



UNIVERSIDAD POLITÉCNICA DE PUEBLA

PROGRAMA ACADÉMICO DE
INGENIERÍA EN INFORMÁTICA

**Sistema Médico-Humano para el Almacenamiento y
Consulta de Enfermedades Genéticas**

Oscar Osvaldo Cortés González

Reporte Técnico PII-12-08-09

COMITÉ EVALUADOR

M.C. Javier Caldera Miguel (*Asesor*)

Dr. Jorge de la Calleja Mora (*Sinodal*)

Dra. María Auxilio Medina Nieto (*Sinodal*)

PROFESOR(A) DE PROYECTO DE INVESTIGACIÓN II

Dra. María Auxilio Medina Nieto

Juan C. Bonilla, Puebla

Agosto 2009

ÍNDICE

Capítulo 1. Planteamiento del problema

1.1 Introducción.....	5
1.2 Objetivo general.....	7
1.3 Objetivos específicos.....	7
1.4 Justificación.....	8
1.5 Cronograma de actividades.....	9
1.6 Recursos.....	10
1.7 Alcances y limitaciones.....	11

Capítulo 2. Marco teórico

2.1 La ciencia de la genética.....	12
2.1.1 Concepto e historia de la genética.....	12
2.1.2 Definición de gen.....	14
2.1.3 Genotipo y fenotipo / transmisión y expresión.....	15
2.1.4 Enfermedades genéticas.....	18
2.2 Ámbitos de la genética.....	23
2.2.1 Genética humana.....	23
2.2.2 Genética médica.....	24
2.2.3 Genética molecular (biología molecular).....	25
2.3 Ácido desoxirribonucleico (ADN).....	26
2.3.1 Definición de ácido desoxirribonucleico (ADN).....	26
2.3.2 Estructura del ADN.....	26
2.3.3 Estructuras en doble hélice.....	28
2.3.4 Modificaciones químicas en el ADN (modificaciones de bases).....	28
2.3.5 Daños en el ADN.....	28
2.3.6 Ácido ribonucleico mensajero.....	29
2.3.6.1 Procesamiento del ARN mensajero en células procariotas...	29
2.3.6.2 ARN mensajero en células procariotas.....	31
2.4 Cromosoma.....	31
2.4.1 Definición de cromosoma.....	31
2.4.2 Elementos en la estructura cromosómica.....	32
2.4.3 Estructura externa de los cromosomas (número, forma y tamaño)...	34
2.4.4 Número de cromosomas.....	34
2.4.5 Cromosomas sexuales.....	35
2.4.6 Forma de los cromosomas.....	35
2.4.7 Cromosomas Humanos.....	36
2.4.8 Tipos especiales de cromosomas.....	37
2.5 Bases de Datos (BD).....	37
2.5.1 Concepto de base de datos.....	37
2.5.2 Tipos de base de datos.....	38
2.5.3 Modelos de bases de datos.....	38
2.6 Trabajo relacionado.....	39

Capítulo 3 Diseño de investigación

3.1 Introducción.....	41
3.2 Documento de requerimientos.....	41
3.3 Modelo Entidad – Relación (E-R).....	42
3.4 Diccionario de datos.....	42
3.5 Casos de uso.....	44
3.6 Especificación de actores.....	47
3.7 Diagrama conceptual.....	47
3.8 Datos de entrada, datos de salida y descripción por módulo.....	48
3.9 Recolección de datos.....	49
3.9.1 Manejo de datos.....	50
3.9.2 Representación de resultados.....	50

Capítulo 4. Implementación

4.1 Implementación de las funciones Consultar, Modificar y Eliminar.....	51
4.2 Pantalla principal del sistema.....	52
4.2.1 Descripción de la barra de menús.....	53
4.2.2 Ejemplos de la barra de menús.....	54

Capítulo 5. Pruebas..... 57

Capítulo 6. Conclusiones..... 64

Referencias..... 65

Sistema Médico-Humano para el almacenamiento y consulta de enfermedades genéticas.

Resumen

Se desarrolló un sistema para almacenar y consultar enfermedades genéticas, este sistema fue probado por especialistas en genética. La información que se manejó fue acerca de los 23 pares de cromosomas, mostrando información e imágenes que describen las alteraciones o deformaciones de los 23 pares de cromosomas, el sistema cuenta con las opciones de Agregar, Consultar, Modificar, y Eliminar información acerca de los cromosomas.

Este sistema se probó y se validó por personal de área de genética, quienes realizaron pruebas con este sistema. Una de las pruebas que se realizaron fue con el par de cromosoma número 22, y aunque los resultados obtenidos fueron exitosos se pretende que el sistema trabaje de forma inteligente.

Este sistema será una herramienta útil que se aportará a la medicina específicamente en el área de la genética para la realización de consultas de enfermedades en los pares cromosómicos.

Capítulo 1. Planteamiento del problema

1.1 Introducción

Actualmente, es posible localizar, identificar, amplificar y secuenciar los genes que codifican las proteínas esenciales y los elementos reguladores de su expresión. También es factible expresar genes en sistemas apropiados para analizar su función o producir grandes cantidades de las proteínas que codifican.

La genética es la ciencia que se ocupa del estudio de la variación y de la herencia de todos los organismos vivos. El término “genética” fue propuesto en 1905 por William Bateson. El término de “gen” fue utilizado por primera vez en 1909 por Wilhelm Johansen para referirse a las unidades Mendelianas de la herencia [1].

La definición de genética como ciencia que se encarga del estudio de la variación y herencia de todos los seres vivos es muy amplia. Por esta razón se abarcarán dos áreas específicas de la genética: genética humana y genética médica. La genética humana es el estudio de la variación y herencia en los humanos, y la genética médica se ocupa del estudio de la variación genética humana con significado médico [2]. Además de estas dos áreas se necesita la Genética Molecular como apoyo complementario para este proyecto.

Actualmente se pueden desarrollar diversas herramientas para dar solución a gran variedad de problemas mediante soluciones informáticas (software). Con el propósito de realizar un software que puede dar respuesta a consultas relacionadas con las diferentes enfermedades genéticas, se reunirá información a cerca de anomalías en los genes humanos, esta

información se almacenará en un sistema médico-humano el cual será la base para almacenar la información de los genes humanos para posteriormente hacer consultas, todo esto gracias a que se tendrá almacenada y actualizada la información de los genes humanos.

1.2 Objetivo General

Construir un sistema médico-humano para el almacenamiento y consulta de enfermedades genéticas.

1.3 Objetivos específicos

- Elaborar descripciones de genes humanos los cuales sean afectados por enfermedades genéticas.
- Implementar un sistema de almacenamiento de información de los cromosomas del ser humano.
- Mostrar características de anomalías en los genes a través de la información almacenada en el sistema

1.4 Justificación.

Ante el crecimiento tecnológico en diversas áreas de la medicina, es necesario desarrollar herramientas informáticas para la identificación eficaz de enfermedades, sobre todo en el área genética ya que ahí se pueden encontrar la gran mayoría de las enfermedades y esto es por el gran avance científico en esta rama de la biología. Con el propósito de contribuir en un sistema de detección de enfermedades se almacenarán y se consultarán enfermedades

cromosómicas, donde se construirá un sistema médico-humano. Es decir, esta herramienta de almacenamiento permitirá manipular la información cromosómica, minimizando el tiempo en la obtención de datos cromosómicos, para que posteriormente se encuentren patrones que definan a una enfermedad y así se pueda detectar de forma inteligente.

1.5 Cronograma de Actividades

Tareas	Tiempo (semanas)															
	Septiembre				Octubre				Noviembre				Diciembre			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1. Elaboración de la propuesta de investigación	■	■														
2. Revisión de la propuesta de investigación		■	■													
3. Revisión literaria			■	■	■	■	■	■	■	■	■	■	■	■	■	■
4. Presentación de la propuesta de investigación			■	■	■											
5. Elaboración del marco teórico	■	■	■	■	■	■	■	■	■	■	■	■				
6. Investigar genes humanos						■	■	■	■	■	■	■	■	■	■	■
7. Investigar genotipo y fenotipo						■	■									
8. Investigar enfermedades genéticas							■	■	■							
9. Investigar ámbitos de la genética							■	■	■	■	■	■				
10. Revisión literaria acerca de genes humanos						■	■	■	■	■	■	■	■	■	■	■
11. Tratamiento de la información									■	■	■	■	■	■	■	■
12. Investigar tipos de base de datos												■				
Tareas	Tiempo (semanas)															
	Enero				Febrero				Marzo				Abril			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
13. Desarrollo del sistema	■	■	■	■	■	■	■	■								
14. Implementación del sistema									■	■						
15. Pruebas del sistema											■	■	■			
16. Análisis de los resultados												■	■	■		
17. Elaboración del reporte de investigación													■	■	■	■

1.6 Recursos

Recursos de hardware

- Una computadora de escritorio con procesador Intel Pentium IV a 1.4 GHz-
- 512 de memoria RAM.
- Unidad de escritura y lectura de CD.
- 2 puertos USB.
- 1 disco duro con 20 GB disponibles
- 1 memoria USB de 1 GB.

Software

- Sistema operativo Windows XP profesional.
- Navegador web Mozilla Fire Fox versión 2.0
- Paquetería Microsoft Office 2007.
- Visualizador de archivos pdf Adobe Reader 8.0
- Descompresor de archivos Winrar 2.0.
- Manejador de Base de Datos Access 2003

1.7 Alcances y limitaciones.

- El presente trabajo sólo reunirá información concerniente a los cromosomas humanos para almacenar y consultar enfermedades cromosómicas.
- Sólo se podrá almacenar información del cromosoma número 21 ya que posteriormente se recopilará la información de los restantes cromosomas humanos.
- Este sistema de almacenamiento permitirá manipular la información cromosómica que caracteriza a enfermedades hereditarias.

Capítulo 2. Marco teórico

2.1 La ciencia de la Genética.

2.1.1 Concepto e historia de la genética

Se podría decir que la revolución más importante en la biología moderna es la llamada "Revolución Genética". El impacto del conocimiento íntimo de nuestra información genética ha transformado casi todos los aspectos de la actividad humana y en especial a la medicina [3].

Hay evidencia a través de relieves esculpidos en la roca que en el antiguo Egipto el cultivo de plantas y la doma de animales eran actividades cotidianas (2700-2200 a. de JC). También han aparecido pinturas de una antigüedad de 8.000 años (6.000 a. de JC) en Argelia que muestran la conducción de ganado. Posiblemente estas actividades debieron requerir el reconocimiento de las características deseables y su selección. La polinización de palmeras datileras (883-859 a. de JC) en gravados de Egipto, indica conocimientos detallados de la historia natural conducente a la fertilización. Unos gravados hallados en Caldea (Irak) de hace unos 6.000 años ilustran pedigríes en los que se documenta la transmisión de la crin de los caballos. El reconocimiento de las características en los animales y plantas debió favorecer el reconocimiento de las características humanas y cómo algunas de éstas se transmitían a la descendencia [4].

Durante la civilización griega surgieron fundamentalmente 3 ideas referentes a las leyes de la reproducción y herencia cuya influencia se extendió hasta el siglo XIX: la de la pangénesis, la de la epigénesis y la del preformacionismo. La pangénesis postulaba que el semen se formaba como suma de pequeñas partículas procedentes de todas las partes del

cuerpo que, circulando por la sangre, llegaban hasta el testículo. Estas partículas, representativas de los rasgos de cualquier parte del cuerpo, se transmitían durante el acto sexual a la descendencia. La epigénesis establecía que los órganos del adulto no existen al principio, sino que se forman durante el desarrollo. Esta idea era contraria a la del preformacionismo, que establecía la existencia de un homúnculo dentro del espermatozoide que contenía todos los órganos ya formados. La influencia de las ideas iniciadas en la antigua Grecia quedó incluso reflejada en la teoría de la evolución de los caracteres adquiridos formulada por J.B. Lamarck. Según esta teoría, los caracteres adquiridos podrían transmitirse a futuras generaciones. Según [1] la siguiente tabla muestra los hechos más importantes en el campo de la genética.

