



UNIVERSIDAD  
POLITECNICA  
DE VALENCIA



Centro de Investigación en Métodos  
de Producción de Software

## Tesis de Máster

Máster de Posgrado Oficial en Ingeniería del Software, Métodos  
Formales y Sistemas de Información

# Integración de Haplotipos al Modelo Conceptual del Genoma Humano utilizando la Metodología SILE

Julio de 2013, Valencia

*José Fabián Reyes Román*

Jore1@posgrado.upv.es

*Director: Óscar Pastor López*

opastor@pros.upv.es

*Co-director: Mercedes R. Fernández Alcalá*

mfernandez@pros.upv.es

# Índice General

## Contenido

<b>Agradecimientos.....</b>	<b>4</b>
<b>Capítulo 1 Introducción .....</b>	<b>5</b>
1.1 Motivación .....	6
1.2 Problemática.....	6
1.3 Objetivos.....	7
1.4 Contexto.....	7
1.5 Estructura de la tesis.....	8
<b>Capítulo 2 Estado del Arte.....</b>	<b>10</b>
2.1 Empresas del Entorno Genómico.....	10
2.2 Sensibilidad al Alcohol.....	11
2.3 Haplotipos.....	12
2.4 Manipulación de Datos (BD Genómicas).....	13
2.5 Bases de Datos basadas en Modelado Conceptual .....	13
<b>Capítulo 3 Solución Propuesta.....</b>	<b>14</b>
3.1 Modelado de Procesos de Negocio de una Empresa Genómica.....	14
3.1.1 Descripción de la Empresa & Organigrama .....	14
3.1.2 Diagramas de los Procesos de Negocio (BPMN) .....	17
3.2 Especificación de Requisitos: Visión General.....	21
3.2.1 Lista Definitiva de Requisitos del Nuevo Sistema (LDR).....	21
3.2.2 Casos de Uso & Diagrama de Clases .....	23
<b>Capítulo 4 Implementación de SILE para el Diagnóstico de la Sensibilidad al Alcohol ..</b>	<b>33</b>
4.1 Metodología SILE aplicada a la Sensibilidad al Alcohol.....	33
4.1.1 Search.....	34
4.1.1.1 Haplotipos.....	34

4.1.2	Identification.....	35
4.1.2.1	Haplotipo detectado en la Sensibilidad al Alcohol.....	38
4.1.2.2	Estudio HapMap (Frequencies) .....	38
4.1.3	Load.....	41
4.1.3.1	Modelo Conceptual del Genoma Humano (v3) e Integración de los Haplotipos.....	41
4.1.3.2	Carga de Base de Datos.....	44
4.1.3.2.1	Carga Selectiva .....	46
4.1.4	Exploitation.....	47
4.1.4.1	Diseño e Implementación del Servicio Web.....	47
4.1.4.1.1	Diagrama Entidad-Relación Normalizado .....	48
4.1.4.1.2	Descripción de Tablas Normalizadas.....	50
4.1.4.1.3	Diseño de Interfaces de Usuario .....	52
4.1.4.2	Contenidos GenesLove.me .....	58
4.1.4.3	Seguridad.....	59
	<b>Capítulo 5 Conclusiones &amp; Trabajo Futuro .....</b>	<b>60</b>
5.1	Trabajo Futuro.....	61
	<b>Referencias Bibliográficas .....</b>	<b>62</b>
	<b>Apéndice A.....</b>	<b>66</b>
A.1	Índice de Figuras.....	66
	<b>Apéndice B.....</b>	<b>68</b>
B.1	Glosario.....	68
	<b>Apéndice C.....</b>	<b>70</b>
C.1	Interfaz GemBiosoft.....	70
C.2	Consentimiento Informado.....	72
C.3	Infraestructura de la Empresa.....	74
C.4	Estructura Tesis de Máster .....	75

# Agradecimientos

## ***A Dios:***

Por darme la fortaleza para luchar y lograr mis metas, por darme salud y ayudarme en los momentos difíciles, por demostrarme que contigo todo es posible.

## ***A mis padres: Nilda Román Cedeño & José Fabián Reyes Olea***

Les dedico este y todos mis logros, gracias a ustedes por darme la vida y enseñarme a ser una persona de bien, por el ejemplo que me dan día a día, por demostrarme que todo es posible cuando uno se lo propone; por estar conmigo en todo momento, por su apoyo y ayuda incondicional, por sus consejos e ideas que me han ayudado a crecer y a ser quien soy. *¡Los quiero muchísimo!*

## ***A mi hermana: Nilda Josefina Reyes Román “Vanessa”***

Por estar junto a mí en todo momento y por todas nuestras aventuras. Que Dios te bendiga mucho hermanita ¡bella!

## ***Flia. Reyes Olea & Román Cedeño:***

Son de lo mejor, gracias por estar conmigo y ser como son, se les quiere. ¡¡¡Bendiciones!!!

## ***Centro de Investigación en Métodos de Producción de Software (ProS):***

A todos les agradezco por brindarme la oportunidad de aprender y crecer junto a ustedes, en especial a todo el Grupo Genoma (en especial: Ainoha, Ana Ma., Roberto, Verónica, MariaJo, Jorge, Ángel), que siempre estuvieron disponibles para colaborar y mostrarme el mejor camino en mi trayecto a la meta.

## ***Óscar Pastor López & Mercedes R. Fernández Alcalá:***

Muchísimas gracias por permitirme trabajar junto a ustedes y por todos los momentos que pasamos juntos, por introducirme en este mundo, por ayudarme a crecer y sacar lo mejor de mí en el día a día.

## ***GemBiosoft***

Por la oportunidad de participar con ustedes, por hacer mis días agradables y por hacerme sentir parte de la familia siempre. Un agradecimiento especial para la Dra. Ana Levin, por tener siempre un consejo/sugerencia para mejorar y crecer, y por sus grandes aportaciones biológicas.

Por último, agradecer a mis compañeros en esta aventura: Fausto Nelson & Alexander Santana, por su apoyo y consejos siempre. A todos mis compañeros del máster y profesores. A mi gran familia en Valencia de la Iglesia Rey de Reyes “Se les quiere mucho”; a la OAI (Ofic. Acción Internacional) por su colaboración y ayuda; a todos los compañeros de RD-VLC que estuvieron conmigo siempre (Rosa Matos, Keyla Rijo, Aury Charles) y todos en general.

# Capítulo 1

## Introducción

En la actualidad la aplicación y oferta de test genéticos directos al consumidor le han facilitado un sin número de ventajas al ámbito bioinformático. Por igual ha dado un gran giro en el campo genómico, puesto que en nuestros días los usuarios tienen la facilidad de acceder a servicios que le permitan conocer u obtener un diagnóstico precoz de alguna (s) enfermedad (es) de origen genético.

La heterogeneidad de los Sistemas de Información Genómicos (**GeIS**) trae de la mano un gran reto el cual todos conocemos hoy en día como el “caos genómico”.

La gestión de bases de datos para la manipulación de estos datos, brinda muchos beneficios a la comunidad científica y tecnológica. El campo bioinformático exige tener un estricto control y dominio en la manipulación de los datos, porque están conformados por grandes cantidades de información que provienen de distintas fuentes de datos. Es por ello que se deben realizar estudios profundos que nos ayuden a implementar mecanismos para mejorar en aspectos de integridad, manipulación de datos, entre otros, permitiendo garantizar resultados más fiables para los usuarios.

Como ejemplo de esta heterogeneidad podemos encontrar algunas fuentes de datos actuales, como: NCBI [1] [2], OMIM [3] y Ensembl [4] nos brindan un conjunto muy extenso de información, por lo que debemos extraer datos concisos que ayuden a generar resultados con mayor precisión. Cada vez más siguen surgiendo herramientas y metodologías que permiten facilitar la labor, y sobretodo sacar el máximo provecho de los avances médicos para brindar a la comunidad alternativas que sirvan para tratar y/o mejorar su salud y calidad de vida.

Todo esto ha sido posible por el constante crecimiento en el área de la bioinformática, lo que ha permitido dar pasos agigantados en estos tiempos. Una metodología que fue creada para mejorar el tratamiento de los datos es SILE (Search-Identification-Load-Exploitation), esta fue desarrollada por el grupo de investigación ProS, y su finalidad es mejorar los procesos de carga de las bases de datos del genoma humano.

Son muchas las ventajas de trabajar en este campo, el cual se mantiene en un avance continuo. Donde cada vez más nos acercamos a una medicina personalizada al 100%.

