan las reversiones de las mutaciones.

COROLARIO 1. *Existe un automorfismo diagonal $f : (C_g)^N \rightarrow (C_g)^N$, que transforma α a β si y solo si para todos los pares de coordenadas $\alpha_i, \beta_i \in Z_{64}, i = 1, \dots, N$, de los vectores α, β , se cumple las desigualdades, $m_{\beta_i} \leq m_{\alpha_i}$. De acuerdo al Corolario, si $\alpha_i = 2^{6-m_{\alpha_i}}x_i$ y $\beta_i = 2^{6-m_{\beta_i}}y_i$*

son las componentes de los vectores α, β respectivamente entonces, la representación matricial A del endomorfismo f sobre una base del módulo $(C_g)^N$, los elementos de la diagonal son calculados por:

$$a_{ii} = 2^{m_{\alpha_i} - m_{\beta_i}}y_i x_i^{-1}.$$

Se espera que los mutantes generadores de la diabetes tipo 2 tienen orden igual o más pequeño que el del tipo del cual multaron, por lo cual deben satisfacer el Corolario, lo cual indica que existe un automorfismo diagonal que transforma el gen natural en el mutante.

Se puede observar que si para todo par de componentes α_i, β_i el orden se mantiene, en la matriz diagonal se encuentran sólo números impares y el endomorfismo es un automorfismo. Por lo tanto, si en una mutación puntual el orden del codon se mantiene entonces existe un automorfismo diagonal local que la transforma el tipo natural en el mutante. Por otra parte, en el gen natural podemos encontrar múltiples sustituciones de bases en un gen que mantiene el orden de codones. En este caso hay un automorfismo diagonal que va del de tipo natural al gen mutante.

4. Descripción del gen CAPN10 y la DM2. La diabetes mellitus de tipo 2 - DM2 tiene un fuerte componente genético. Se ha estudiado casi 50 candidatos de genes, sin embargo las ideas divergen de acuerdo a los grupos étnicos, exposiciones ambientales, interacción de genes- ambiente.

Algunos de estos genes se cree que están implicados en la función pancreática de células β , acción de la insulina/metabolismo de la glucosa, u otras condiciones metabólicas que aumentan el riesgo diabetes tipo 2 (por ejemplo, la ingesta de energía/gastos, metabolismo de los lípidos)[12].

Por ser el primer gen con estas características identificado por clonaje posicional en la región de NIDDDM1 (no insulino dependiente diabetes mellitus 2) se eligió el CAPN10.

Este gen codifica una cisteína proteasa calcio dependiente intracelular, es decir ubica expresado [16], llamada Calpaína 10. Este gen tiene 15 exones y 14 intrones con una longitud de 31 kilobases[13], Horikawa [9] secuenciaron una región de 66 kilobases, incluyendo el gen de la calpaína 10, y encontraron 179 polimorfismos, principalmente del tipo de la sustitución de un simple nucleótido, SNP.

4.0.1. Análisis genético. Localización citogénica: 2q37.3, es decir en el brazo largo del cromosoma 2 en la posición 37.3

Localización molecular: Se inicio en 240,586,716 bp del pter y finaliza en 240,617,705 bp del pter

Tamaño: 30,990 bases, figura 4.1³

³<http://www.genecards.org/cgi-bin/carddisp.pl?gene=CAPN10>

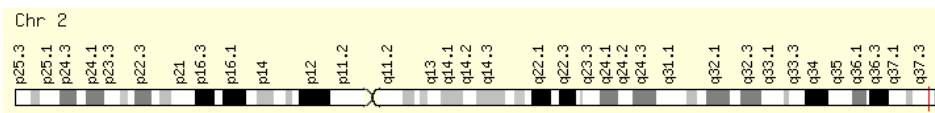


FIGURA 4.1. CAPN10 en el cromosoma 2 marcado con línea roja. tomado de Human Genes Database

4.0.2. Ubicación Regional del gen CAPN10 . Sobre el cromosoma 2 se ha definido una región compuesta por 193 genes ⁴

La longitud de esta región comprende el intervalo (238,586,716 - 242,193,529) de bases. Lo cual significa que tiene una longitud de 3 606 813 pb, equivalente a 3.6 cM aproximadamente

A continuación en la Tabla (4.1) se presenta 10 genes, extraídos del intervalo arriba, con sus respectivas posiciones, donde se observa claramente el gen CAPN10 en el orden 86.