Tabla 2.1. Hechos importantes en el campo de la genética [1].

Año	Autores Principales	Descripciones, Descubrimientos o hechos.
1859	Charles Darwin	Origen de las especies de Darwin
1865	Gregory Mendel	Herencia unitaria. Distribución independiente. Segregación.
1869	Friedrich Miescher	Aislamiento del DNA por primera vez
1902	Walter Sutton	Teoría cromosómica de la herencia
1905	William Bateson	Término de genética
1909	Wilhelm Johansen	Términos de GEN, genotipo y fenotipo
1952	Alfred Hersey, Martha Chase	Los genes están hechos de DNA
1953	Francis Crick, James Watson	DNA es una doble hélice.
1955	Joe Hin Tjio	El hombre tiene 46 cromosomas
1961	Brenner , Jacob , Meselson	El mRNA porta información
1966	Nirenberg, Khorana, Ochoa	Código genético. Síntesis DNA y RNA.

2.1.2 Definición de Gen

“Wilhelm Johansen utiliza por primera vez el término de gen para referirse a las unidades mendelianas de herencia. También se distingue entre genotipo (lo heredado) y fenotipo (la

apariciencia externa)” [1]. La genética es el campo de la biología que se encarga del estudio de la herencia, es decir, el modo en que los rasgos y las características se transmiten de padres a hijos (como el color de los ojos y una posibilidad mayor de contraer una enfermedad determinada).

El objeto de estudio de la genética son los patrones de herencia. Los genes están formados por segmentos de una molécula que codifica la información genética en las células denominada ácido desoxirribonucleico (ADN).

La genética se subdivide en distintas ramas [3]: la clásica o mendeliana (que se centra en el estudio de los cromosomas y los genes y en cómo se heredan de generación en generación), la cuantitativa (analiza el impacto de múltiples genes sobre el fenotipo), la molecular (estudia el ADN, su composición y la forma en que se duplica), la genética de poblaciones y evolutiva (centra su atención en el comportamiento de los genes en una población y de cómo esto determina la evolución de los organismos) y la genética del desarrollo (cómo los genes que controlan el desarrollo de los organismos).

2.1.3 Genotipo y fenotipo / transmisión y expresión

La genética debe su existencia al hecho genético: los organismos son portadores de información codificada. Esto, que hoy parece obvio, en su momento fue extraño, anti-intuitivo. La revolución informática y la teoría de la información no habían mostrado la lógica e "intuitividad" de estos aspectos: no hay nada en sistemas no vivos excepto los hechos por el hombre que se corresponda con el genotipo [5].

Fue Mendel el primero en captar la naturaleza dual de los organismos, su dicotomía entre su genotipo y su fenotipo (a pesar que estos conceptos fueron introducidos por el danés W.

Johannsen en 1911). Lo esencial del mendelismo fue el percatarse de la ruptura, nunca antes clara, entre el proceso de herencia y el proceso de desarrollo, entre transmisión y expresión. Se heredan un conjunto de factores internos, los genes, y el estado genético interno de cada individuo. Su *genotipo* es una consecuencia de las leyes dinámicas que regulan el paso de estas entidades de padres a hijos. Las dos leyes de la herencia son leyes de transmisión, no hacen ninguna referencia a la apariencia del organismo, (el *fenotipo*). El fenotipo, con respecto a la herencia, es un epifenómeno sin interés, pues éste resulta de un proceso causal diferente: el proceso epigenético de la ontogenia, que depende del estado de los genes pero no de las leyes de su herencia (Lewontin 1992).

Luego entonces se define que genotipo y fenotipo es [6]:

Fenotipo: la clase de la que se es miembro según las cualidades físicas observables en un organismo, incluyendo su morfología, fisiología y conducta a todos los niveles de descripción.

Genotipo: La clase de la que se es miembro según el estado de los factores hereditarios internos de un organismo, sus genes.

El fenotipo y el genotipo se identifican a un sólo nivel: el del ADN. Por primera vez en la historia ahora el genotipo también es fenotipo, es un carácter observable, expresión de la realidad material del genotipo. “Un conocimiento completo del sistema genético requiere conocer cómo el genotipo se relaciona con el fenotipo y cómo el fenotipo a su vez se relaciona con el genotipo (puesto que las leyes que van de genotipo a fenotipo no tienen necesariamente que ser las mismas leyes que van de fenotipo a genotipo, como lo muestra, la existencia de dominancia y la redundancia de código), por último, el genotipo parental

llega a convertirse en genotipos hijos, ver la Figura 1 que muestra la relación entre los espacios fenotípicos y genotípicos. Mientras que este último proceso prácticamente está resuelto, sólo existe un conocimiento limitado de las rutas causales de los otros procesos [7]”.

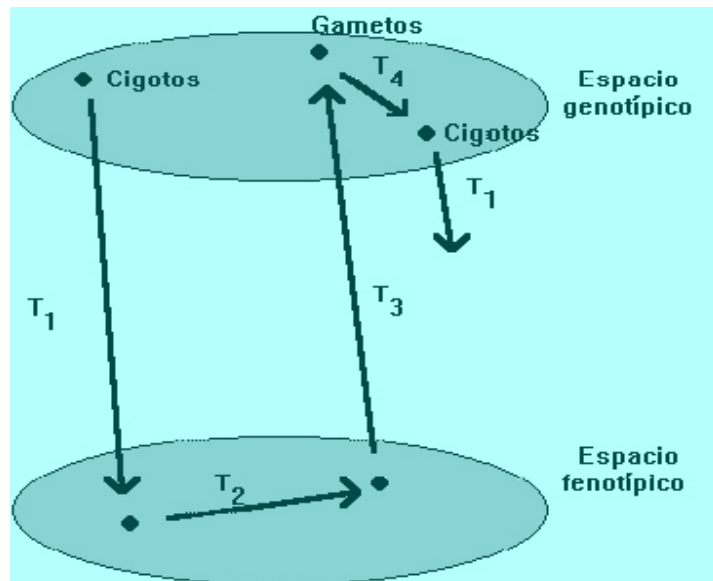


Figura 2.1. Relación entre los espacios fenotípicos y genotípicos [7].

T1, T2, T3 y T4 son las leyes de transformación requeridas para obtener una descripción completa de ambos espacios entre generaciones. Según [Lewontin 1971], T1 son las leyes epigenéticas que convierten los cigotos en fenotipos adultos. T2 son las leyes de apareamiento, migración y selección natural que modifican la distribución de los fenotipos dentro de una generación. T3 simboliza las reglas que permiten inferir la distribución de genotipos a partir de los diferentes fenotipos adultos. T4 son las leyes genéticas de Mendel y Morgan (las leyes de transmisión) que permiten predecir, a partir de la distribución de los genotipos parentales, la distribución de los genotipos en la próxima generación [8].

La relación entre el fenotipo y el genotipo es compleja debido a, las relaciones entre alelos de un gen llamadas relaciones de dominancia y las interacciones entre genes, otras relaciones se generan por la secuencia de ambientes por los que pasa cada genotipo durante su desarrollo [9]. La descripción del fenotipo de un individuo tiene una dimensión temporal. Cuando el fenotipo se describe a un nivel próximo del genotipo, el componente de interacción entre genes y el ruido asociado al desarrollo es menor y puede determinarse con más claridad las relaciones entre ambos niveles, el nivel de descripción más bajo posible: es el nivel del ADN. La secuencia de un gen determina completamente su genotipo y, es posible inferir el fenotipo del genotipo. El nivel inmediatamente superior, el ácido ribonucleico (ARN) mensajero, presenta componentes de elaboración de mensajes, tales como edición o procesamiento del ARN. El siguiente nivel, la proteína especificada por los genes, presenta una relación exhaustiva (de uno a muchos) debido a la degeneración del código, además en este nivel hay, una modificación de la estructura secundaria y terciaria bajo la influencia de genes distintos al que especifica la proteína. La división, la migración y la diferenciación celular que sigue a la síntesis proteica introducen un número creciente de interacciones, incrementando con lo anterior una mayor contingencia a las relaciones entre el fenotipo y el genotipo.

2.1.4 Enfermedades genéticas

Existen varios tipos de enfermedades genéticas. El modo en el que se hereda el trastorno puede ayudar a determinar los riesgos en un embarazo y el riesgo de que vuelva a recurrir en futuros hijos. Los riesgos de tener un bebé con un defecto congénito debido a una anomalía genética pueden aumentar cuando:

- Los padres tienen otro hijo con un trastorno genético.
- Existen antecedentes familiares de trastornos genéticos.
- Uno de los padres tiene una anomalía cromosómica.
- El feto presenta anomalías que se detectan en la ecografía

Los diversos tipos de enfermedades genéticas con las que el sistema trabajará son:

- Anomalías cromosómicas
- Defectos de un único gen
- Problemas multifactoriales
- Problemas teratogénicos

a) Anomalías cromosómicas

Las anomalías cromosómicas pueden ser heredadas de los padres o pueden aparecer sin que haya antecedentes familiares. Las anomalías cromosómicas más frecuentes son las siguientes [10]:

- 1 Aneuploidia.- Cantidad de cromosomas superior o inferior a la normal, que incluye síntomas como:
 - *El síndrome de Down (trisomía 21)*. Las células contienen tres cromosomas 21
 - *El síndrome de Turner*. Uno de los cromosomas sexuales no se transfiere, dejando un cromosoma X solo, o un total de 45 cromosomas x's.
- 2 Deleción.- Falta parte de un cromosoma o parte del código de ADN.

- 3 Inversión.- Se presenta cuando un cromosoma se desprende y la parte del cromosoma que se desprende se invierte y se vuelve a insertar. Las inversiones pueden causar defectos congénitos o no, según su estructura exacta.
- 4 Translocación.- El término translocación se utiliza cuando se presentan modificaciones en la ubicación de determinado material cromosómico.
 - o *Translocación recíproca*.- El ADN se intercambia equitativamente entre cromosomas sin que se agregue o se pierda ninguno. Un padre con una translocación equilibrada es sano, pero corre el riesgo de transmitir cromosomas no equilibrados en el embarazo.
 - o *Translocación robertsoniana*.- Un translocación balanceada en la que un cromosoma se une al extremo de otro.
- 5 Mosaicismo.- Presencia de dos o más patrones cromosómicos en las células de un individuo, que origina dos o más líneas celulares (por ejemplo, algunas con 46 cromosomas y otras con 47).

b) Defectos de un único gen

Los defectos de un único gen se conocen como trastornos hereditarios mendelianos, debido al primer trabajo en genética de Gregory Mendel. En estos trastornos, un sólo gen es responsable de un defecto o anomalía. Los trastornos de un único gen normalmente tienen mayores riesgos de ser heredados y pueden ser [11]:

- 1 Dominante. Se produce una anomalía cuando sólo uno de los genes de uno de los padres es anormal. Si el padre tiene el trastorno, el bebé tiene un 50% de posibilidades de heredarlo. Algunos ejemplos son:

- *Acondroplasia*.- Desarrollo imperfecto de los huesos que causa el enanismo.
 - *Síndrome de Marfan*.- Trastorno del tejido conectivo que provoca extremidades largas y defectos cardíacos.
- 2 Recesivo.- Sólo se produce una anomalía cuando ambos padres tienen genes anormales. Si ambos padres son portadores, el bebé tiene un 25% de posibilidades de tener el trastorno. Como ejemplo se incluye lo siguiente:
- *Fibrosis quísticas*.- Trastorno glandular que produce exceso de mucus en los pulmones y problemas en la función pancreática y la absorción de los alimentos.
 - *Anemia drepanocítica*.- Trastorno que produce glóbulos rojos anormales.
 - *Enfermedad de Tay Sachs*.- Trastorno autosómico recesivo hereditario que produce la degeneración progresiva del sistema nervioso central, con resultados fatales (normalmente alrededor de los 5 años de edad).
- 3 Trastorno ligado al cromosoma X. El trastorno está determinado por los genes del cromosoma X. Los principales afectados y quienes tienen el trastorno son los hombres. Las hijas de hombres que sufren el trastorno son portadoras del rasgo y tienen una posibilidad cada dos de transferirlo a sus hijos. Los hijos varones de las mujeres portadoras tienen una posibilidad cada dos de tener el trastorno. Como ejemplo se incluye lo siguiente:
- *Distrofia muscular de Duchenne*.- Enfermedad de debilidad y desgaste muscular.
 - *Hemofilia*.- Trastorno hemorrágico causado por bajos niveles, o ausencia, de una proteína de la sangre que es esencial para la coagulación.