## 1.1 Motivación

Debido al gran crecimiento que se observa en el estudio de los Sistemas de Información Genómicos, y ver como investigadores de la comunidad médica/científica buscan de manera constante sistemas que les permitan obtener diagnósticos eficientes y fiables. Son algunas de las causas que demuestran la gran necesidad de realizar análisis apoyados en modelos de negocio y requerimientos, puesto que miramos más allá de lo existente y confirmamos que es lo que realmente se necesita o escasea.

El modelado conceptual nos ayuda a crear sistemas potentes, y nos garantizan una mejora continua de los Sistemas de Información Genómicos (GeIS). Adentrarnos en este campo nos abre puertas, nos ayuda a crecer y a desarrollar nuevas herramientas que permitan generar resultados de manera rápida y segura.

La bioinformática ha evolucionado bastante en los últimos años, lo que le hace un campo muy atractivo para aplicar los avances tecnológicos y conocimientos adquiridos con el fin de mejorar y optimizar los procesos actuales.

## 1.2 Problemática

La gran demanda actual de test genéticos directos al consumidor ha despertado alrededor del mundo el interés por conocer y estudiar los procesos que estos requieren. Es por ello que surge la necesidad de explotar los conocimientos en BPMN y análisis para estudiar y a su vez evaluar los procesos de negocio que participan en esta actividad, con fines de aportar ideas innovadoras que permitan mejorar la gestión de los servicios de cara al público en general.

La implementación de los estudios realizados nos llevó a aplicar la metodología SILE para que nos facilitase cada una de las tareas, y con la recopilación de información surgió el problema del tratamiento de los haplotipos, un punto importante para la generación de los diagnósticos y que no era considerado en el modelo conceptual del genoma humano.

Con la implementación de la metodología SILE, buscamos erradicar la problemática en la gestión de los datos; con esta idea creamos bases de datos filtradas, donde solamente manejamos datos depurados y validados, ayudándonos a generar diagnósticos más precisos y a optimizar nuestra base de datos del genoma humano. Dejando de lado la carga masiva, en donde tomamos todos los datos encontrados, pasamos a una carga selectiva que mejora circunstancialmente la integridad y calidad de los datos. Esta metodología la ponemos en práctica con la Sensibilidad al Alcohol, teniendo en cuenta que la misma puede ser aplicada a cualquier enfermedad.

Por estas causas y muchas más, hemos querido abordar estos problemas con el fin de traer soluciones que permitan seguir avanzando en esta área tan interesante.

### 1.3 Objetivos

El objetivo general es presentar el modelo de negocio para los servicios web orientados a brindar test genéticos directos al consumidor y realizar la implementación para la Sensibilidad al Alcohol, integrando al modelo conceptual del genoma humano el estudio/caso de los Haplotipos.

#### Objetivos Específicos

- ✓ Especificar el modelo de negocio para el diseño e implementación de servicios web que permitan gestionar tests genéticos directos al consumidor.
  - Realizar un análisis exhaustivo de las necesidades de los usuarios que buscan alternativas sencillas para conocer su predisposición a padecer ciertas enfermedades de origen genético.
- ✓ Aplicar SILE para mejorar la calidad de los datos y optimización de tiempos.
  - Mediante la implementación de SILE (Search-Identification-Load-Exploitation) lograr generar resultados más aproximados a la realidad, trabajar con datos tratados que nos ayuden a mejorar la calidad de los mismos en la generación de los diagnósticos.
- ✓ Estudiar el caso de los Haplotipos en el diagnóstico de enfermedades de origen genético e integrar en el modelo conceptual del genoma humano.
  - Identificar la repercusión de los haplotipos, y ver la importancia a la hora de generar los diagnósticos. La integración de los haplotipos nos ayuda a mejorar los resultados de forma directa, porque son un conjunto de parámetros que se deben revisar al momento de diagnosticar una enfermedad. Tomando la Sensibilidad al Alcohol como caso de prueba.
- ✓ Explotar la experiencia en Ingeniería de Requisitos y BPMN en el área de bioinformática.
  - Realizar estudios y análisis que aborden esta área. Aprovechar los conocimientos en estos campos para crear modelos que se puedan adaptar a cualquier problemática con el paso de los años. Fortalecer los modelos bioinformáticos de manera que estos cumplan con las necesidades de los usuarios reales y potenciales.

### 1.4 Contexto

Esta tesis de máster ha sido elaborada en el Centro de Investigación en Métodos de Producción de Software (ProS) de la Universidad Politécnica de Valencia (UPV). Durante los últimos años, el Centro ProS ha enfocado varias de sus líneas de investigación al campo de la bioinformática [5].

En la actualidad se están llevando un sin número de investigaciones y/o estudios de cómo el Modelado Conceptual puede ayudar en el ámbito de la Genómica a crear Sistemas de Información que permitan mejorar y obtener grandes ventajas/beneficios para el proceso de los diagnósticos genómicos de los pacientes.

La presente tesis de máster, aborda un análisis completo por medio de la aplicación de los procesos de negocio e implementación de servicios web para el diagnóstico genómico. La empresa en que se plantearon estos análisis fue GemBiosoft [6], la cual tiene como misión fomentar el uso de los sistemas de información (SI) en el área de la bioinformática.

En las etapas iniciales desarrollamos los BPMN de la empresa y una vez hecho, buscamos la forma de implementarlos y para ello tomamos como caso de referencia la Sensibilidad al Alcohol, en la cual aplicamos la metodología SILE y esto nos llevó a la problemática de los haplotipos, la cual solucionamos con la aportación/integración al modelo conceptual del genoma humano y base de datos para cubrir la misma.

Con la ayuda de la empresa se desarrollaron las distintas fases de estudios para determinar los factores claves e incidencias altas para la mejora de los servicios de test genéticos directos al consumidor. Gracias a los distintos medios como: entrevistas, brainstorming, reuniones, revisiones semanales, consultas con especialistas en biología, entre otros nos ayudó a estar claros sobre las metas a alcanzar al final del camino.

## 1.5 Estructura de la tesis:

La estructura utilizada en el transcurso de esta tesis de máster es como sigue:

### **Capítulo 1. Introducción**

Este capítulo explicará cual ha sido la motivación de realizar este trabajo, dejando claro cuál es el problema que tenemos actualmente con los Sistemas de Información Genómicos, conociendo los retos que traen consigo este campo de estudio. Un área que cada vez es más atractiva y llamativa por los avances alcanzados por la bioinformática en nuestros días, haciendo que el campo investigativo y médico busque las herramientas que les permitan optimizar la generación de los diagnósticos genéticos.

### **Capítulo 2. Estado del Arte**

En esta sección presentamos un estudio de la situación actual de la bioinformática junto a los populares/famosos test genéticos; además de un análisis realizado a distintas empresas a nivel nacional como internacional. También comentamos sobre las bases genéticas para el estudio de la sensibilidad al alcohol y como la misma está ligada a la solución del tratamiento de los haplotipos.

### **Capítulo 3. Solución Propuesta**

En este capítulo presentamos el proceso de negocio de los diagnósticos genéticos, mostramos las etapas que se deben llevar a cabo para llegar al resultado definitivo. También ofrecemos una visión general de los requisitos del servicio web para la puesta en marcha de los test genéticos directos al consumidor.



**Capítulo 4. Implementación de SILE para el Diagnóstico de la Sensibilidad al Alcohol**

En esta sección desarrollamos la metodología SILE aplicándola a la sensibilidad al alcohol, además presentamos como surge la necesidad de integración de los haplotipos, su importancia y utilidad al implementarlo dentro del modelo conceptual del genoma humano para la mejora de los diagnósticos genéticos.

**Capítulo 5. Conclusiones & Trabajo Futuro**

Finalmente, en esta sección presentamos las conclusiones que nos quedan al realizar este trabajo de investigación, y mostramos las posibles líneas de estudio para el trabajo futuro en este ámbito.

## Capítulo 2

# Estado del Arte

En la actualidad la bioinformática desempeña un rol importante, debido a los aportes y avances que ha facilitado al sector médico y tecnológico. En este caso para los test genéticos ha servido de proyección para brindar diagnósticos que ayuden a prevenir y/o tratar enfermedades de origen genético.

El campo de la bioinformática está en un constante crecimiento, así como por igual el estudio de las enfermedades que están asociadas a la genética. A partir del 1977 se inicia la secuenciación de ADN y el desarrollo de software para analizarlo, con lo que rápidamente para el año siguiente se publica la primera secuencia de genes completa de un organismo [7].