TABLA 4.1
10 genes y el gen CAPN10

No	Inicio	Fin	GEN	Descripción
81	240560054	240564014	DUSP28 Exon structure	dual specificity phosphatase 28
82	240564016	240566475	GC02M240564	
83	240565804	240581372	RNPEPL1 Exon structure	arginyl aminopeptidase (aminopeptidase B)-like 1
84	240582700	240586699	CAPN10-AS1 Exon structure	CAPN10 antisense RNA 1 (head to head)
85	240585546	240585574	PIR33196 Exon structure	
86	240586716	240617705	CAPN10 Exon structure	calpain 10
87	240605408	240631259	GPR35 Exon structure	G protein-coupled receptor 35
88	240672403	240673097	LOC100420500 Exon structure	
89	240676416	240685768	AQP12B Exon structure	aquaporin 12B
90	240685942	240690414	LOC285191 Exon structure	Uncharacterized LOC285191 (est)

5. Mutaciones y Combinaciones haplotípicas de los polimorfismos SNP-43, SNP-19 y SNP-63 del gen CAPN10 . Horikawa et al. [9](2000), dice que diversos estudios indican que las combinaciones haplotípicas de los polimorfismos SNP-43, SNP-19 y SNP-63 mostraban asociación con mayor riesgo de DM2 en diferentes poblaciones. Veamos las características de ellos [10].

El SNP43, sustitución de una guanina(G) por una adenina(A) está ubicado en el intron 3 que tiene una longitud de 922bs y en la posición rs3792267= 240 591 757 (alelo 1= G y alelo 2=A), en la siguiente secuencia(tomados de la base NCBI ⁵) de 1001 bp, que contienen al intron 3, se encuentra en el lugar 501, los alelos A/G marcado con R. CACAGAGAAC ATGTGTGCCG TCCTCCTTAT

```

TTTATCGGCC CCAGCAAGAA AGATGCTTCT
TTATATTTGT TGTGGAGTGG TTGGGACAGG CAGACTCATT GTGTAGTCGT TGGGGAGGAA
TGAGGCTACC CCAGCATACC AACACTTGTG TATCACGGTG CTTGCTGGCT CAGGGGACCA
GGACCCTCAC CATGAGTCAT AATTGAATAG CCTCCCTCT TAGAATGCAT TTGTCTTCTT
GCCAAAGGCA ACTGGACTGA CAGGCAGGCA GGAAGCTGG TGAACATGGG AAGGCTGGCT
GGTGACATCA GTGCCAGTG AGCCCTTCCA TCCCAAGGGC TGTTTTAGGA AAAGCAGGGT
TGGAGCTTGA GAGCCAAGGG ATGTGGGCAT CCATAGCTTC CACGCCTCCT GCCCTGCTCC
TGTGCCACA CCGGATGCCA GAGAGTTTCT GTGTGTGGGC AGAGGACTGC AGGGCGCTCA CGCTTGCT-
GT GAAGTAAGGC
R
    
```

⁴https://genecards.weizmann.ac.il/geneloc-bin/display_map.pl?chr_nr=2&range_type=gc_id&gc_id=GC02P240586

⁵<https://www.ncbi.nlm.nih.gov/SNP/snp-ref.cgi?rs=rs3792267>

TTTGAAGGTG AGGCTAAGCC TTGACTTGGT GAGGATGAGG AAGAAGGCAG AGGGGAGTAA
 AGAGGTGGGA TTGAGGCAGC GGTGGACGA TTTGGGGTGC TACAGACCAT GGGAATCAGA
 GAGGGGGCCA TGCTCAATGC CAGAGGCTCA CTCCCATGGT GATTGTGTCC CCTAGGGTCC
 ATGGGTCTA CGAGCACCTG TGGGCCGGC AGGTGGCCGA TGCCCTGGT GACCTGACCC
 GCGGCCTGGC AGAAAGATGG AACCTGAAGG GCGTAGCAGG AAGCGGAGGC CAGCAGGACA
 GGCCAGGCCG CTGGGAGCAC AGGACTTGTC GGCAGCTGCT CCACCTGAAG GACCAGTGTC
 TGATCAGCTG CTGCGTGCTC AGCCCCAGAG CAGGTGAGGC ACGTGCCAG CATGGGAGGG
 CTGCAGCCAG CTGCCCCCCT ACTGCCAGGC CTCAGGCACA CTGTAGCTTT TTATGTGACT GGCTACACAG
 CCCTGTCAGG