c) **Problemas multifactoriales.**- “Algunos defectos congénitos no siguen el patrón de un único gen ni de anomalía cromosómica sino que se deben a varios problemas, o al efecto combinado de los genes y el ambiente. Es difícil predecir la herencia de anomalías causadas por factores múltiples” [12]. Algunos ejemplos son los defectos cardíacos, el labio leporino o el paladar hendido y los defectos del tubo neural (defectos en la columna o en el cerebro).

d) **Problemas teratogénicos.**- Se sabe que ciertas sustancias producen anomalías en los bebés. Muchos defectos congénitos se producen cuando el feto es expuesto a teratógenos (sustancias que causan anomalías) durante el primer trimestre del embarazo, cuando los órganos están en formación [13]. Algunos teratógenos conocidos son los siguientes:

- Algunos medicamentos
- El alcohol
- Exposición a altos niveles de radiación
- El plomo
- Determinadas infecciones (como la rubéola)

Los problemas genéticos se diagnostican en las familias con riesgo de sufrir enfermedades genéticas pueden consultar a un asesor especializado en genética. Este especialista se apoyará en la citogenética, ésta es una rama de la genética que se ocupa del estudio de los cromosomas, También puede ser necesario controlar el ADN de cada padre para conocer algunos patrones genéticos hereditarios, o en ocasiones se dispone de pruebas prenatales para detectar problemas en el feto. Las pruebas pueden incluir las ecografías (utilizar ondas sonoras para observar las estructuras internas), el muestreo de vellosidades coriónicas (examinar los tejidos que rodean al feto), o la amniocentesis (tomar una muestra del líquido amniótico) [2].

2.2 Ámbitos de la genética

2.2.1 Genética humana

Cada ser humano está formado por trillones de células. En cada una de estas células existen 46 cromosomas: 44 cromosomas autosómicos y 2 cromosomas sexuales (XX en las mujeres y XY en los varones). Los 46 cromosomas corresponden a 23 aportados por el espermatozoide y 23 aportados por el óvulo, de modo que al ocurrir la fecundación se constituyen los 46 cromosomas del cigoto y de todas las células derivadas del cigoto inicial.

En los cromosomas residen los genes, que corresponden a las unidades de la herencia. La información genética se encuentra codificada en pequeños trozos de la molécula de ADN. Los genes poseen una secuencia específica con una función particular. Generalmente, esta secuencia génica específica determina una función específica, como por ejemplo, la

formación de una proteína que cumple un rol específico en las complejas vías metabólicas que presentan las diferentes células de nuestro organismo.

La genética humana es la disciplina biológica que se preocupa de la manera cómo se transmiten (herencia) los caracteres de padres a hijos a lo largo de las generaciones, y de las semejanzas y diferencias entre padres e hijos, que son determinadas por la herencia y el ambiente [1].

2.2.2 Genética médica

La genética médica es la aplicación de la genética en la medicina. La genética médica tiene amplios y variados campos de aplicación, puede incluir campos individuales como la genética clínica, la genética bioquímica (la genética bioquímica es un híbrido de la bioquímica principalmente de aminoácidos y proteínas y la genética), citogenética (la citogenética es un híbrido de la citología y la genética; incluye el estudio de cromosomas bajo el microscopio), genética molecular (la genética molecular es un híbrido de la bioquímica del ADN y ARN con la genética), la genética de enfermedades comunes. A cualquier otro campo de genética médica se le considera híbrido de la medicina clínica y medicina genética [10].

2.2.3 Genética molecular

La biología molecular (o biología molecular) se basa en estudiar la estructura de genes, de su expresión y del control de su expresión. La biología molecular está condicionada a trabajar, principalmente, con moléculas de ácido desoxirribonucleico (DNA) y de ácido ribonucleico mensajero (mRNA). Las aplicaciones de la biología molecular son múltiples y muy prometedoras. Estas aplicaciones pueden ser utilizadas en un laboratorio de

investigación para estudiar la estructura y la expresión de un gen (o de una parte de un gen), en la industria para conseguir producir, a partir de células, proteínas útiles para el hombre, o incluso en un laboratorio de biología clínica para efectuar algunos diagnósticos. Además, se espera mucho de la terapia genética [14].

2.3 Ácido desoxirribonucleico (ADN)

2.3.1 Definición de ácido desoxirribonucleico (ADN).

El ADN es un polímero de nucleótidos (en otras palabras un polinucleótido). Un polímero es un compuesto constituido por muchas unidades simples conectadas entre sí, como semeja una cadena formada por varios eslabones. En el ADN. Cada eslabón es un nucleótido y cada uno de estos nucleótidos está formado por azúcar (desoxirribosa), una base nitrogenada (adenina→A, timina→T, citosina→C o guanina→G) y finalmente un grupo de fosfato que actúa como enganche de cada eslabón con el siguiente. Lo que diferencia a cada eslabón (nucleótido) de otros es, la base nitrogenada y por consiguiente la secuencia del ADN se define nombrando sólo la secuencia de sus bases. En los organismos vivos, el ADN se presenta como una doble cadena de nucleótidos, en la que las dos hebras están unidas entre sí por unas conexiones denominadas puentes de hidrogeno [15].

2.3.2 Estructura del ADN

El ADN es una molécula formada por dos cadenas dispuestas de forma antiparalela y con las bases nitrogenadas enfrentadas. En su estructura tridimensional, se distinguen distintos niveles [16]:

1. Estructura primaria:

Secuencia de nucleótidos encadenados. En estas cadenas es donde se encuentra la información genética, y dado que el esqueleto es el mismo para todos, la diferencia de la información radica en la distinta secuencia de bases nitrogenadas. Esta secuencia presenta un código, que determina una información u otra, según el orden de las bases.

2. Estructura secundaria:

Es una estructura en doble hélice. Permite explicar el almacenamiento de la información genética y el mecanismo de duplicación del ADN. Es una cadena doble, dextrógira o levógira, según el tipo de ADN. Ambas cadenas son complementarias, pues la adenina y la guanina de una cadena se unen, respectivamente, a la timina y la citosina de la otra. Ambas cadenas son antiparalelas, pues el extremo 3' de una se enfrenta al extremo 5' de la homóloga.

3. Estructura terciaria:

Se refiere a cómo se almacena el ADN en un espacio reducido, para formar los cromosomas. Varía según se trate de organismos procariotas o eucariotas:

1. En procariotas el ADN se pliega como una súper-hélice, generalmente en forma circular y asociada a una pequeña cantidad de proteínas.

2. En eucariotas, dado que la cantidad de ADN de cada cromosoma es muy grande, el empaquetamiento ha de ser más complejo y compacto; para ello se necesita la presencia de proteínas, como las histonas y otras proteínas de naturaleza no histónica (en los espermatozoides estas proteínas son las protaminas).

2.3.3 Estructuras en doble hélice

El ADN posee demasiadas conformaciones. Sin embargo, en organismos vivos sólo se han observado las conformaciones ADN-A, ADN-B y ADN-Z. La clasificación que adopta el ADN depende de su secuencia, la cantidad y dirección de super-enrolamiento que presentan la presencia de modificaciones químicas en las bases y las condiciones de la solución. De las tres conformaciones, la forma "B" es la más común en las condiciones existentes en las células. Las dos dobles hélices alternativas del ADN difieren en su geometría y dimensiones. La forma "A" es una espiral que gira hacia la derecha más amplia que la "B", con una hendidura menor superficial y más amplia, y una hendidura mayor más estrecha y profunda. La forma "A" ocurre en condiciones no fisiológicas en formas deshidratadas de ADN [16].

2.3.4 Modificaciones químicas en el ADN (modificaciones de bases)

La expresión de los genes está influenciada por la forma en la que el ADN está empaquetado en cromosomas, en una estructura denominada cromatina. Las modificaciones de bases pueden estar implicadas en el empaquetamiento del ADN: las regiones que presentan una expresión génica baja o nula normalmente contienen niveles altos de metilación de las bases citosina [17].

2.3.5 Daño en el ADN

El ADN puede resultar dañado por muchos tipos de mutágenos, que cambian la secuencia del ADN tales como: agentes alquilantes, radiación electromagnética de alta energía. Por lo tanto el daño producido en el ADN depende del tipo de mutágeno. El daño en el ADN inicia una respuesta que activa diferentes mecanismos de reparación que reconocen lesiones específicas en el ADN, que son reparadas en el momento para recuperar la secuencia

original del ADN. Asimismo, el daño en el ADN provoca una parada en el ciclo celular, que conlleva la alteración de numerosos procesos fisiológicos, que a su vez implica síntesis, transporte y degradación de proteínas. Si el daño genómico es demasiado grande para que pueda ser reparado, los mecanismos de control inducirán la activación de una serie de rutas celulares que culminarán en la muerte celular [17].

2.3.6 Ácido ribonucleico mensajero

El ácido ribonucleico mensajero (ARN mensajero o ARNm) contiene información genética procedente del ADN que utiliza en la síntesis de proteínas (determina el orden en que se unirán los aminoácidos). El ARN mensajero es un ácido nucleico monocatenario, al contrario que el ADN que es bicatenario [18].

2.3.6.1 Procesamiento del ARN mensajero en células eucariotas

El ARN precursor (pre-ARN) o transcrito primario es obtenido del ARN mensajero después de la transcripción en la mayoría de los casos no se libera del complejo de transcripción en forma totalmente activa, sino que ha de sufrir modificaciones antes de ejercer la función de procesamiento o maduración del ARN. Dentro de estas modificaciones se encuentra la eliminación de fragmentos, la adición de nuevos fragmentos no codificados en el DNA y la modificación covalente de algunas bases nitrogenadas. Específicamente, el procesamiento del ARN en eucariotas comprende diferentes fases [19]:

1. Adición al extremo 5' de la estructura denominada caperuza o casquete que es un nucleótido modificado de guanina. Esta caperuza es necesaria para el proceso normal de traducción del ARN y para mantener su estabilidad; esto es crítico para el reconocimiento y el acceso apropiado del ribosoma.

2. Poliadenilación: es la adición al extremo 3' de una cola poli-A, una secuencia larga de poliadenilato, es decir, un tramo de RNA cuyas bases son todas adenina. Esta cola protege al ARNm frente a la degradación, aumentando su vida media en el citosol, de modo que se puede sintetizar mayor cantidad de proteína.

3. Cuando el ARN mensajero sufre la eliminación de secuencias internas, no codificantes, llamadas intrones, el proceso de retirada de los intrones y conexión o empalme de los exones se llama ajuste, o corte y empalme. A veces un mismo transcrito primario o pre-ARNm se puede ajustar de diversas maneras, permitiendo que con un solo gen se obtengan varias proteínas diferentes; a este fenómeno se le llama ajuste alternativo.

4. El ARN mensajero maduro es trasladado al citosol de la célula, en el caso de los seres eucariontes, a través de poros de la membrana nuclear.

5. Una vez en el citoplasma, el ARNm se acoplan los ribosomas, que son la maquinaria encargada de la síntesis proteica. En procariontes, la unión de los ribosomas ocurre mientras la cadena de ARNm está siendo sintetizada.

6. Después de cierta cantidad de tiempo el ARNm se degrada en sus nucleótidos componentes, generalmente con la ayuda de ribonucleasas.

2.3.6.2 ARN mensajero en células procariotas

El proceso de transcripción y el de traducción se realizan de manera similar que en las células eucariotas sin embargo, la diferencia fundamental radica en que, en las procariotas, el ARN mensajero no pasa por un proceso de maduración y, por lo tanto, no se le añade

caperuza ni cola, ni se le quitan intrones. Además no tiene que salir del núcleo como en las eucariotas, porque en las células procariotas no hay un núcleo definido [19].