Las ventajas de los test genéticos son innumerables, puesto que nos permiten identificar mutaciones o alteraciones en los genes, siendo de gran utilidad e interés para la medicina clínica y favoreciendo el Diagnóstico precoz de enfermedades [8] [9]. El entorno bioinformático nos ayuda de igual manera en: el manejo de bases de datos biológicas; procesos metabólicos y genética de poblaciones; Inteligencia artificial, entre otros más [10].

Son muchas las áreas beneficiadas con la explotación de la bioinformática, y de igual forma son diversas las ramas de investigación que esta posee, siendo algunas: el análisis de secuencias, anotación de genomas, biología evolutiva computacional, análisis de la expresión génica, análisis de la expresión de proteínas, análisis de mutaciones en el cáncer, genómica comparativa, modelado de sistemas biológicos, acoplamiento proteína-proteína, etc. [7]

En el 2008 ya existían alrededor de 1,200 tests [8] [9] pero eran limitados y costosos, por lo que las empresas buscaban la forma de disminuir el coste para facilitar el acceso a la población desde la comodidad de sus hogares. Fue gracias a los avances tecnológicos en secuenciación que permitió obtener secuencias en menor precio y tiempo con la llamada “Next Generation Sequencing” [11] [12].

### 2.1 Empresas del Entorno Genómico

Actualmente existe un alto número de empresas que ofertan servicios de test genéticos directos al consumidor, con el fin de facilitar un diagnóstico sobre la predisposición de un paciente a padecer ciertas enfermedades.

Una de las empresas de mayor repercusión en este sector y que es importante mencionar es 23andMe, la cual ha sido una gran propulsora en el diagnóstico genético. Esta empresa estadounidense cuenta con una alta gama de servicios. Las informaciones que ofrecen están relacionadas con la historia genética (ancestros) y salud personal (riesgo sobre enfermedades), los cuales son presentados en su gran mayoría en base a probabilidades [13] [14].

De igual forma en España se han desarrollado empresas de este tipo como por ejemplo Genotest o Imegen, todas con el mismo objetivo de brindar a los usuarios test genéticos de una manera sencilla y sobre todo aportándole un diagnóstico que le permita tomar acciones de lugar para mejorar su calidad de vida.

Algunas de las empresas analizadas se encuentran en la tabla 2.0, en donde estuvimos evaluando distintos puntos de interés para lanzar este tipo de empresa.

Empresa	Ciudad	Promociona Tiempo de Entrega?	ANUNCIA: Precio Std. (Disponible desde...)	Firma de Consentimiento (Caso Si que implica?)	Avalado por Médico	Requisitos de Uso	Envío de Muestras al Lab. Incluido en Precio (Especificado)?	Facilidad para Comprar Diagnosticos Geneticos	Cómo ofrece el Servicio (Análisis Genético)?
BIOARRAY	Alicante	NO	NO	SI	SI	SI	SI	NO	Laboratorio/Clinica
CIALAB	Alicante	NO	NO	SI	SI	SI	SI	NO	Laboratorio/Clinica
Genomic Genetics	Barcelona	NO	NO	NO	SI	SI	NO	NO	Laboratorio/Clinica
23andMe	California, US	SI	SI	SI	NO	SI	NO	SI	Web
deCODEme	Islandia	SI	SI	SI	NO	NO	NO	SI	Web
CircaGen	Madrid	NO	NO	SI	SI	NO	NO	NO	Laboratorio/Clinica
GENOTEST	Madrid	NO	SI	NO	NO	SI	NO	SI	Web
IMEGEN	Valencia	NO	NO	SI	SI	SI	NO	SI	Web
CAGT	Zaragoza	NO	SI	SI	SI	SI	SI	SI	Web

Tabla 2.1 Comparativa de Empresas que ofrecen test genéticos directos al consumidor

Son muchos los puntos que se deben de tratar, debemos ser conscientes de que estos servicios deben ser de ayuda para los usuarios. Las empresas cada vez más se acercan al público y rompen con esos muros que impedían avanzar en este campo, contemplando todos los aspectos como: éticos, morales, legales, entre otros.

Estas empresas trabajan con información delicada y sensible, por lo que deben contar con altos niveles de seguridad y protección de los datos, además de tomar todas las medidas legales de lugar. Por ejemplo para estos estudios se necesita la adopción de las siguientes leyes:

- ⇒ Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal (Texto consolidado a fecha 5 de marzo de 2011).
- ⇒ Periodos de Conservación de Datos Personales en Investigación Biomédica
- ⇒ Ley General de Sanidad 86

Son muchos los factores a tomar en cuenta para la puesta en marcha de servicios que te brindan una predisposición para padecer ciertas enfermedades. Con el paso de los años conoceremos más y más de nuestro genoma por los avances que siguen surgiendo.

## 2.2 Sensibilidad al Alcohol

La Sensibilidad al Alcohol es el grado de metabolización el alcohol ingerido. Provocando que a mismas concentraciones de alcohol el efecto sobre el organismo tenga efectos claramente diferentes. Este grado de sensibilidad está claramente diferenciado poblacionalmente e individualmente. Teniendo asociado un componente genético y metabólico directo.

Diversos estudios a través de los años han sustentado y justificado la implicación genética para padecer esta sensibilidad.

Para ejemplificar estos casos, vemos que en el año 2003 se publica (presenta) en una Revista de Neurología el artículo [15] que nos explicaba la relación genética con esta sensibilidad.

Con el paso de los años se continuó investigando y profundizando; es para los años 2010 y 2011 en donde se presentaron muchas aportaciones, así como nuevos genes que sustentaban estos estudios. Entre las publicaciones se encuentran [16] [17].

Los avances médicos han demostrado que con el transcurso de los años que por nuestros genes podemos ser sensibles o no al alcohol.

Durante las etapas de búsqueda y estudio encontramos una variante particular dentro del manejo de las variaciones, puesto que algunas solamente eran de importancia cuando eran tomadas y/o analizadas en conjunto, siendo este fenómeno conocido como Haplotipos.

### 2.3 Haplotipos

Un haplotipo está constituido por un grupo de SNP de un gen que, al estar muy cercanos, tienden a heredarse juntos. Esto quiere decir que los alelos de un haplotipo no son separados en la fase de recombinación y pueden transmitirse “en bloque”, pudiéndose establecer, de esta forma, combinaciones de variaciones dentro de un gen que pueden estar implicadas en el fenotipo [18].

“Un haplotipo es el conjunto de alelos sobre un mismo cromosoma, el cual se puede construir a partir de diversos marcadores genéticos, entre ellos las repeticiones cortas en tandem o STR por su sigla en inglés (Short Tandem Repeats)” [46].

En el genoma humano, los haplotipos tienden a abarcar una distancia de, aproximadamente, 60,000 pares de bases (pb) y contienen hasta 60 SNP que son transmitidos en grupo de padres a hijos [18]. La ventaja principal de utilizar los haplotipos en nuestros estudios de asociación genética es que nos permiten distinguir mejor las variantes o alelos de un determinado gen, por lo que aumenta el poder de detección de los tests genéticos.

En la actualidad son muchos los interesados en estudiar y aprender sobre los haplotipos, y en nuestro caso encontramos información valiosa generada por el Proyecto HapMap.

#### HapMap

Debido a la identificación reciente de los haplotipos en el genoma humano y los cambios más o menos variables con respecto a la población, existen grupos de investigación enfocados en esta área. Uno de ellos es el Proyecto Internacional HapMap, el cual consiste en un esfuerzo multi-país para identificar y catalogar las similitudes y diferencias genéticas en los seres humanos [19].

La utilización de esta información de HapMap permitirá a los investigadores encontrar los genes que afectan la salud, la enfermedad y las respuestas individuales a los medicamentos y los factores ambientales [19]. “El objetivo principal de HapMap es comparar las secuencias genéticas de diferentes individuos para identificar regiones cromosómicas donde se comparten variantes genéticas.”

## 2.4 Manipulación de Datos (BD Genómicas)

Son muchos los estudios llevados a cabo por la comunidad médica y en tecnologías de la información para encontrar soluciones a los grandes problemas de gestión de las bases de datos genómicas, principalmente por tener que manejar/administrar fuentes de datos de gran tamaño, aumentándonos la inversión en tiempo, almacenamiento, búsqueda, entre otros de poca utilidad.

Con las herramientas actuales y los motores de búsqueda, como son: dbSNP [20] [21], OMIM [3], entre otras más son incapaces de resolver problemas específicos, somos conscientes de que hay mucho donde mejorar, por ejemplo el uso de bases de datos refinadas (filtradas) nos ayuda a optimizar el rendimiento y ha obtener una mayor precisión y calidad.