El SNP19, poco estudiado, ubicado en el intron 6 y no es propiamente de nucleótido único, sino de inserción/delección(indel) con 2 o 3 repeticiones de 32 pares de bases[12](2 repeticiones, una secuencia de 32 bp, tres repeticiones de una secuencia de 32bp; alelo1= 2 repeticiones y alelo2 = 3 repeticiones). El SNP 19 tiene la posición rs3842570= 240 594 876 b. en el intron 6 de longitud 315bp entre (240 594 710 - 240 595 023)

A continuación presentamos la ubicación del C (inlet) en la posición 61 de 121 bases.
 CGTGGAGAGA TGATCTGTG CAGGAGCCG GGAGGAGGT GATGATCTG TCCAGGAGC

N

TGGGAGGAG GTGGGCTTGT GGGAGGGGCT GGCTCTGTCT GTGGCCGTAG CTGCTGCTTA

y el SNP63 en el intron 13 es una sustitución de C por T. A continuación presentamos el lugar del SNP63 en la posición alelica 251 de un total de 501 bases

CCTTGCCAGG GTGTCAAACC TGTGTCTCGG GACACTGCTG TTAGGTCTCC AGCTTCCCTA
 TCAGGGCCT CAGCACCCAG TCCTACCAGT GCTCCCGCCT CCGTCCCCA GCTGGCTGGG
 CCTGCAGCCC CCTCCTGTGC CCGAGCTGG CCGGGCCCGC AGCCACTCC CTGGTCACTG
 GATGTTGCTG ACACCTCACT CGGTCAGAGC CBTAGCACCC AAGGGGGGCC AGGGCCTGAC
 GGGGTGGAG

Y

GAGGGTGGG CCGCGTCTGT GCAGGCTCAG AAGCTTCCCTA AGAGGCTGGA GAGTGAACC
 TTCAGGCACC ACGACTGCC TCCTCCCTGC CCACGTCCT GTTTCTCCA GATGGGCCT
 TGGCCTTGGC TAGGTGTTGA TCAGGAGCTG GGAGTGCTGC GCCCCGCCA ACTTCTCCA
 ACTCCAGCCA GGCACCTCA GTGAGGCCTC AGCCACTGC GCCTATTTG CTTCTCCTT
 GGAGCCCTC

5.1. Modelo para la región del CAPN10. Ahora interpretemos la secuencia de ADN y las mutaciones SNP del CAPN10:

En primer lugar la mutación SNP43, se puede representar por un homomorfismo diagonal sobre la \mathbb{Z}_{64} álgebra $(c_g)^N$, considerado como un r-módulo de dimensión $N = 162$, que corresponden a la longitud el gen aproximadamente que 30 989 pb.

En este contexto se tiene una matriz de orden N^2 , donde en la posición 26 se realiza la mutación G/A .

Por tanto, para mejor presentar la matriz, exhibimos un bloque de $N=4$, que contiene al codón con el alelo G. Se tiene que el vector $\alpha = TAAGGC GTTTGA$ transformado a la porción del vector $\beta = TAAGGC ATTTGA$, en términos de coordenadas del \mathbb{Z}_{64} -álgebra $(C_g)^4$ se tiene $\alpha = (12, 37, 59, 44)$ y $\beta = (12, 37, 51, 44)$, que se puede escribir como

$$(12, 37, 51, 44) = (12, 37, 59, 44) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 56 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Para el SNP 63 en la región 3' se tiene la siguiente transformación matricial, exhibiendo un bloque de $N=4$, que contiene al codón con el alelo G, igualmente el vector $\alpha = GTG GAG CGA GGG$ transformado a la porción del vector $\beta = GTG GAG TGA GGG$, que en términos de coordenadas del \mathbb{Z}_{64} -álgebra $(C_g)^4$ se tiene $\alpha = (58, 10, 36, 42)$ y $\beta = (58, 10, 44, 42)$, que se puede escribir como

$$(58, 10, 44, 42) = (58, 10, 36, 42) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Sin embargo para modelar el SNP19 que consiste en delecciones/inserciones, entonces es necesario extender el grupo base del código genético. Sánchez [20] propone la extensión formada por el grupo base $\{O, A, C, G, U\}$, que siguiendo un proceso muy similar al presentado con las 4 bases, se construye el grupo cíclico del código genético extendido C_e formado por tripletas del código base

extendido, y por tanto el grupo es isomórfico a \mathbb{Z}_{125} . Pero además se interpreta al grupo (C_e, \oplus) como un \mathbb{Z}_{125} -módulo sobre el anillo \mathbb{Z}_{125} .