2.4 Cromosoma

2.4.1 Definición de cromosoma

Se denomina cromosoma a cada uno de los pequeños cuerpos en forma de bastoncillos en que se organiza la cromatina del núcleo celular durante las divisiones celulares (mitosis y meiosis). La cromatina es un material microscópico que lleva la información genética de los organismos eucariotas y está constituida por ADN asociado a proteínas especiales llamadas histonas. Este material se encuentra en el núcleo de las células eucariotas y se visualiza como una maraña de hilos delgados. Cuando el núcleo celular comienza el proceso de división (cariocinesis), esa maraña de hilos inicia un fenómeno de condensación progresivo que finaliza en la formación de entidades discretas e independientes: los cromosomas. Por lo tanto, cromatina y cromosoma son dos aspectos morfológicamente distintos de una misma entidad celular [20].

2.4.2 Elementos en la estructura cromosómica

La organización de la cromatina no es uniforme a lo largo de la estructura del cromosoma. Por esta razón, se puede distinguir [21]: los centrómeros (o constricciones primarias), los telómeros (o extremos cromosómicos), las regiones organizadoras del nucléolo y los cromómeros.

Centrómetros

El centrómero es la constricción primaria que, utilizando tinciones tradicionales, aparece menos teñida que el resto del cromosoma. Es la zona por la que el cromosoma interacciona

con las fibras del huso acromático desde profase hasta anafase, tanto en mitosis como en meiosis, y es responsable de realizar y regular los movimientos cromosómicos que tienen lugar durante estas fases. Las estructuras centroméricas que interactúan con las fibras del huso se denominan cinetocoros. Además, el centrómero contribuye a la nucleación de la cohesión de las cromátidas hermanas. En la estructura del centrómero intervienen tanto el ADN centromérico, que consta fundamentalmente de heterocromatina constitutiva, como proteínas centroméricas.

Telómeros

Los telómeros son los extremos de los cromosomas. Son regiones de ADN no codificante, altamente repetitivas, cuya función principal es la estabilidad estructural de los cromosomas en las células eucariotas, la división celular y el tiempo de vida de las estirpes celulares. Además están involucradas en enfermedades tan importantes como el cáncer. En los organismos procariontes, los cromosomas son circulares y no poseen telómeros

Regiones organizadoras del nucléolo

Además de las constricciones primarias, en algunos cromosomas se puede distinguir otro tipo de "adelgazamiento" denominada constricción secundaria. Tales regiones se denominan regiones organizadoras del nucléolo (RON). Las secuencias de ADN ribosómico quedan englobadas dentro del nucléolo, que permanece adosado a las RONS durante buena parte del ciclo celular. Los cromosomas con regiones organizadoras dentro del nucléolo en muchos casos presentan un segmento que une a esta región con el telómero, el cual se denomina satélite o trabante

Cromómeros

Los cromómeros son engrosamientos o regiones más compactadas de la eucromatina, que se distribuyen de manera más o menos uniforme a lo largo de los cromosomas y se pueden visualizar durante las fases de la mitosis o de la meiosis de menor condensación de la cromatina. Su naturaleza molecular sigue siendo controvertida, pero podrían ser consecuencia de un cierto grado de compartimentalización en la distribución de las secuencias de ADN y en la organización de los cromosomas.

2.4.3 Estructura externa de los cromosomas (número, forma y tamaño)

El estudio de la estructura externa de los cromosomas de cualquier especie eucariótica consiste en analizar la forma, tamaño y número de los cromosomas que posee. El mejor momento para llevar a cabo este estudio puede ser aquel en el que los cromosomas han alcanzado su máximo grado de contracción y tienen sus bordes perfectamente definidos. Dicho momento suele ser la metafase mitótica. El estudio de la estructura externa de los cromosomas culmina con la obtención del cariotipo [22].

2.4.4 Número de cromosomas

Usualmente las especies animales y vegetales tienen un número de cromosomas constante y determinado que constituyen la ley de la constancia numérica de los cromosomas (cariotipo), aunque existen especies con una alta variabilidad cariotípica, no sólo en número sino en forma y tamaño de los cromosomas.

El número de cromosomas de una especie diploide se identifica como $2n$ mientras que ese número en una especie (o fase vital) haploide se identifica con la letra n . En aquellas

especies que presentan un número repetido de cromosomas superior a dos complementos se habla de poliploidía, representándose el múltiplo por delante de la letra n. Con la indicación x se quiere expresar el número básico de cromosomas de una especie que presenta individuos con diversos grados de ploidía o el de una línea filogenética a partir de la cual diversos taxones han alcanzado situaciones aneuploides variadas [10].

2.4.5 Cromosomas sexuales

En muchos organismos, uno de los pares de los cromosomas homólogos es distinto al resto, realizando la determinación del sexo del individuo. A estos cromosomas se les llama cromosomas sexuales o heterocromosomas e incluso gonosomas, porque determinan el sexo, por esta razón se tienen los siguientes sistemas de determinación [23]:

Sistema de determinación XY: es propio del ser humano y muchos otros animales. Las hembras, siendo XX, darán gametos iguales con cromosoma X, sexo homogamético y los machos, siendo XY, darán dos tipos de gametos, uno con el cromosoma X y otro con el cromosoma Y.

Sistema de determinación ZW: en otras especies ocurre lo contrario, el sexo masculino es homogamético (ZZ) y el femenino heterogamético (ZW).

Sistema de determinación XO: otras especies que no tienen el cromosoma Y, determinándose el sexo por el número de cromosomas X, macho XO y hembra XX.

2.4.6 Forma de los cromosomas

La forma de los cromosomas es para todas las células somáticas constante y característica de cada especie. La forma depende fundamentalmente de las constricciones que presente el cromosoma y de su localización en la cromátida. El cromosoma se encuentra constituido básicamente por el centrómero que divide el cromosoma en un brazo corto o brazo p y un brazo largo o brazo q. Algunos cromosomas presentan satélites en el brazo corto.

Según la posición del centrómero, los cromosomas se clasifican en [24]:

Metacéntricos: El centrómero se localiza a mitad del cromosoma y los dos brazos presentan igual longitud.

Submetacéntricos: La longitud de un brazo del cromosoma es algo mayor que la del otro.

Acrocéntricos: Un brazo es muy corto (p) y el otro largo (q).

Telocéntricos: Sólo se aprecia un brazo del cromosoma al estar el centrómero en el extremo.

2.4.7 Cromosomas humanos

El ser humano presenta 23 pares de cromosomas en sus células somáticas: 22 autosomas y un par de cromosomas sexuales (dos X en el caso de las mujeres y un cromosoma X y un Y en el caso de los varones). El tamaño total aproximado del genoma humano es de 3200 millones de pares de bases de ADN (3200 Mb) que contienen unos 20.000-25.000 genes. De las 3200 Mb unas 2950 Mb corresponden a eucromatina y unas 250 Mb a heterocromatina. La secuencia de ADN que conforma el genoma humano contiene

codificada la información necesaria para la expresión, altamente coordinada y adaptable al ambiente, del proteoma humano, es decir, del conjunto de proteínas del ser humano [1].

2.4.8 Tipos especiales de cromosomas

Existen algunos tipos de cromosomas presentes sólo en algunos tipos celulares o en poblaciones concretas de una especie. Entre ellos, destacan los cromosomas politénicos, en escobilla, cromosomas B e isocromosomas [25].

2.5 Bases de datos

2.5.1 Concepto de base de datos

Una base de datos es un “almacén” que permite guardar grandes cantidades de información de forma organizada para después encontrar y utilizar fácilmente la información. El término de bases de datos fue escuchado por primera vez en 1963, en un simposio celebrado en California, USA. Se define una base de datos como una serie de datos organizados y relacionados entre sí, los cuales son recolectados y explotados por los sistemas de información (empresa o particular). Desde el punto de vista informático, la base de datos es un sistema formado por un conjunto de datos almacenados en discos que permiten el acceso directo a ellos y un conjunto de programas que manipulen ese conjunto de datos. Cada base de datos se compone de una o más tablas que guarda un conjunto de datos. Cada tabla tiene una o más columnas y filas. Las columnas guardan una parte de la información sobre cada elemento que queramos guardar en la tabla, cada fila de la tabla conforma un registro [25].

2.5.2 Tipos de base de datos

Entre los diferentes tipos de base de datos, podemos encontrar los siguientes [26]:

- **MySql:** es una BD con licencia pública general (GPL) basada en un servidor. Se caracteriza por su rapidez. No es recomendable usar para grandes volúmenes de datos.
- **PostgreSql y Oracle:** Son sistemas de base de datos poderosos. Administra muy bien grandes cantidades de datos, y suelen ser utilizadas en intranets y sistemas de gran calibre.
- **Access:** Es una base de datos desarrollada por Microsoft.
- **Microsoft SQL Server:** es una base de datos más potente que access desarrollada por Microsoft. Se utiliza para manejar grandes volúmenes de informaciones.

2.5.3 Modelos de bases de datos

Entre los diferentes modelos de bases de datos podemos encontrar los siguientes [27]:

a) **Bases de datos jerárquicas.**- Estas son bases de datos que, como su nombre lo indican, almacenan su información en una estructura jerárquica. En este modelo los datos se organizan en una forma similar a un árbol (visto al revés), en donde un *nodo padre* de información puede tener varios *hijos*. El nodo que no tiene padres se le conoce como *raíz*, y a los nodos que no tienen hijos se les conoce como *hojas*. Una de las principales limitaciones de este modelo, es su incapacidad de representar eficientemente la redundancia de datos.

b) Bases de datos de red.- Este es un modelo ligeramente distinto del jerárquico, en donde su diferencia fundamental es la modificación del concepto de un *nodo*, permitiendo que un mismo nodo tenga varios padres (algo no permitido en el modelo jerárquico).

c) Bases de datos relacionales.- Este es el modelo más utilizado en la actualidad para modelar problemas reales y administrar datos dinámicamente. Cuando fue creado a finales de los años sesenta, no tardó en consolidarse como un nuevo paradigma en los modelos de base de datos. Su idea fundamental se basa en el concepto de "tablas", que a su vez se componen de *registros* (las filas de una tabla) y *campos* (las columnas de una tabla). En este modelo, el lugar y la forma en que se almacenen los datos no tienen relevancia (a diferencia de otros modelos como el jerárquico y el de red). Esto tiene la considerable ventaja de que es más fácil de entender y de utilizar para un usuario casual de la base de datos. La información puede ser recuperada o almacenada por medio de "consultas" que ofrecen una amplia flexibilidad y poder para administrar la información.

d) Bases de datos orientados a objetos.- Este modelo, es reciente, y propio de los modelos informáticos orientados a objetos, trata de almacenar en la base de datos los *objetos* completos (estado y comportamiento).

2.6 Trabajo Relacionado

La investigación de "CLC bio" creó el software "CLC DNA Workbench¹". Este software crea un entorno informático que permite a los usuarios hacer una gran cantidad de análisis de DNA combinados con la gestión de datos con opciones gráficas, además CLC DNA Workbench está disponible en las plataformas Windows, Mac OS X y Linux.

El software CLC DNA Workbench tiene las siguientes características [1]:

- Acceso a herramientas integradas de investigación
- El trabajo de investigación será fácil de realizar
- Se tendrá una solución bioinformática desarrollada continuamente
- Se podrá modificar el banco de trabajo para requisitos particulares

¹Disponible en <http://www.clcbio.com/>

Por otro lado la investigación de “National Center for Biotechnology Information (NCBI)” creó el software “Genbank²” cuyo propósito es proporcionar acceso dentro de la comunidad científica a la información reciente y comprensiva de la secuencia de DNA. NCBI no propone ninguna restricción en el uso o distribución de “Genbank”.