Desarrollando las tareas de SILE avanzamos y damos un paso hacia adelante, porque obtenemos una definición precisa de lo que se quiere, buscando la forma de ir mejorando cada vez más. La complejidad existente en la relación de los genes (variaciones, mutaciones) y una población específica es un punto a tratar, puesto que vemos como unos son más frecuentes en una población u otra. Con la integración de aspectos poblacionales dentro del modelo conceptual del genoma humano se brindan servicios personalizados y más confiables.

## 2.5 Bases de Datos basadas en Modelado Conceptual

Las bases de datos utilizadas son generadas a partir de modelos conceptuales (del genoma) que nos permiten abordar todas las informaciones necesarias para trabajar en esta área. Somos conscientes de que la utilización de modelos conceptuales dentro de la bioinformática ha sido leve. Por lo que han surgido muchas fuentes de datos de información genómica sin una estructura concreta [31].

A pesar de que ha faltado más estímulo para tratar con este tema, existen algunas propuestas orientadas al estudio genómico. Los primeros trabajos de modelado conceptual del genoma, los llevó a cabo Paton [47] [48], en la cual ya describían el genoma de célula eucariota, la interacción entre proteínas, el transcriptoma, entre otros. Otro proyecto en el que aplicaron principios de modelado conceptual al genoma fue con Ram [49], estos estuvieron más enfocados al contexto de las proteínas.

El uso del modelado conceptual en este contexto es de gran utilidad en nuestros días, por lo que cada vez más son muchos los que deciden colaborar en este campo. Un ejemplo de ello es el Grupo de Investigación ProS [5] de la Universidad Politécnica de Valencia, el cual ha dedicado una de sus líneas para desarrollar y potenciar modelados conceptuales en este enfoque.

## Capítulo 3

# Solución Propuesta

### 3.1 Modelado de Procesos de Negocio de una Empresa Genómica

#### 3.1.1 Descripción de la Empresa & Organigrama

La empresa GemBiosoft es un spin off de la Universidad Politécnica de Valencia, fundada en el 2010. GemBiosoft tiene una estrecha relación con el centro de investigación PROS y la experiencia en su desarrollo dirigido por modelos [6].

Entre finales de los noventa y principio de los 2000 aplicaron ideas con éxito centrándose fundamentalmente en los Sistemas de Información e Inteligencia Ambiental. El objetivo principal esta empresa es definir el Modelo Conceptual del Genoma Humano, obteniendo un esquema preciso que gestione los datos genómicos y consolide la información que está en continuo crecimiento en el campo del genoma.

GemBiosoft se creó con la idea de plasmar todas estas ideas y experiencias en la industria Bioinformática. Se ha enfocado en desarrollar y explorar este reto en los Sistemas de Información Genómicos, donde los Sistemas de Información comienza a ser un componente imprescindible en la gestión de datos genómicos y el esquema conceptual del genoma humano el cual encauza el proceso de captura, almacenamiento y explotación de datos.

La misión de GemBiosoft es fomentar el uso de los sistemas de información en el área de la Bioinformática, proporcionando soluciones de alta calidad para la gestión de datos genómicos y los cuidados sanitarios informatizados. Su visión es convertirse en una empresa líder en el diseño y optimización de sistemas de información de calidad para las áreas de la genómica y la salud.

GemBiosoft cuenta con un equipo formado por personas interdisciplinarias – ingenieros y doctorados – capacitados para ofrecer soluciones a empresas de investigación y desarrollo aplicando tecnologías de la información en la bioinformática y la salud.

Entre sus participantes se encuentran: Centro de Investigación en Métodos de Producción de Software (PROS) de la Universidad Politécnica de Valencia, PrincipiaTech y el Instituto de Medicina Genómica (IMEGEN).

Sus investigaciones están orientadas al Modelado Conceptual y a Publicaciones Científicas.

Con respecto a la oferta comercial GemBiosoft dispone de dos líneas principales de negocio:

1. Plataforma Software para la gestión y operación de datos genómicos, con el apoyo de nuestro Esquema Conceptual del Genoma Humano.
2. Auditoría, consultoría y desarrollo de servicios respaldados por nuestros expertos.

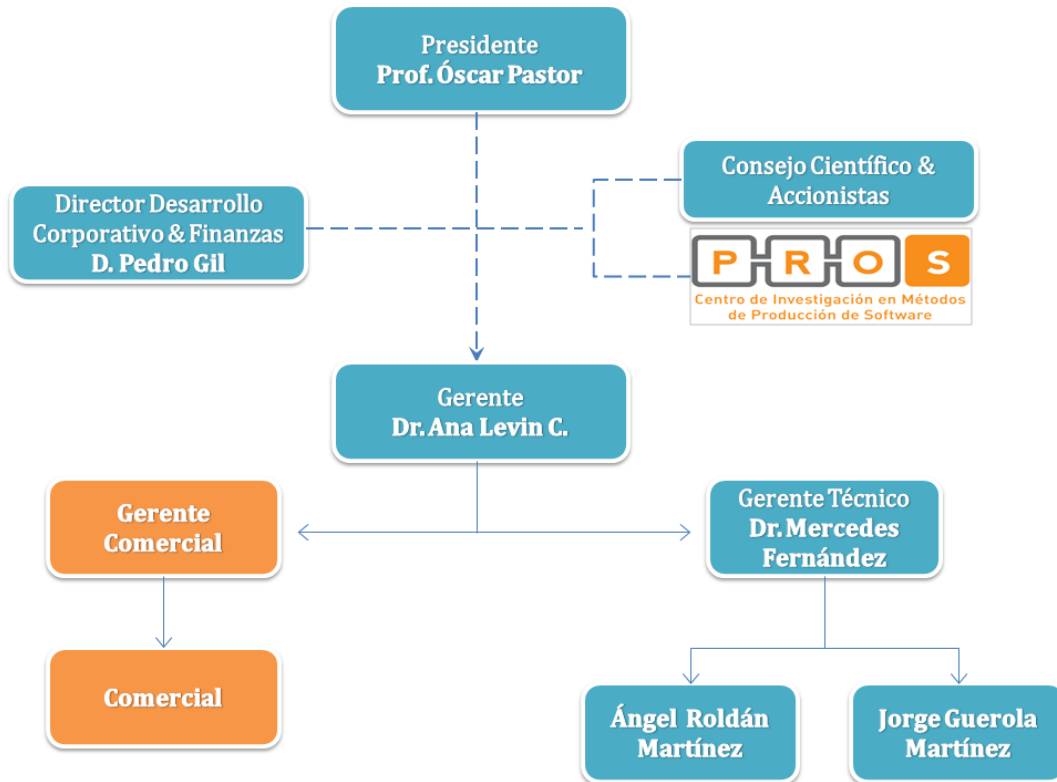


Figura 3.1 Organigrama de la Empresa



Figura 3.2 Logo de la Empresa

### Descripción Textual del Caso (*Textual Case Description*)

GemBiosoft es una empresa enfocada en desarrollar y explorar el reto existente actualmente en los Sistemas de Información Genómicos, y se mantiene trabajando día a día en la búsqueda de mejores soluciones en el tratamiento y Diagnóstico de enfermedades por medio del estudio/comparación de las secuencias de genes de los pacientes.

#### Descripción del Proceso

GemBiosoft cuenta con una oferta de diagnósticos genómicos de algunas enfermedades muy comunes en nuestros días y para ello cuenta con un sitio web donde los usuarios pueden acceder sin necesidad de estar registrados para consultar toda la información sobre las enfermedades cubiertas, sus características, tratamiento, y sobre todo cuenta con tutoriales y videos del proceso, permitiendo esto dejar claros a los usuarios de cómo se manejan estos estudios.

Existen 3 formas de acceso al sistema, ya sea como usuarios del tipo cliente, proveedor y/o administrador. Dependiendo el caso se realizan unas u otras tareas.

Al usuario realizar su autenticación accede a una zona o área donde realiza todas las operaciones que le conciernen.

**Caso de Usuario Cliente (Paciente):** Al acceder a su suite de cliente puede hacer contrataciones de los servicios ofertados, donde este procede a seleccionar los servicios que le interesa y luego realiza el pago de los mismos. Puede también revisar el perfil público que consiste en visualizar información sobre las enfermedades tratadas, ver videos o tutoriales. En su área ver las notificaciones que va recibiendo ya sean sobre sus resultados (diagnósticos) o cualquier alerta/mensaje de interés. Están las opciones de edición de perfiles, ver/consultar resultados de diagnósticos, contando en este último con la ventaja de mantener un histórico de todos sus los diagnósticos reportados en el servicio.