De manera que, cuando analizamos las secuencias de longitud N . En las secuencia genómica del ADN se encuentra los marcos de lectura abierta(ORF) que son construcciones de bloques de genes.

Y si analizamos los múltiples arreglos de sucesiones hallamos subregiones donde no hay espacios introducidos y sólo encontraron mutaciones de sustitución. Estos bloques construídos corresponden a exones completos o subregiones, que pueden ser descritos por medio del grupo (C_g, \oplus) , mientras que las regiones donde aparecen saltos como resultado de las mutaciones indel(insertión,delección), pueden ser descritos por medio del grupo monogénico (C_e, \oplus)

Esto es fundamentalmente una aplicación del teorema fundamental para grupos abelianos finitos. Donde dice que cualquier grupo abeliano finito es isomórfico a un producto directo de grupos cíclicos de orden potencia de un primo. en particular el teorema presenta una descomposición del grupo abeliano finito como producto directo de grupos cíclicos que son bloques de genes.

En el presenta caso el grupo definido sobre el espacio de secuencias S formado por todos los posibles arreglos de secuencias $64^{N_1}125^{N_2}$, de longitud $N = N_1 + N_2$, $N_1 = n_1 + n_2 + \dots + n_p$ y $N_2 = m_1 + m_2 + \dots m_p$, es un grupo heterocíclico. Este grupo se descompone en suma directa de p-grupos homocíclicos y cada uno de ellos se descompone ne suma directa de p-grupos cíclicos con el mismo orden

Considerando la region mayor para la inlet, el haplotipo SPN43, SPN19 y SPN63 es una composición de endomorfismo sobre el espacio descompuesto en suma directa

$$S = (\mathbb{Z}_{2^6})^{N_1} \times (\mathbb{Z}_{5^3})^M \times (\mathbb{Z}_{2^6})^{N_2}$$

donde $N = N_1 + N_2 + M = 162 + 93 = 255$ (30,900+ 35100pb= 66kb= 66,000pb),
Tal efecto del haplotipo se representa por una matriz diagonal de orden 255×255 .

6. Conclusiones.

1. Se ha presentado los fundamentos de la genética, y el funcionamiento del código genético, quien describe el proceso de la síntesis de proteínas.
2. Se ha descrito las mutaciones genéticas, principalmente el haplotipo formado por las mutaciones SPN 43,19,63, causantes de la DM2.
3. Se identificó el endomorfismo del haplotipo sobre el sistema algebraico de codones(grupo de bases nitrogenadas y su extensión) que es un producto directo de grupos finitos, $(\mathbb{Z}_{2^6})^{81} \times (\mathbb{Z}_{5^3})^{93} \times (\mathbb{Z}_{2^6})^{81}$.