“Genbank” es una secuencia base de datos genética que contiene aproximadamente 37.893.844.733 bases en 32.549.400 expedientes. Los expedientes de “GenBank” se anotan usando un sistema estándar de términos biológicos. “GenBank” es parte de la colaboración internacional de base de datos de la secuencia de nucleótido, que abarca el banco de datos de la DNA de Japón (DDBJ), del laboratorio de biología molecular europeo (EMBL), y de “GenBank” en NCBI.

² Disponible en <http://www.ncbi.nlm.nih.gov/>

Capítulo 3. Diseño de investigación

3.1 Introducción

En este capítulo se describe la forma en la que será desarrollada la investigación tomando en cuenta los requerimientos, casos de uso y casos de prueba, además de especificar cómo se hará la recolección de datos.

3.2 Documento de requerimientos

a) Requerimientos funcionales

1. Agregar nuevos genes
2. Consultar información acerca de los genes ya agregados.
 - 2.1 Consultar por nombre del gen.
 - 2.2 Consultar por número de gen.
3. Modificar información de los genes ya agregados.
4. Eliminar genes

b) Requerimientos no funcionales

Disponibilidad: El acceso a los datos se realizará mediante una interfaz gráfica ya que éstos de almacenarán en una base de datos.

Portabilidad: El sistema para consultas **G-H** se realizará bajo la plataforma de Windows XP.

Adaptabilidad: El sistema estará disponible a futuros cambios para nuevas implementaciones y/o modificaciones.

Calidad: El sistema permitirá realizar búsqueda de información de manera eficaz y eficiente para mostrar resultados en forma de tabla.

3.3 Modelo Entidad – Relación (E-R)

En este apartado se especifica la composición del modelo Entidad-Relación. En la Figura 3.1 se muestra el diagrama E-R.

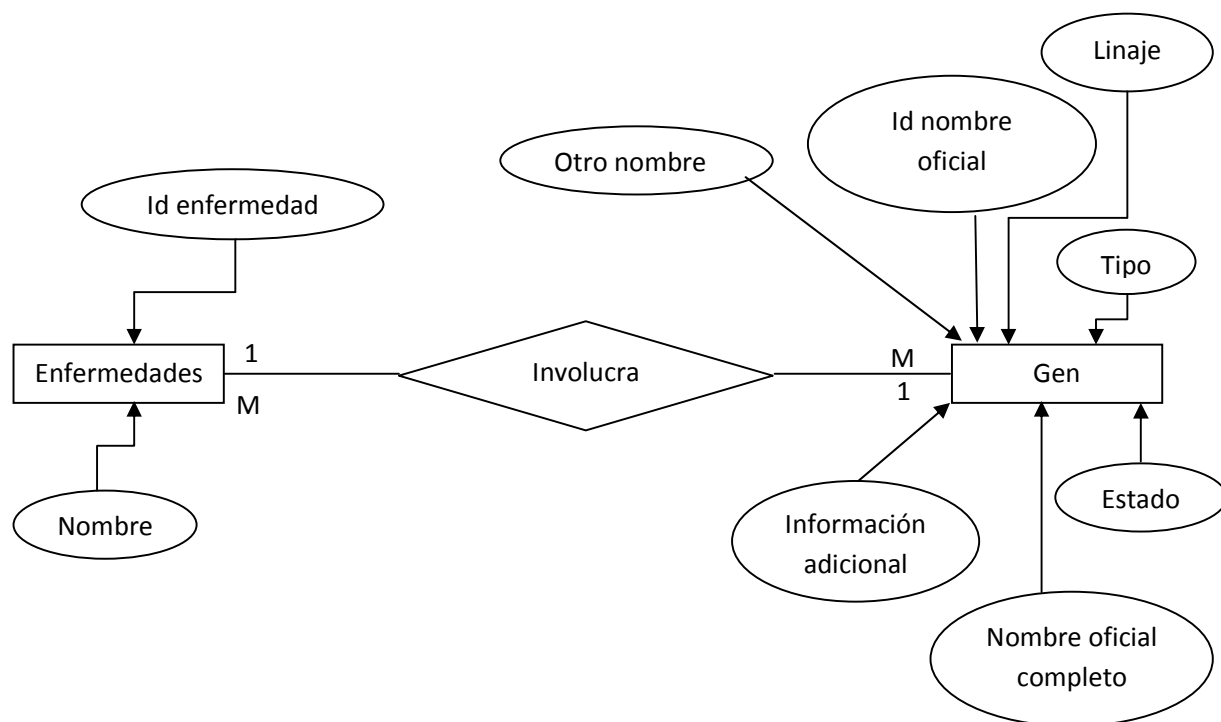


Figura 3.1. Modelo E-R.

3.4 Diccionario de datos

Este apartado define las características lógicas y puntuales de los datos que se van a utilizar en el sistema.

Enfermedades = Id enfermedad + nombre

Id enfermedad = {carácter}

Nombre = {carácter}

Carácter = [A-Z | a-z | 0-9]

Gen = id nombre oficial + nombre oficial completo + linaje + otro nombre + tipo + estado
+ Información adicional

Id nombre oficial = {carácter}

Nombre oficial completo = {carácter}

Linaje = {carácter}

Otro nombre = {carácter}

Tipo = {carácter}

Estado = {carácter}

Información adicional = {carácter}

Carácter = [A-Z | a-z | 0-9]

3.5 Casos de uso

Esta sección especifica las actividades que serán manipuladas por el usuario una vez implementado el sistema. En la Figura 3.2 se muestran las actividades que el usuario podrá realizar.

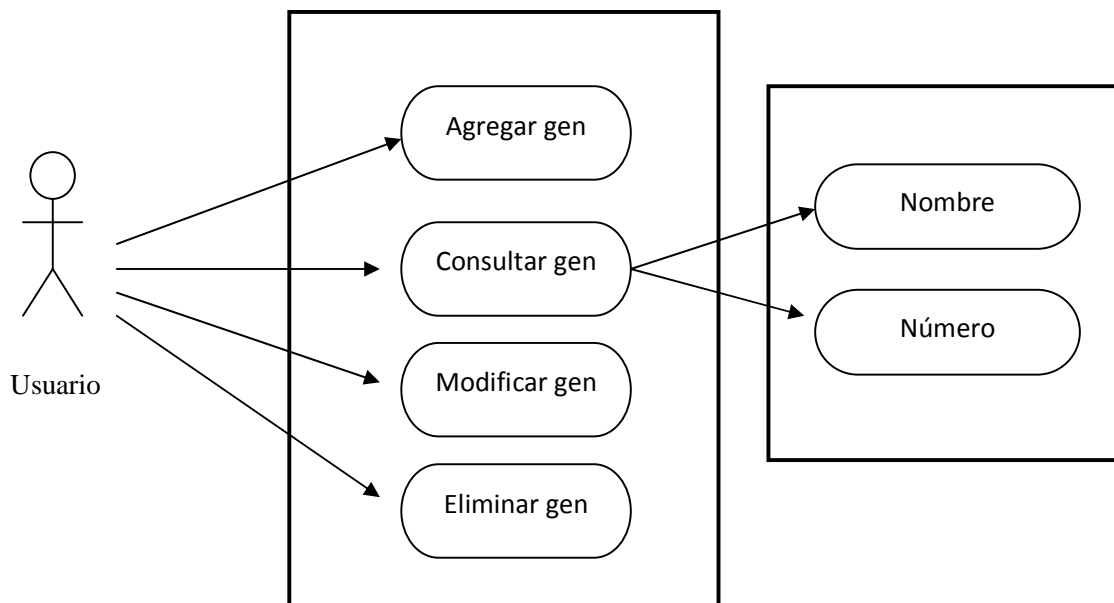


Figura 3.2. Casos de uso principales.

Casos de uso específico

Nombre: CU1 Agregar gen.

Descripción: En este caso de uso se añadirá un nuevo gen a la base de datos.

Datos de entrada: Símbolo oficial, nombre oficial completo, linaje, también conocido como, tipo de gen, estatus, información adicional.

Garantías de éxito: Al salir la pantalla Agregar, todos los datos introducidos se almacenarán.

Datos de salida: mensaje de recopilación satisfactoria o mensaje de error debido a que los datos introducidos no sean validos y/o falte algún campo por llegar.

Nombre: CU2 consultar gen por número.

Descripción: En este caso de uso se realiza una consulta de un gen por medio del número de gen.

Datos de entrada: numero de gen.

Datos de salida: mensaje de búsqueda satisfactoria o mensaje de error debido a que los datos no coincidan.

Nombre: CU3 consultar gen por nombre.

Descripción: En este caso de uso se realiza una consulta de un gen por medio de su nombre.

Datos de entrada: nombre de gen.

Datos de salida: mensaje de búsqueda satisfactoria o mensaje de error donde debido a que los datos no coincidan.

Nombre: CU4 Modificar gen.

Descripción: En este caso de uso se modificar la información de un gen previamente almacenado en la base de datos.

Datos de entrada: Nombre del gen o Número del gen.

Garantías de éxito: En la pantalla Modificar, se muestra la información del gen para hacer las modificaciones requeridas y al salir de esta pantalla, todos los cambios realizados se almacenarán.

Datos de salida: mensaje de Modificación satisfactoria o mensaje de error debido a que los datos introducidos no sean validos y/o falte algún campo por llegar.

Nombre: CU5 Eliminar gen.

Descripción: En este caso de uso se elimina toda la información de un gen previamente almacenado en la base de datos.

Datos de entrada: Nombre del gen o Número del gen.

Garantías de éxito: En la pantalla Eliminar, se muestra la información del gen elegido y al salir de esta pantalla, la información del gen ya no se mostrara.

Datos de salida: mensaje de Eliminación satisfactoria o mensaje de error debido a que los datos de entrada no son validos.

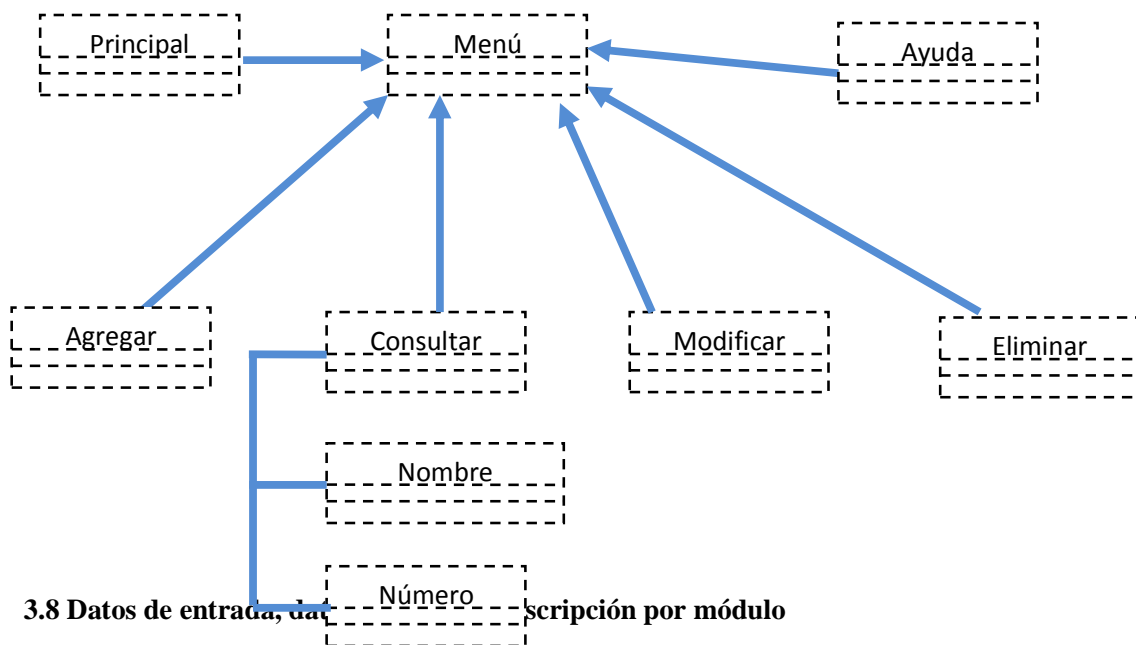
3.6 Especificación de actores

En la Tabla 3.1 se describe a la persona que manipulará el sistema describiendo sus características y la relación que tendrá con el sistema.