**Caso de Usuario Proveedor:** esta zona le permite a los proveedores acceder para llevar a cabo el reporte de las muestras que recibe por parte del cliente. Tan pronto llegan las muestras a su centro estos proceden a hacer la activación de la muestra recibida simplemente registrando el código de la muestra para realizar el cambio de estado en que se encuentran las muestras. Cuenta con las opciones de edición de perfil, asignar el estatus en que están las muestras y consultar todas las muestras que tienen tanto en proceso como registradas.

**Caso de Usuario Administrador:** este es un usuario que tiene como objetivo realizar las labores de mantenimiento y publicación de resultados, este usuario al trabajar con la secuencia que recibe del proveedor procede a analizarlos por medio de DIAGEN y cuando tiene el Diagnóstico realiza la publicación dentro del sitio web, de esta forma se notifica a los clientes (pacientes) por vía email que ya sus resultados están disponibles. Dentro de su suite cuenta además con las facilidades para realizar la publicación de publicidad y/o promociones sobre enfermedades nuevas (disponibles de estudio) y ver/consultar reportes de los pagos y usos que han sido realizados dentro de un periodo de tiempo (personalizado).

La empresa se mantiene trabajando en el estudio de nuevas enfermedades, y mejorando en el día a día de manera constante para brindar un servicio con el que los clientes se sientan cómodos y sobre todo tengan la confianza/seguridad de acceder a las facilidades y/o servicios que brindan.



### 3.1.2 Diagramas de los Procesos de Negocio (BPMN)

#### El Proceso del Diagnóstico Genómico

Actualmente se ofertan análisis genéticos con el fin de conocer y/o detectar el padecimiento o predisposición de una persona para contraer alguna enfermedad de carácter hereditario.

Por ello, cada vez se buscan los mecanismos (medios) para generar diagnósticos genéticos que permitan a los clientes (pacientes) acceder a estos análisis con el fin de orientar y brindar un tratamiento de prevención.

La Figura 3.3 nos muestra un diagrama BPMN que describe el proceso de Diagnóstico Genético (desde cuando el cliente (paciente) decide solicitar el servicio hasta que recibe el resultado de su Diagnóstico Genético). En este proceso participan tres actores: 1) Cliente (paciente) que solicita el servicio para determinar si padece o no una enfermedad; 2) La Empresa, en este caso GemBiosoft que se encarga de gestionar y realizar el Diagnóstico; 3) Proveedores que en este caso preparan las secuencias de referencia para el análisis.

El proceso general se inicia en el momento que el cliente (paciente) accede al servicio web y solicita el análisis genético (t1). La empresa (GemBiosoft) procesa esta solicitud y procede a enviar el colector de muestra al cliente (t2). Cuando el cliente recibe el colector procede realizar la activación en el servicio web utilizando su identificador (t3), luego coloca las muestras y lo envía nuevamente a la empresa (t4). Al recibir las muestras la empresa comprueba y/o valida que cumple con los requerimientos necesarios para el estudio y notifica al cliente sobre la recepción (t5). Se procede a determinar que proveedor realizará la secuenciación y se le envía la muestra (t6). El proveedor seleccionado recibe las muestras y le notifica a la empresa del estado (t7). Se inicia la preparación de la secuencia de referencia llevando a cabo los pasos pertinentes del proveedor (t8). Se realiza el envío de la secuencia de referencia (fichero) a la empresa (t9). La empresa notifica al proveedor de la recepción efectiva de la secuencia y procede a realizar el diagnóstico (t10). Luego se procede a generar el informe del diagnóstico definitivo (t11). La empresa procede a la publicación en el servicio web del diagnóstico genético (resultado) del cliente (paciente) y a su vez le notifica de la disponibilidad para que este pueda consultarlos (t12). Para finalizar el proceso el cliente (paciente) accede al servicio web y consulta sus diagnósticos (resultados) (t13).

El objetivo la empresa es brindar servicios web que permitan a los clientes realizar la compra de test genéticos de forma sencilla y asequible con el fin de realizar diagnósticos que le ayuden a conocer la predisposición y/o padecimiento de una enfermedad hereditaria, por medio del tratamiento de secuencias de referencia del ADN que lleven a determinar resultados de utilidad (calidad).

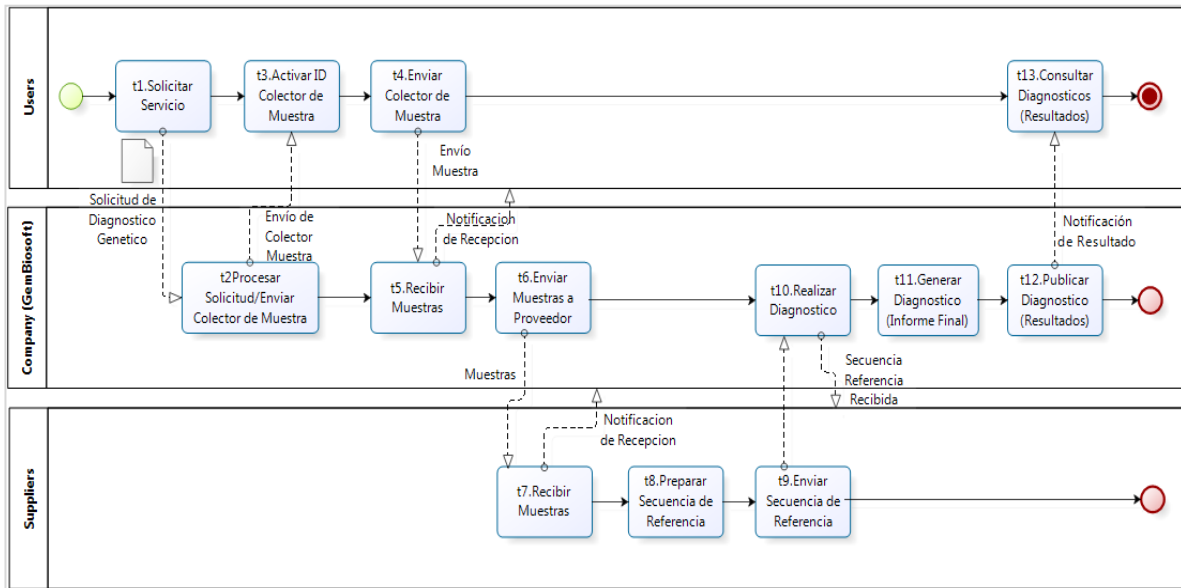


Figura 3.3 El Proceso del Diagnóstico Genético

La empresa ofrece una variedad de servicios, pero para el caso de los test genéticos se debe contar con los Colectores de Muestra, ya que tan pronto se recibe una solicitud se debe de hacer llegar lo más pronto posible al cliente (paciente).

Entre los procesos manejados se tiene en cuenta la reposición de estos equipos, los cuales son suplidos por un determinado proveedor.

La Figura 3.4 nos muestra el diagrama BPMN de los procesos llevados a cabo para la reposición de los colectores de muestra, en esta tarea se cuenta con la participación de dos actores: 1) La Empresa GemBiosoft, la cual mantiene un stock de colectores y cuando se encuentran por debajo del mínimo proceden a realizar sus pedidos; 2) Proveedores son los encargados de suplir de los equipos necesarios.