Referencias

- [1] ANGEL MARTIN MUNICIO. **Presente y Futuro de la Biotecnología** (visitado en 25/05/15) <http://www.rac.es/ficheros/doc/00323.pdf>
- [2] ANTONIO J. CARUZ A. *Proyecto Genoma Humano* (acceso 10/06/15) <http://www.ujen.es /investiga/inmunoge/ gmo/tema-genoma-humano.pdf>
- [3] Centro de Biotecnología de la Universidad de Concepción, (accesado el 01/06/15) <http://www.centrobiotecnologia.cl/index.php/que-es-la-biotecnologia>
- [4] GLORIA ROMERO VASQUEZ. *Biotecnología: generalidades, riegos y beneficios*, UNED, 2008
- [5] INGRID BRENA SESMA. *Análisis genético y manipulación genética en varios documentos internacionales*,(visitado 3/06/15) <http://biblio.juridicas.unam.mx/libros/5/2252/9.pdf>
- [6] Crick, F.H.C, *The Origin of the Genetic Code*. J.Nol.Biol.38-367(1968).
- [7] FRALEIGH, J.B. *First course in abstract Algebra*, Edit Pearson, 7th Edition, 2003.
- [8] GARCÍA-ESCALANTE, MARIA , SUAREZ-SOLIS, VICTO; LÓPEZ-ÁVILA, MARIA; PINTO-ESCALANTE, DORIS; LAVIADA-MOLINA, HUGO. *Efecto de los polimorfismos Gly972Arg del gen IRS1, SNP43 del gen CAPN10 y Pro12Ala del gen PPARG2 sobre la falla secundaria a sulfonilureas y metformina en pacientes con diabetes tipo 2 de Yucatán, México*.m Investigación Clínica, 50(1), 65-76, 2009.
- [9] HORIKAWA Y, ODA N, COX N J, LI X, ORHO-MELANDER M, HARA M, HINOKIO Y, LINDNER TH, MASHIMA H, SCHWARZ PE, DEL BOSQUE-PLATA L, HORIKAWA Y, ODA Y, YOSHIUCHI I, COLILLA S, POLONSKI KS, WEI S, CONCANNON P, IWASAKI N, SCHULZE J, BAIER LJ, BOGARDUS C, GROOP L, BOERWINKLE E, HANIS CL Y BELL GI. *Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus*. Nat Genet; 26: 163-175;2000.
- [10] KAHN, RONALD, ET AL. *Joslin's Diabetes Mellitus*, 14 edition, Lippincots Williamas & Wilkins, Philadelphia. 2005, en (<http://www.meduweb.com/threads/30450-Joslin-s-Diabetes-Mellitus-Free-Book-PDF>)
- [11] LANG, S. , *Algebra*. Addison Wesley Publishing Company COLUMBIA University. New York. (1965).
- [12] LOYA MÉNDEZ, YOLANDA; GILBERTO LEAL,GILBERTO; SÁNCHEZ GONZÁLEZ, ADRIANA; PORTILLO REYES, VERÓNICA; REYES RUVALCABA, DAVID Y BOJÓRQUEZ RANGEL, GUILLERMO. *Variantes genotípicas del SNP -19 del gen de la CAPN 10 y su relación con la diabetes mellitus tipo 2 en una población de Ciudad Juárez, México*, Nutrición Hospitalaria, 31(2), 74-750, 2015.

- [13] PAREDES ANAYA, MÓNICA; LIZARASO SOTO, FRANK; LISSÓN ABANTO, ROSA, RODRÍGUEZ, GIOVANNA; CALDERÓN, JORGE; RODRÍGUEZ ZÁRATE, EDUARDO; NORES, GABRIELA Y FUJITA ALARCÓN, RICARDO. *Variación y distribución genética de los SNP's 19, 43 Y 63 del gen de susceptibilidad de diabetes tipo 2 Calpain 10 (Capn10) en la población peruana*, Horizonte Médico, rev. de la Fac. de Medicina de la USMP, vol 5(2), 2005.
- [14] RAMÍREZ-BELLO, JULIÁN; VARGAS-ALARCÓN, GILBERTO; TOVILLA-ZÁRATE, CARLOS Y FRAGOSO-JOSÉ MANUEL. *Polimorfismos de un solo nucleótido (SNP): implicaciones funcionales de los SNP reguladores (rSNP) y de los SNP-ARN estructurales (srSNP) en enfermedades complejas*, Gaceta Médica de Mexico, 149, 220-8, 2013.
- [15] RAJ, RESAL; RAMTEKE, PRAMOD. *Polymorphisms of Calpain 10 (Capn10) In Type 2 Diabetes Mellitus – A Review*, International Journal of Scientific and Research Publications, Volume 2, Issue 11, November 2012
- [16] SALAMANCA-GÓMEZ, FABIO. *Un nuevo gen de predisposición a la diabetes tipo 2*, Gaceta Médica de la Academia nacional de medicina-México, 137(1), 2001.
- [17] SÁNCHEZ, R., MORGADO, E., GRAU, R. *A Genetic Code Boolean Structure. I. The Meaning of Boolean Deductions*. Bull. Math. Biol. (article in press) doi:10.1016/j.bulm.2004.05.005 (2004)
- [18] SANCHEZ, R., PORFETTI, L.A., GRAU, R., Y MORGADO, E.R. *A New DNA Sequences Vector Spaces a Genetic Code Galois Field*. Match Commun. Math Comput. Chem. Vol 54, 1-3-28(2005b).
- [19] SANCHEZ, R., MORGADO, E., GRAU, R. *Gene Algebra from a Genetic Code Algebraic structure*. J. Math Biol. 51, 431-457(2005c).
- [20] SANCHEZ, R., BARRETO, J., MORGADO, E. Y GRAU R. *Abelian Finite Group of DNA Genomic Sequences*, MATCH Commun. Math Comput. Chem., 54(3), 2005.
- [21] VALERO, C.; HERNANDEZ-CHICO, C. *Comprendiendo una enfermedad hereditaria: conceptos básicos de genética* Asociación española de neurofibromatosis, Madrid, 2011.