Tabla 3.1. Especificación del actor Administrador o Usuario.

Nombre del actor:	Administrador o Usuario
Descripción:	Persona encargada de la manipulación del sistema.
Características:	No cuenta con una autenticación de usuario.
Relaciones:	Interactúa con el sistema para llevar a cabo la manipulación de los genes, además podrá guardar las modificaciones que crea pertinentes.

3.7 Diagrama conceptual



3.8 Datos de entrada, datos de salida y descripción por módulo

En las siguientes tablas se muestran los datos de entrada así como los de salida además de la descripción de cada uno de los módulos pertenecientes al sistema.

Tabla 3.2. Descripción del módulo Ayuda.

Nombre del módulo	Ayuda
Descripción	Contendrá la descripción de cada una de las funciones.
Datos de entrada	Tarea a consultar
Datos de salida	Descripción de las funciones

Tabla 3.3. Descripción del Módulo Agregar.

Nombre del módulo	Agregar
Descripción	Añade un nuevo gen al sistema.
Datos de entrada	Símbolo oficial, nombre oficial completo, linaje, también conocido como, tipo de gen, estatus, información adicional
Datos de salida	Mensaje de recopilación satisfactoria

Tabla 3.4. Descripción del Módulo Consultar.

Nombre del módulo	Consultar
Descripción	Muestra la información del gen solicitado.
Datos de entrada	Nombre del gen, Número del gen
Datos de salida	Pantalla con la información correspondiente al gen solicitado

Tabla 3.5. Descripción del Módulo Modificar.

Nombre del módulo	Modificar
Descripción	Muestra la información del gen solicitado para realizar las

	modificaciones.
Datos de entrada	Nombre del gen, Número del gen
Datos de salida	Mensaje de Modificación satisfactoria

Tabla 3.6. Descripción del Módulo Eliminar.

Nombre del módulo	Eliminar
Descripción	Muestra la información del gen solicitado para eliminarlo
Datos de entrada	Nombre del gen, Número del gen
Datos de salida	Mensaje de eliminación del gen satisfactoria

3.9 Recolección de datos

En esta investigación la información recolectada es acerca de las enfermedades genéticas en los seres humanos, por este motivo es importante saber las características de los genes, fuente de donde es adquirida la información, los tipos de usuarios y el tipo de formato en el que los datos serán manipulados.

3.9.1. Manejo de datos

Los datos a manipular en este proyecto se organizarán de acuerdo a la secuencia o segmento del gen. La información los genes a utilizar serán obtenidos de internet¹.

La información que se obtenga será manipulada sólo por el usuario, quien podrá respaldar la información de su interés. Él usuario final es el único que podrá dar mantenimiento a la base de datos relacional centralizada respecto a la información obtenida después de que la búsqueda han sido procesada

¹ Disponible en <http://www.ncbi.nlm.nih.gov/>

3.9.2. Representación de resultados

El medio por el cual los datos serán recolectados será mediante internet¹. Posteriormente en la base de datos (BD) que se diseñará se almacenarán los datos recolectados para el posterior uso del usuario.

Capítulo 4. Implementación

En este capítulo se hace una descripción del funcionamiento de la interfaz del sistema, mostrando la manipulación de las funciones que se ejecutan en los módulos de altas, bajas, consultas y modificaciones.

4.1 Implementación de las funciones Consultar, Modificar y Eliminar

Para las funciones de Consultar, Modificar y Eliminar se usa la siguiente función, la cual tiene por prioridad checar si existe el gen para después llevar a cabo la instrucción solicitada.

Ejemplo para la Función **Eliminar** Gen

Declarar una variable de tipo carácter

X= ¿Está seguro de querer eliminar este gen del sistema?

Si $X \neq 1$ entonces

Cancelar opción

De lo contrario

Borrar gen seleccionado

Si gen seleccionado = a el número de genes registrados en la base de datos entonces

Moverse al registro anterior

De lo contrario

Moverse al registro siguiente

4.2 Pantalla principal del sistema

En la Figura 4.1 se muestra la pantalla principal al sistema, en esta se puede visualizar un menú [1] que contiene las funciones agregar, consultar modificar y eliminar, también son parte del menú las opciones salir y ayuda. En la parte central se encuentra el título que describe lo que se pretende alcanzar con este sistema, además en la parte inferior izquierda de la pantalla se encuentran los datos (Nombre, correo electrónico y teléfono celular) de quien desarrolló el sistema.



Figura 4.1.

Pantalla

Principal del sistema.

4.2.1 Descripción de la barra de menús

Menú Opciones

En la Figura 4.2 se muestran las opciones *Agregar Gen* (permite agregar información de genes), *Consultar Gen* (permite buscar algún gen en el sistema), *Modificar Gen* (permite hacer cambios en la información de los genes) y *Eliminar Gen* (permite eliminar genes)

Figura 4.2. Pantalla Menú Opciones



Menú Salir

Esta opción permite salir del sistema.

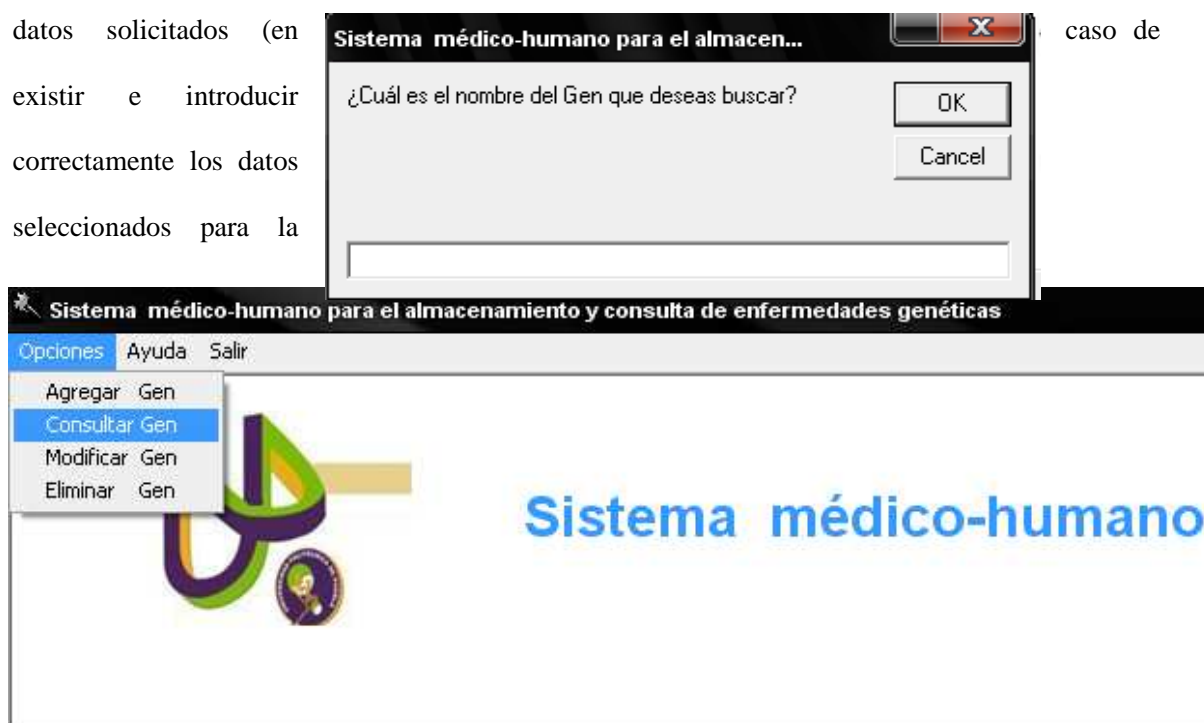
Menú Ayuda

La opción *Acerca del sistema* contenida en el menú Ayuda.

4.2.2 Ejemplos de la barra de menús

Consultar Gen

En la Figura 4.3 se muestra un ejemplo para buscar un gen. Una vez en la opción Consultar Gen del menú opciones, se da clic en la opción búsqueda de Gen por nombre o en búsqueda de gen por Número posterior a esta acción aparece un cuadro de dialogo, el cual pregunta por el nombre del gen que desea buscar (o puede preguntar qué número de gen desea buscar, esto depende del modo de búsqueda que el usuario seleccione) para finalizar la consulta del gen se introducen los datos a buscar y se da clic en el botón aceptar del cuadro de dialogo. Finalmente el sistema muestra los datos solicitados (en caso de existir e introducir correctamente los datos seleccionados para la



búsqueda) o en su defecto aparece un mensaje que muestra el posible problema por el cual no arroja resultados positivos la búsqueda.

Figura 4.3. Pantalla Consultar Gen

En la Figura 4.4 se muestra los datos seleccionados en la pantalla de búsquedas.

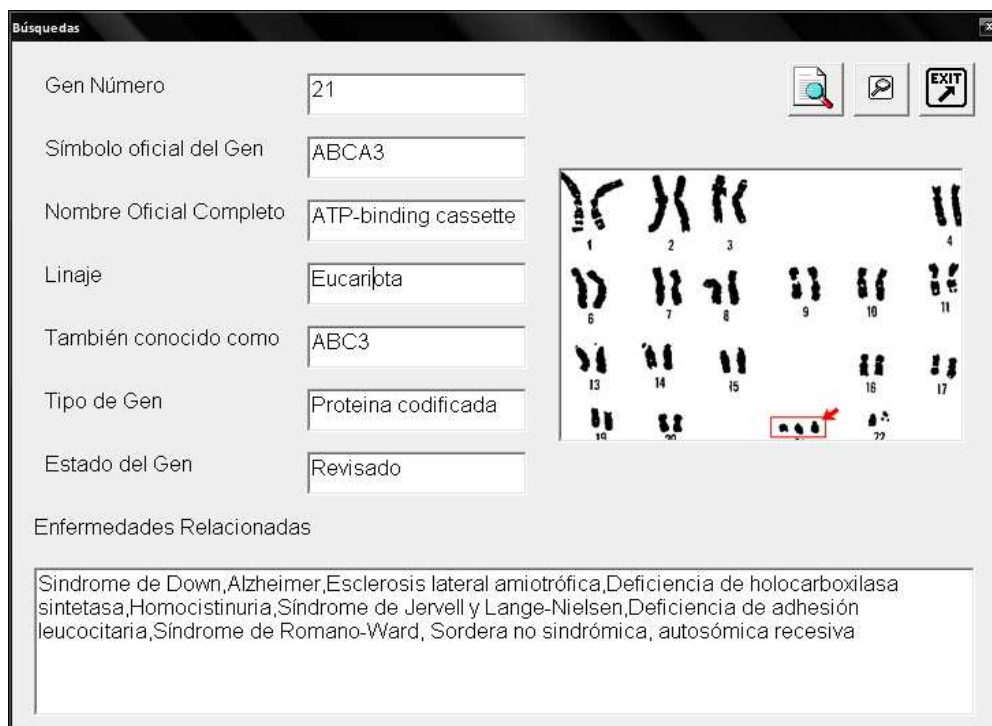


Figura 4.4. Pantalla datos seleccionados

En la figura 4.7 se muestra el cuadro de dialogo al no encontrar los datos solicitados.

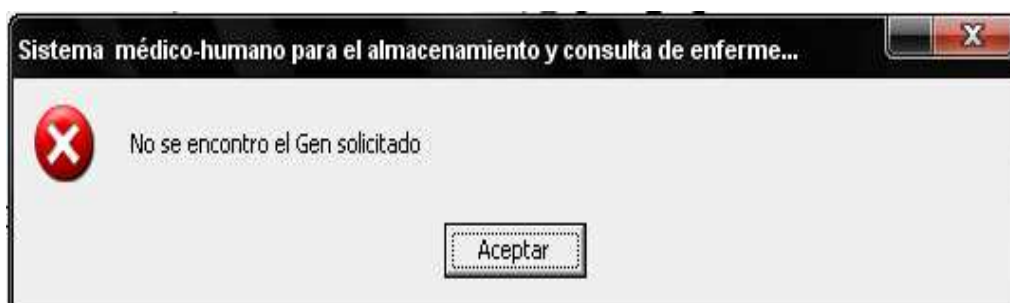


Figura 4.5. Cuadro de dialogo.