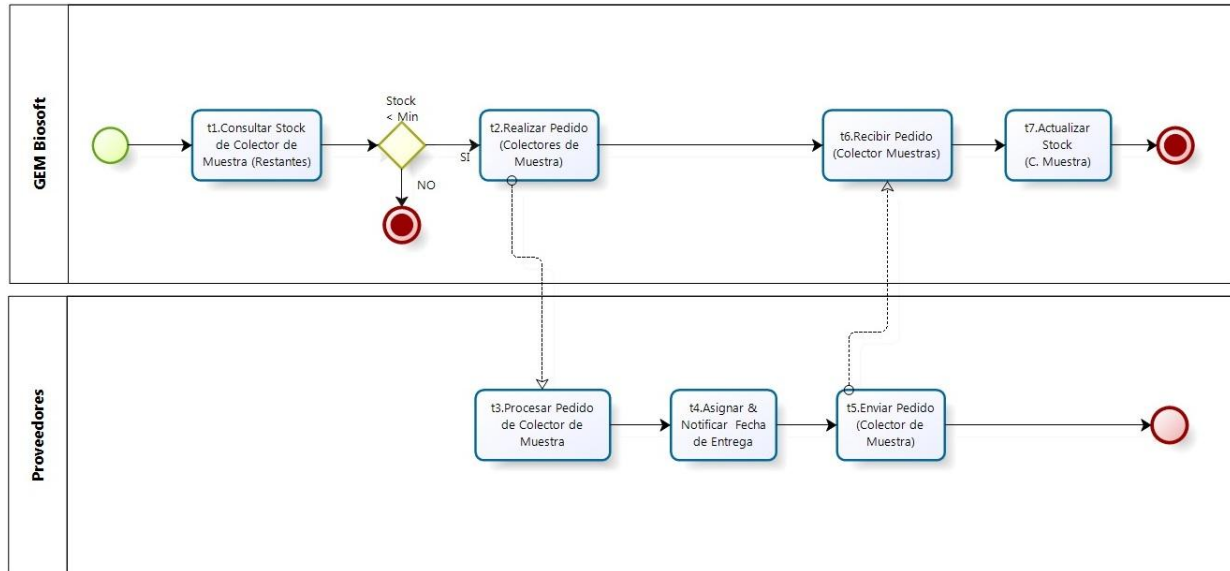


Figura 3.4 El Proceso de Reposición de Equipos (Colector de Muestras)

El proceso general se inicializa cuando la empresa realiza la consulta su stock de colectores de muestra y en caso de estar en el mínimo se realiza el siguiente proceso, en caso contrario se finaliza (t1). Si se da el caso “Stock < Min” se procede a hacer el pedido al proveedor (t2). El proveedor seleccionado recibe el pedido y lo procesa (t3). El proveedor asigna y notifica la fecha de entrega (t4). Se procede a enviar el pedido hasta la empresa (t5). La empresa recibe los colectores de muestra solicitados (t6) y finaliza el proceso actualizando el stock con las nuevas adquisiciones (t7).

## Diagramas de los Procesos de Negocio

La figura a continuación nos presenta el flujo que se realiza para los servicios web ofertados por la empresa:

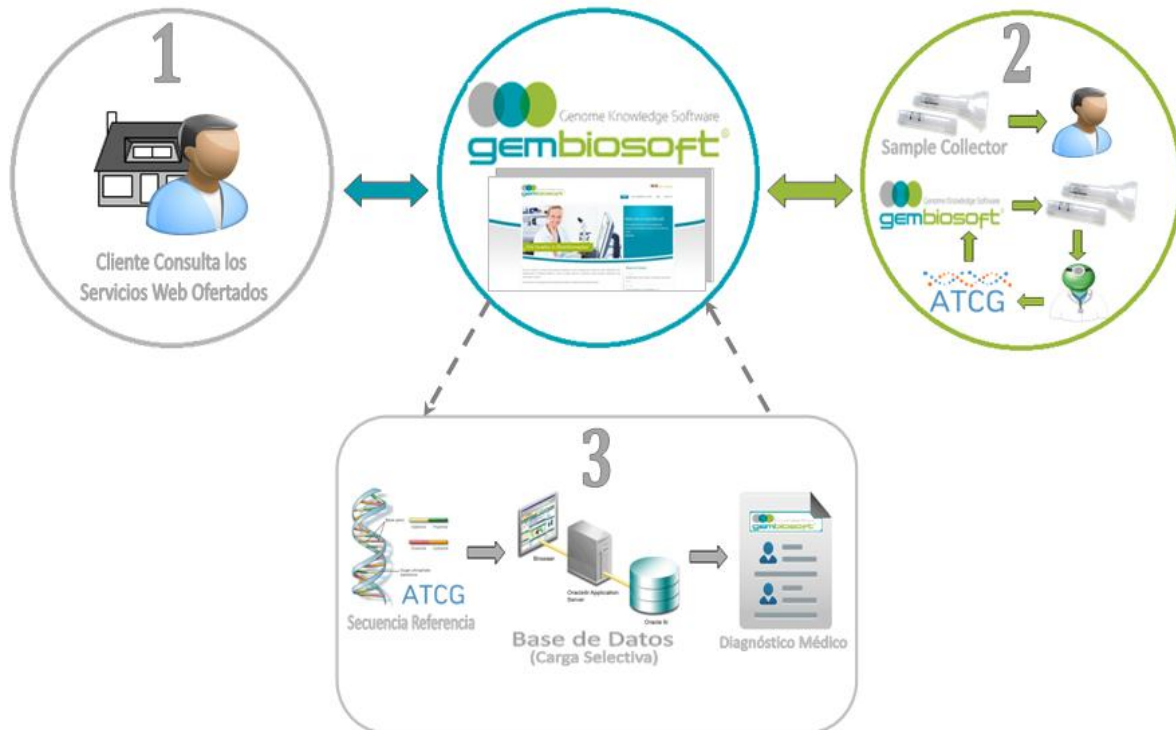


Figura 3.5 Procesos del Servicio Web

1. El primer paso y que inicializa todo es cuando el usuario consulta los servicios web ofertados. Y la decisión de comprar alguno de los test genéticos activan/habilitan la interacción con la empresa.
2. La empresa envía al usuario (paciente) un respectivo Colector de Muestra donde este deberá hacer llegar la muestra a analizar.
3. El usuario recibe el colector y lo rellena con la muestra correspondiente. Al finalizar este paso procede a enviar la muestra junto con el consentimiento informado para los test a la empresa.
4. La empresa comprueba que las muestras estén en condiciones, luego procede a seleccionar el proveedor que realizará la secuenciación de referencia; y le hace llegar la muestra.
5. El proveedor (secuenciador, genetista) toma la muestra y prepara la secuencia de referencia, al finalizar este proceso envían la secuencia por medio de un fichero (FASTA, Sequencher u otro) a la empresa.
6. La empresa toma la secuencia de referencia, y realiza el diagnóstico con la comparación en su base de datos (carga selectiva), lo que genera un diagnóstico médico para el usuario (paciente).
7. La empresa realiza la publicación en su sitio web de los diagnósticos y le notifica a sus usuarios para que accedan a ver sus resultados.
8. El usuario (paciente) accede y consulta sus diagnósticos (resultados).

## 3.2 Especificación de Requisitos: Visión General

### 3.2.1 Lista Definitiva de Requisitos del Nuevo Sistema (LDR)

El sistema debe permitir:

No.	Requisito	Tipo	Usuarios/Roles				Prioridad
			Anónimos	Clientes (Pacientes)	Administrador	Proveedores	
1	La creación de clientes (pacientes) para el proceso de solicitud de diagnósticos médicos. (Datos personales, entre otros -email, nombres, apellidos, sexo, fecha de nacimiento, dirección, código postal, teléfono-). a. Esto implica la creación de usuario para acceder al sistema.	Funcional		X	X		Alta
2	Contar con una suite (sección) del cliente (paciente) donde podrá consultar sus diagnósticos y visualizar una tabla con los resultados de los servicios contratados (Intolerancia a la lactosa, Diabetes, etc.), para esto debe estar logueado en el sistema a. Dentro de esta suite/zona se debe permitir acceder a la zona de compra y siempre indicar las facilidades para añadir cualquier servicio al carrito de compra.	Funcional		X			Alta
3	La selección de los servicios contratados por el cliente para el proceso de análisis/diagnóstico	Funcional		X			Alta
4	El pago de los servicios contratados por medio de Paypal con tarjetas de crédito, debe existir la facilidad de pago por medio de transferencias bancarias (cuentas)	Funcional		X			Alta
5	La consulta y/o generación de reportes sobre las siguientes informaciones: a. De los diagnósticos obtenidos (Resultados)	Funcional		X	X		Alta
6	Publicar resultados del análisis de los clientes (pacientes) en la web. Este a su vez que sea publicado debe enviar un email de notificación al cliente informándole de los resultados	Funcional			X		Alta
7	La Creación de proveedores para el proceso de reportes sobre el estado de las muestras, datos a tomar del proveedor: id_proveedor, nombre, descripción, especialidad, fecha_entrada, Dirección, Teléfono, representante, email, estatus, nota, etc.	Funcional			X		Alta
8	Contar con una zona de Proveedores, en esta sección los mismos podrán acceder para registrar la recepción de las muestras por parte de los clientes (pacientes) y reportar el estado de las muestras	Funcional				X	Alta
9	El inicio de sesión o login. (Se debe considerar secciones/zona para clientes (pacientes), para administración y para hospitales (proveedores, genetistas))	Funcional	X	X	X	X	Media
10	1. La visualización de consultas/reportes de las muestras registradas/proceso. a. Esto debe estar disponible para los proveedores y administradores.	Funcional			X	X	Media

Tabla 3.6 Lista Definitiva de Requisitos del Nuevo Sistema (LDR) (1)

El sistema debe permitir:

No.	Requisito	Tipo	Usuarios/Roles				Prioridad
			Anónimos	Cientes (Pacientes)	Administrador	Proveedores	
11	La consulta de información (básica) sobre el proceso de diagnóstico e información sobre las enfermedades cubiertas por la empresa sin necesidad de registro en la aplicación (usuarios anónimos). Brindar la facilidad de un perfil público donde se pueda ver información detallada de enfermedades, tutoriales, videos, etc. a. Mantener las opciones de compra y/o añadir al carrito para estos usuarios, y en caso de ser solicitados indicarles los pasos de lugar (inicio sesión/registro) para el proceso de adquisición del servicio.	Funcional	X	X	X	X	Media
12	La edición de los perfiles de los clientes (modificar datos)	Funcional		X			Media
13	Activar y/o indicar dentro del sistema cuando reciba el Colector de Muestra por parte de la empresa, esto con el fin de llevar un seguimiento de las muestras en todos los escenarios presentados a. El cliente (paciente) accederá tan pronto reciba el Colector de Muestra y realizará la activación. La cual servirá de notificación para el control de localización. i. Se debe contemplar al igual que cuando reciba el Colector de Muestra, cuando proceda a retomararlo a la empresa informarlo/notificarlo en el sistema.	Funcional		X			Media
14	Mantener un histórico de los análisis (estudios) realizados y diagnósticos obtenidos (con sus respectivos resultados)	Funcional		X			Media
15	Consultar todos los pagos y usos realizados por los clientes (pacientes). Con la facilidad de ver los reportes en forma mensual, quincenal, o con fecha personalizada a. Acceso sólo a la Administración.	Funcional			X		Media
16	Ingresar/publicar las muestras recibidas (Activación de muestras recibidas), debe permitir el ingreso de la identificación del Id Cliente (muestra recibida) y que de esta forma active o cambie el estado en que se encuentra la muestra	Funcional				X	Media
17	La notificación al cliente de cuando sus resultados estén disponibles en la web a. Cuando los resultados del análisis se encuentren reportados en la web se debe de realizar esta notificación al cliente (paciente) vía email para que pase a revisarlos cuando desee.	Funcional		X			Baja
18	La impresión de los diagnósticos médicos por parte del cliente (paciente) en formato *.pdf	Funcional		X	X		Baja
19	El manejo de promoción y/o publicidad sobre los nuevos servicios (disponibles) por parte de la empresa (estos deben estar en la parte principal)	Funcional			X		Baja
20	Contar con información detallada sobre las enfermedades cubiertas, además de facilitar tutoriales que muestren el tratamiento/forma de enviar las muestras para los tests	No Funcional	X	X	X	X	Baja
21	El cierre de la sesión (usuario) si la cuenta se mantiene inactiva por un periodo de 10 minutos	No Funcional		X	X	X	Baja

Tabla 3.6 Lista Definitiva de Requisitos del Nuevo Sistema (LDR) (2)

### 3.2.2 Casos de Uso & Diagrama de Clases

A partir del análisis realizado vemos los requisitos generales que deberá cumplir el servicio web para brindar a los usuarios (pacientes) una aplicación que les permita la contratación de test genéticos para conocer la predisposición o padecimiento de alguna enfermedad genética. En lo adelante estaremos presentando los distintos casos de uso clasificados por los usuarios que participarán en el sistema.

**Diagrama de Caso de Uso General**

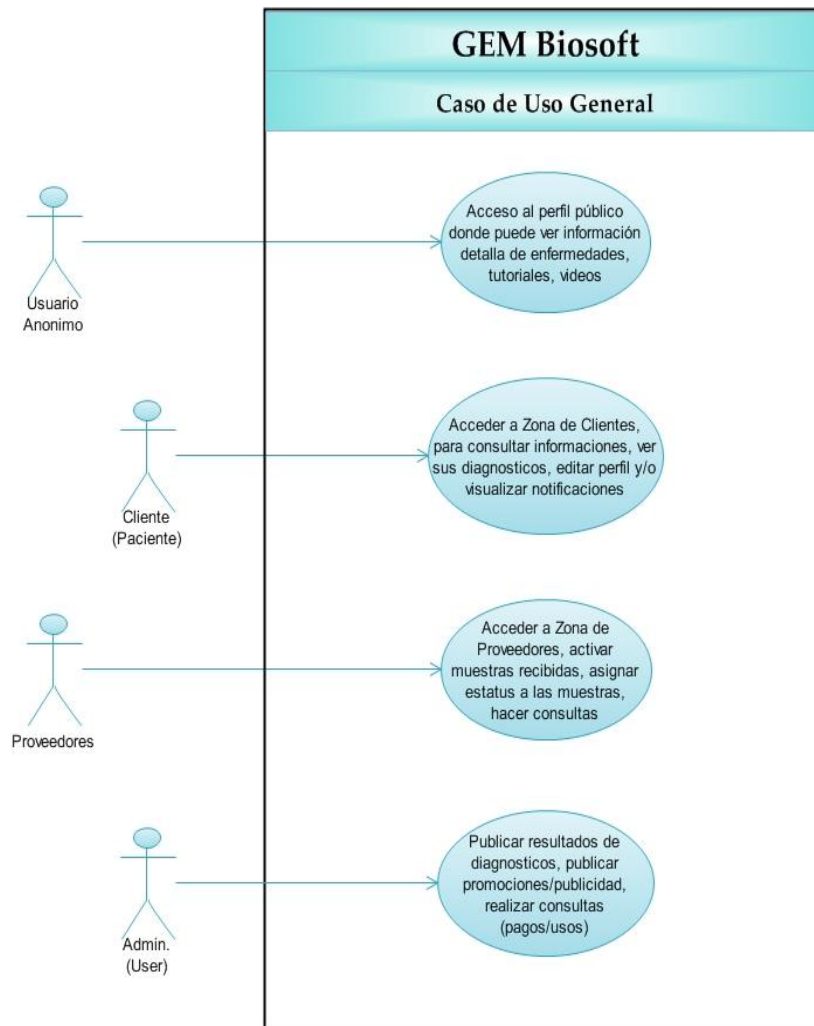


Figura 3.7 Caso de Uso General del Sistema

En el diagrama presentado mostramos los actores del sistema y un breve resumen de las tareas que realizará dentro del servicio web.