Ayuda del sistema

En la Figura 4.8 se muestra la pantalla de la opción Acerca del sistema del menú Ayuda la cual contiene información del sistema.

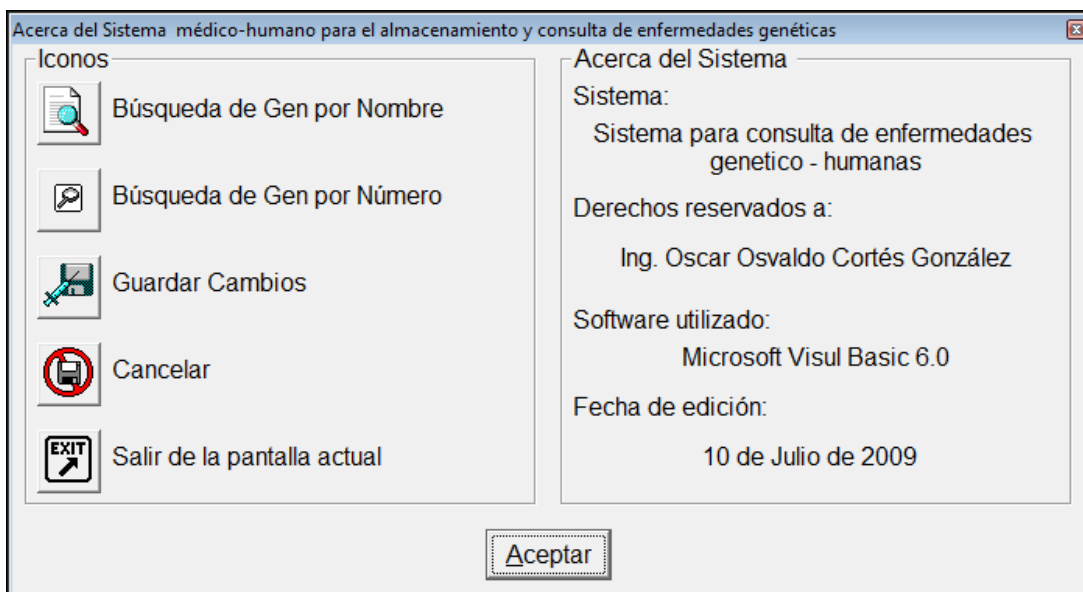


Figura 4.6. Pantalla Acerca del sistema

Capítulo 5. Pruebas

En este capítulo se describen las pruebas que se realizaron para el almacenamiento de información de los cromosomas, en donde las personas especialistas del área utilizaron el sistema y realizaron las tareas de almacenar, consultar, modificar y eliminar, los datos correspondientes a los cromosomas según se presentaban los casos. La Tabla 5.1 muestra la información de los usuarios que realizaron las pruebas, mientras que la Tabla 5.2 muestra las preguntas realizadas para conocer el grado de satisfacción de los usuarios.

Nombre	Perfil	Manejo de BD
David de Jesús Medina Benítez.	Estudiante en Ingeniería en Biotecnología.	http://www.genebase.com/
Juan Carlos Moreno Reyes	Estudiante en Ingeniería en Biotecnología.	http://www.genebase.com/
Miriam Araceli Martínez Gómez	Estudiante en Ingeniería en Biotecnología.	http://www.genebase.com/

Tabla 5.1. Información de los usuarios.

Preguntas	Respuestas		
	Usuario 1	Usuario 2	Usuario 3
Manejo del sistema	Aceptable	Bueno	Bueno
Diseño del sistema	Bueno	Bueno	Bueno
Facilidad para buscar de información	Excelente	Excelente	Excelente
¿La información que maneja el sistema es adecuada?	Si	SI	Si
¿Los datos que se piden ingresar son claros?	Si	Si	Si
¿El sistema necesita algún otro dato?	No	No	No

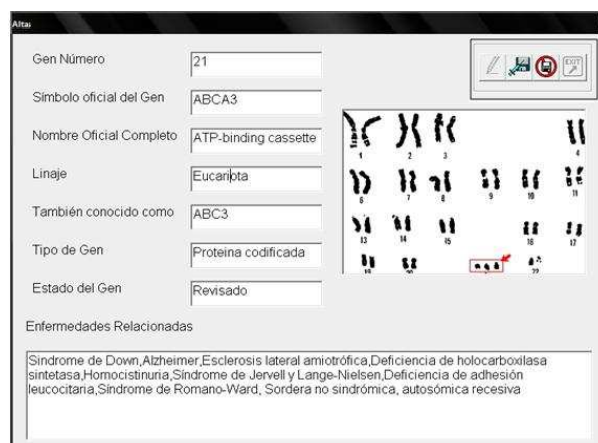
Tabla 5.2 Cuestionario

Prueba 1 Ingresar datos al sistema.

La primera prueba consistió en que el usuario (Especialista en genética) debía ingresar los datos de un cromosoma.

El usuario abrió la pantalla de Altas e introdujo los datos del cromosoma 21 llenando los campos de: Gen Número, Símbolo oficial de Gen, Linaje, También conocido como, Tipo de Gen, Estado del Gen, Enfermedades relacionadas, Ingresar imagen. Posteriormente el usuario guardó los datos y podía salir de la aplicación, en caso de que no quisiera guardar los datos este cancelaba la operación, y también podía salir de la aplicación

La Figura 5.1 muestra la pantalla de Altas en la cual el genetista introdujo información del gen número 21.



The screenshot shows a software interface titled 'Altas'. On the left, there is a form with the following fields and values: 'Gen Número' (21), 'Símbolo oficial del Gen' (ABCA3), 'Nombre Oficial Completo' (ATP-binding cassette), 'Linaje' (Eucariota), 'También conocido como' (ABC3), 'Tipo de Gen' (Proteína codificada), and 'Estado del Gen' (Revisado). To the right of the form is a karyotype image with chromosome 21 highlighted in red. Below the form is a text area labeled 'Enfermedades Relacionadas' containing a list of associated conditions: 'Síndrome de Down, Alzheimer, Esclerosis lateral amiotrófica, Deficiencia de holocarboxilasa sintetasa, Homocistinuria, Síndrome de Jervell y Lange-Nielsen, Deficiencia de adhesión leucocitaria, Síndrome de Romano-Ward, Sordera no síndrómica, autosómica recesiva'. The interface also includes a toolbar with icons for editing, deleting, and saving.

Figura 5.1. Pantalla de Altas

El resultado de esta primera prueba presentó las siguientes observaciones:

- Incluir ayuda rápida que indique el tipo de información requerida para cada uno de los cuadros de texto en la pantalla Altas.
- Modificar el tamaño de letra en el cuadro de texto Enfermedades Relacionadas.
- Ampliar el área donde se muestra la imagen

Prueba 2 Consultar datos.

La segunda prueba consistió en que el usuario debía consultar los datos de un cromosoma almacenado en el sistema.

El usuario abrió la pantalla de Consultas e introdujo el dato solicitado para poder realizar la búsqueda del cromosoma 21. La pantalla que mostró el sistema en esta búsqueda fue la información en los campos: Gen Número, Símbolo oficial de Gen, Linaje, También conocido como, Tipo de Gen, Estado del Gen, Enfermedades relacionadas, Ingresar imagen. Posteriormente el usuario realizó esta operación repetidas veces, finalmente también introdujo información errónea, ante acción el sistema mostró la ventana Dato incorrecto. Finalmente el usuario probó la opción Consultar Gen por Número obteniendo con esto el mismo resultado satisfactorio que con el anterior método búsqueda. La Figura 5.2 muestra la ventana donde el usuario introdujo el nombre del gen

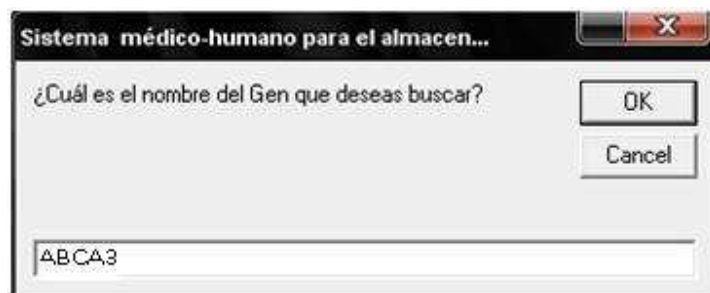


Figura 5.2. Ventana para consultar genes

En la Figura 5.4 se observan los datos que el sistema mostró con la búsqueda solicitada por el usuario. El resultado para esta prueba es el siguiente:

- Unificar el criterio para buscar Gen por nombre. Especificar si se puede escribir con mayúsculas o minúsculas en la ventana para consultar genes.

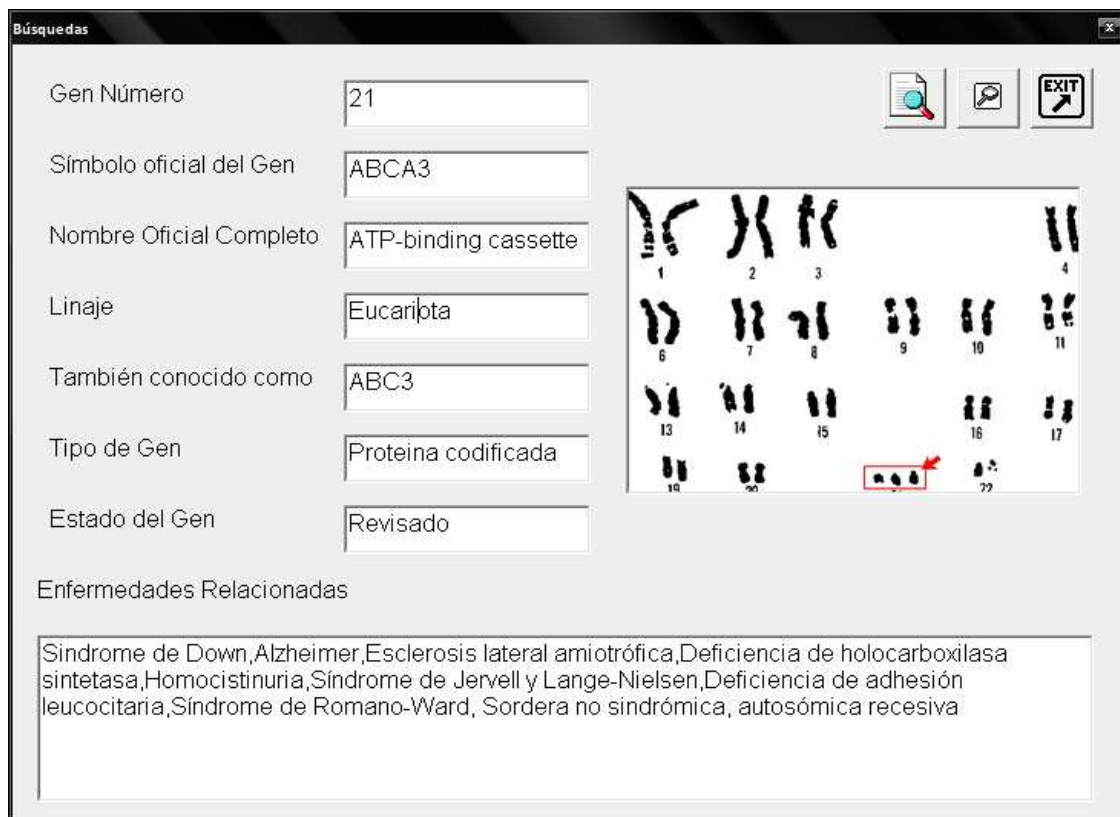


Figura 5.3. Pantalla Búsquedas

Prueba 3 Modificar datos.

La tercera prueba fué que el usuario debía modificar los datos de un cromosoma almacenado en el sistema.

El usuario abrió la pantalla Modificar e introdujo el dato elegido para que el sistema pudiera realizar la búsqueda del cromosoma 21, antes de que el usuario pueda hacer cambios, el sistema pregunta si se desea modificar la información. Después de que el usuario aceptó modificar los datos el sistema muestra la pantalla con la información en los campos: Gen Número, Símbolo oficial de Gen, Linaje, También conocido como, Tipo de Gen, Estado del Gen, Enfermedades relacionadas, Ingresar imagen, todos los anteriores campos se activaron para poder realizar modificaciones. Una vez que el usuario realizó los cambios presionó el botón guardar. Finalmente para verificar los cambios realizados en el gen 21, el usuario realizó una búsqueda, siendo satisfactorio el resultado de esta, pues el sistema

mostro los cambios realizados al gen. La Figura 5.4 muestra la ventana donde el usuario introduce el nombre del gen en el cual desea realizar modificaciones.

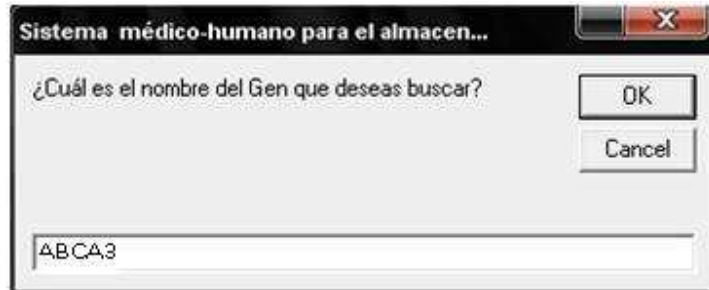


Figura 5.4. Pantalla para consultar un gen

En la Figura 5.5 se observan la venta para confirmar la acción modificar.

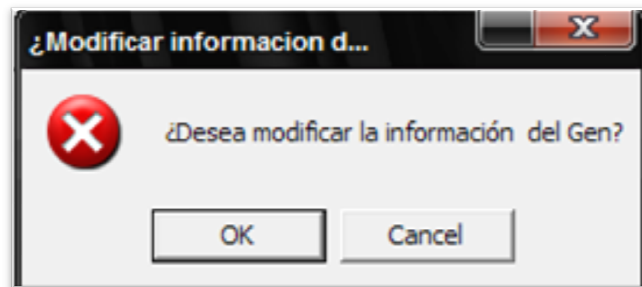


Figura 5.5. Ventana de ratificación modificar

Finalmente la Figura 5.6 muestra la pantalla Modificar con la información del gen solicitada y todos los cuadros de texto activados.

Gen Número: 21

Símbolo oficial del Gen: ABCA3

Nombre Oficial Completo: ATP-binding cassette

Linaje: Eucariota

También conocido como: ABC3

Tipo de Gen: Proteína codificada

Estado del Gen: Revisado

Enfermedades Relacionadas

Síndrome de Down, Alzheimer, Esclerosis lateral amiotrófica, Deficiencia de holocarboxilasa sintetasa, Homocistinuria, Síndrome de Jervell y Lange-Nielsen, Deficiencia de adhesión leucocitaria, Síndrome de Romano-Ward, Sordera no sindrómica, autosómica recesiva

Figura 5.6. Pantalla Modificar

Observación

- Agregar una ventana para confirmar las modificaciones realizadas.

Prueba 4 Eliminar datos.

En la última prueba realizada al sistema el usuario eliminó los genes que había añadido al sistema. Para esto abrió la pantalla eliminar e ingresó el dato del gen para eliminar del sistema. Después de esto el sistema mostró al usuario una pantalla donde pregunta si desea eliminar el gen, a lo cual este aceptó. La Figura 5.7 muestra la ventana donde el usuario introdujo el nombre del gen que desea eliminar.

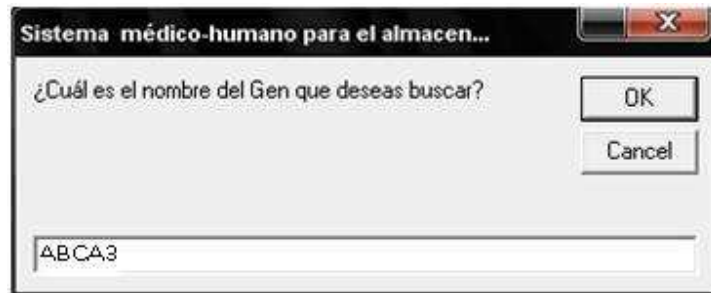


Figura 5.7. Pantalla para buscar gen a eliminar

La Figura 5.8 muestra la ventana para confirmar la eliminación del gen por parte del usuario.

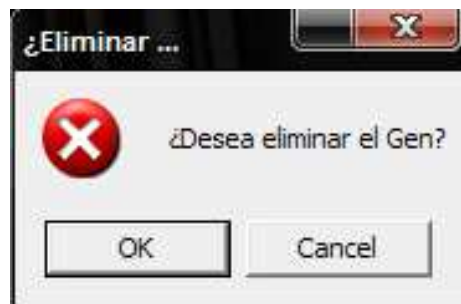


Figura 5.8. Eliminar Gen

Observación

- Agregar una ventana para confirmar la eliminación del Gen

Todas las observaciones realizadas en este capítulo por el especialista en genética fueron corregidas y probadas en una segunda prueba hecha por el mismo usuario.

Capítulo 6. Conclusiones

Se desarrolló un sistema que trabaja con información cromosómica para poder almacenar y consultar enfermedades genéticas. Se implementaron las opciones Agregar, Consultar, Eliminar y Modificar, para manipular la información almacenada. Esta herramienta permitirá que el genetista pueda consultar enfermedades cromosómicas.

Se trabajo con un genetista, quien participó en la validación de requerimientos. La información de los cromosomas que se utilizó en la parte de pruebas, fue en base a las enfermedades relacionadas con el gen número 22. Es importante aclarar que para las pruebas con dicho gen se aplicaron todas las opciones que ofrece el sistema, y las pruebas fueron satisfactorias.

En este proyecto de investigación se cubrieron los objetivos. En el sistema, el almacenamiento y las consultas se realizan manualmente, posteriormente este proyecto será la base de un sistema automático para la detección de enfermedades hereditarias, en donde almacenará no sólo las características de enfermedades, sino también los patrones que las definen.

Referencias.

- [1] Oliva R., Ballesta F., Oriola J., Clària J., 2004, *Genética Médica Publicacions i Edicions UB*
- [2] Jorde, L., Carey, J., Bamshad, M., White R., 2008, *Genética Médica*, ISBN13: 978848174763-8, ELSEVIER
- [3] Soberón M., 2008, *La ingeniería genética, la nueva biotecnología y la era genómica*, ISBN: 9789681664336, Fondo de Cultura Económica.
- [4] Rama S., Costas B., Diane B., Beatty J., *Thinking about Evolution: Historical, Philosophical, and Political Perspectives*, 2000, ISBN 0521620708, 9780521620703, Cambridge University Press.
- [5] Garbán, H., 1999, *Medicina molecular: nueva perspectiva en medicina*, ISBN 1317-987X, Academia Biomédica Digital N°. 2.
- [6] Malcolm S., Goodship J., *Genotype to Phenotype*, 2001, ISBN 0203450426, 9780203450420, Taylor & Francis Group.
- [7] Lewontin Richard C., *The Genetic Basis of Evolutionary Change*, 1974, ISBN 0231033923, 9780231033923, Columbia University Press.
- [8] Mayr E., 1985, *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*, ISBN 0674364465, 9780674364462, Harvard University Press.
- [9] Shmal'gauzen I., Dobzhansky T., *Factors of Evolution: The Theory of Stabilizing Selection*, 1949, Blakiston Co.
- [10] Thompson M., 2004, Nussban R., McInnes R., Thompson J., Willard H., Peral J., Sedó M., *Thompson & Thompson: Genética en medicina*, ISBN 8445812254, 9788445812259, Masson.
- [11] Lyons J., 2006, *Características Reconocibles de la Malformación Humana*, ISBN 8481749478, 9788481749472, Elsevier España.
- [12] Piédrola G., Domínguez R., 2001, *Medicina Preventiva y Salud Pública*, ISBN 8445810243, 9788445810248, Elsevier España.
- [13] Pan American Health Organization, *Organizacion Mundial de la Salud OMS, Organización Panamericana de la Salud*, 1995, ISBN 927531554X, 9789275315545, Pan American Health Org.

- [14] Étienne J., 2001, Bioquímica Genética. Biología Molecular: Biología molecular, ISBN 844580992X, 9788445809921, MASSON
- [15] Clayton J., Dennis C., 2003, 50 years of DNA, ISBN 978-1-4039-1479-8, Palgrave Macmillan Press.
- [16] Subirana, J., 1985, Estructura del AND, ISBN 8420511609, 9788420511603, Alhambra.
- [17] Lewin B., Aguilera A., Casa de Jesús J., Casa de Jesús J., 1996, Genes, ISBN 84-291-1845-4 Reverté S.A
- [18] William J., Kolb D., 2000, Química para el nuevo milenio, ISBN 9701703413, PRENTICE HALL.
- [19] Audesirk T., Audesirk G., Byers B., Escalona H., Escalona R., 2003, Biología La vida en la Tierra, ISBN 9789702603702, PRENTICE HALL.
- [20] Hienz H., 1975, Cromosomas: Introducción a la Citogenética clínica para médicos y estudiantes, ISBN 313464701X, 9783134647013, Alhambra.
- [21] Ramón J., 1996, Citogenética, ISBN 84-8936558X, Complutense, S.A.
- [22] Hartl D., Jones E., 2004, Genetics: Analysis of genes and genomes, ISBN 0-7637-1511-5, Jones and Bartlett Publishers.
- [23] Bachmann K., 1978, Biología para médicos, ISBN 84-291-1804-7, Reverté S.A.
- [24] Vives J., Aguilar J., 2006, Manual de técnicas de laboratorio en Hematología, ISBN 13 978-84-458-1581-6, 84-458-1581-4, Masson
- [25] Jiménez L., Merchant H., 2002, Biología celular y molecular, ISBN 970-26-0387-0, Pearson Educación.
- [25] Date C.J., Introducción a los sistemas de base de datos, 2001, ISBN 9684444192, 9789684444195, Pearson Educación.
- [26] Date C.J., Bases de datos: Una guía práctica, 1987, ISBN 9688580732, 9789688580738, Addison-Wesley Iberoamericana.

[27] Rob P., Coronel C., Sistemas de bases de datos: Diseño, implementación y administración, 2004, ISBN 9706862862, 9789706862860, Cengage Learning Editores.

Índice de Figuras y Tablas.

Índice	
Tabla 2.1. Hechos importantes en el campo de la genética	13
Figura 2.1. Relación entre los espacios fenotípicos y genotípicos	17
Figura 3.1. Modelo E-R	42
Figura 3.2. Casos de uso principales	44
Tabla 3.1. Especificación del actor Administrador o Usuario	47
Tabla 3.2. Descripción del módulo Ayuda	48
Tabla 3.3. Descripción del Módulo Agregar	48
Tabla 3.4. Descripción del Módulo Consultar	48
Tabla 3.5. Descripción del Módulo Modificar	49
Tabla 3.6. Descripción del Módulo Eliminar	49
Figura 4.1. Pantalla Principal del sistema	52
Figura 4.2. Pantalla Menú Opciones	53
Figura 4.3. Pantalla Consultar Gen	54
Figura 4.4. Pantalla datos seleccionados	55
Figura 4.5. Cuadro de dialogo	55
Figura 4.6. Pantalla Acerca del sistema	56
Tabla 5.1 Información de los usuarios	57
Tabla 5.2 Cuestionario	57
Figura 5.1. Pantalla de Altas	58
Figura 5.2. Ventana para consultar genes	59
Figura 5.3. Pantalla Búsquedas	60
Figura 5.4. Pantalla para consultar un gen	61
Figura 5.5. Ventana de ratificación modificar	61
Figura 5.6. Pantalla Modificar	62
Figura 5.7. Pantalla para buscar gen a eliminar	63
Figura 5.8. Eliminar Gen	63