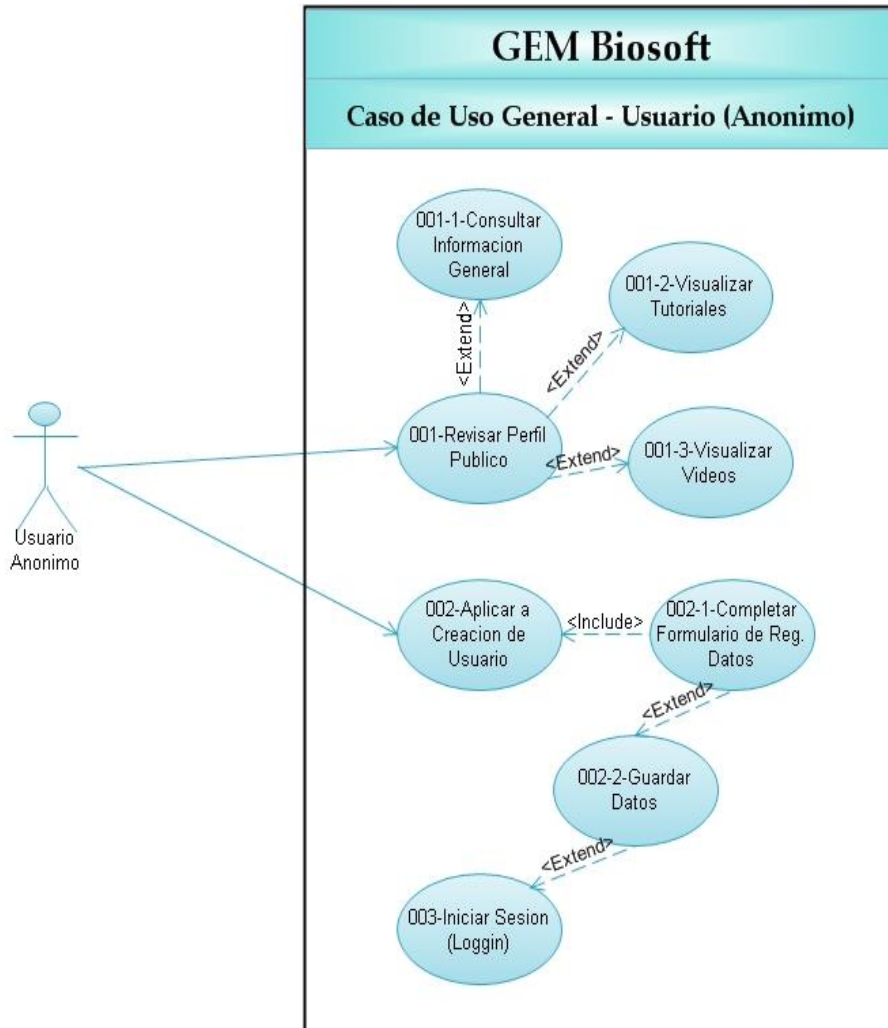


Figura 3.8 Caso de Uso General: Usuario Anónimo

Nombre CU	001-Revisar Perfil Público
Propósito	Obtener información sobre los servicios que ofrecen
Precondiciones	Acceder a la web
Postcondiciones	Se conoce los productos y servicios ofertados por la empresa
Casos de Uso que incluye	Consultar Información General, Visualizar Tutoriales, Visualizar Videos

**El caso de uso está formado por los pasos siguientes:**

1. El usuario accede a la web, por lo que navega todo el perfil público.
2. El usuario puede visualizar distintas opciones, como por ejemplo: a) Consultar la información general sobre las enfermedades; b) Visualizar los tutoriales disponibles; c) Visualizar los videos.
3. Conforme la elección del usuario irá conociendo e informándose sobre los procesos y usos del servicio web.



Nombre CU	002-Aplicar a Creación de Usuario
Propósito	Crear una cuenta de usuario para la web, con la cual tendrá su área de usuario para realizar tareas más específicas (compras, ver diagnósticos, etc.)
Precondiciones	Rellenar todos los datos (obligatorios) en el formulario de registro
Postcondiciones	Usuario y contraseña de acceso a la web enviada a su email
Casos de Uso que incluye	Completar Formulario de Registro de Datos (extiende guardar datos e iniciar sesión)

**El caso de uso está formado por los pasos siguientes:**

1. El usuario dentro de la web procede a rellenar el formulario de registro de datos.
2. Cuando todos los campos están llenos se procede a guardar los datos.
3. El sistema envía un email con el nombre de usuario y contraseña para acceder a la web.

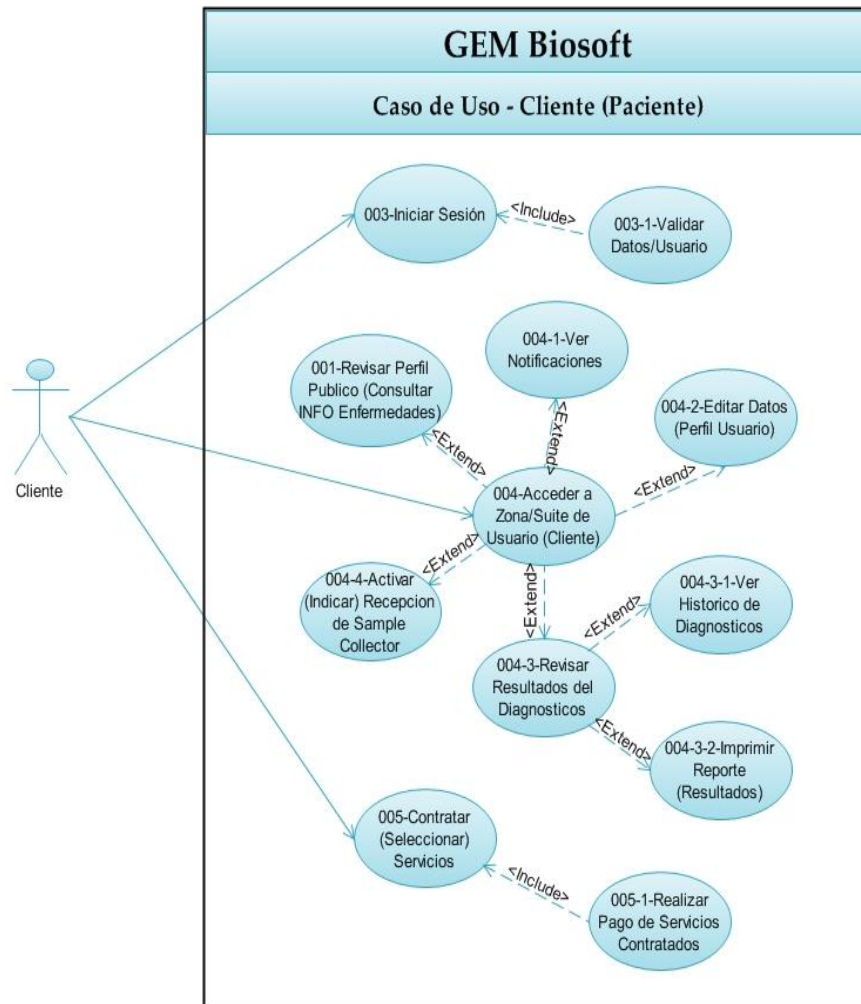


Figura 3.9: Caso de Uso General: Cliente (Paciente)

Nombre CU	003-Iniciar Sesión
Propósito	Acceder a la zona de usuario (acceder a la web como usuario registrado)
Precondiciones	Introducir nombre de usuario y contraseña
Postcondiciones	Acceso a zona de usuario
Casos de Uso que incluye	Validar datos de usuario

**El caso de uso está formado por los pasos siguientes:**

1. El usuario introduce su nombre de usuario y contraseña.
2. El sistema realiza la validación de los datos del usuario, en caso correcto accede a la zona de usuario. En caso contrario, le pide que vuelva a intentarlo (con un límite de 3 intentos).

Nombre CU	004-Acceder a Zona/Suite de Usuario
Propósito	Realizar las tareas de gestión (usuario) como: ver diagnósticos, editar datos, ver notificaciones, etc.
Precondiciones	Realizar el inicio de sesión
Postcondiciones	Gestionar la cuenta de usuario, Consultar estado de productos contratados, entre otros
Casos de Uso que incluye	Revisar perfil público, Ver notificaciones, Editar datos (perfil), Activar recepción del colector de muestra, Revisar resultados de diagnósticos (ver histórico y/o imprimir diagnóstico)

**El caso de uso está formado por los pasos siguientes:**

1. Tras realizar el inicio de sesión, accedemos a nuestra zona/suite de usuario.
2. Una vez nos encontramos en el área de usuario, vemos distintos tipos de opciones:
  - a) Revisar perfil público: con esta opción se navega en la web, viendo todos los contenidos sobre los servicios que se ofrecen.
  - b) Ver notificaciones: en esta parte se mantiene informado al usuario sobre todas las novedades dentro de su cuenta.
  - c) Editar datos (perfil): facilidad para modificar los datos de usuario en algún momento dado.
  - d) Activar recepción del colector de muestra: sección muy importante, ya que cuando el usuario (paciente) recibe el colector de muestra, debe de activar un código que se encuentra en él mismo, esto es con fines de seguimiento de la muestra.
  - e) Revisar resultados de diagnósticos: esta es la opción para visualizar los resultados obtenidos en el diagnóstico genético.
    - I. Aquí se encuentra la facilidad para consultar el histórico de los diagnósticos solicitados e;
    - II. Imprimir los resultados obtenidos.

Nombre CU	005-Contratar (Seleccionar) Servicios
Propósito	Realizar la compra de servicios para el diagnóstico genético
Precondiciones	Seleccionar el servicio deseado
Postcondiciones	Solicitud de compra realizada
Casos de Uso que incluye	Realizar Pago de Servicios Contratados

**El caso de uso está formado por los pasos siguientes:**

1. El usuario cuando está en la web puede seleccionar cualquiera de los servicios ofertados y comprarlos (para esto es opcional el inicio de sesión).
2. Si el usuario está registrado se le solicita que inicie la sesión, en caso contrario, se procede a rellenar un formulario con todos sus datos para el proceso de envío y método de pago.
  - a. Realizar pago de servicios contratados: En esta funcionalidad el usuario selecciona el método de pago (Visa, transferencia bancaria, otros) y realiza el pago correspondiente.
3. El sistema le envía un email de confirmación al usuario con toda la información relacionada a su compra.

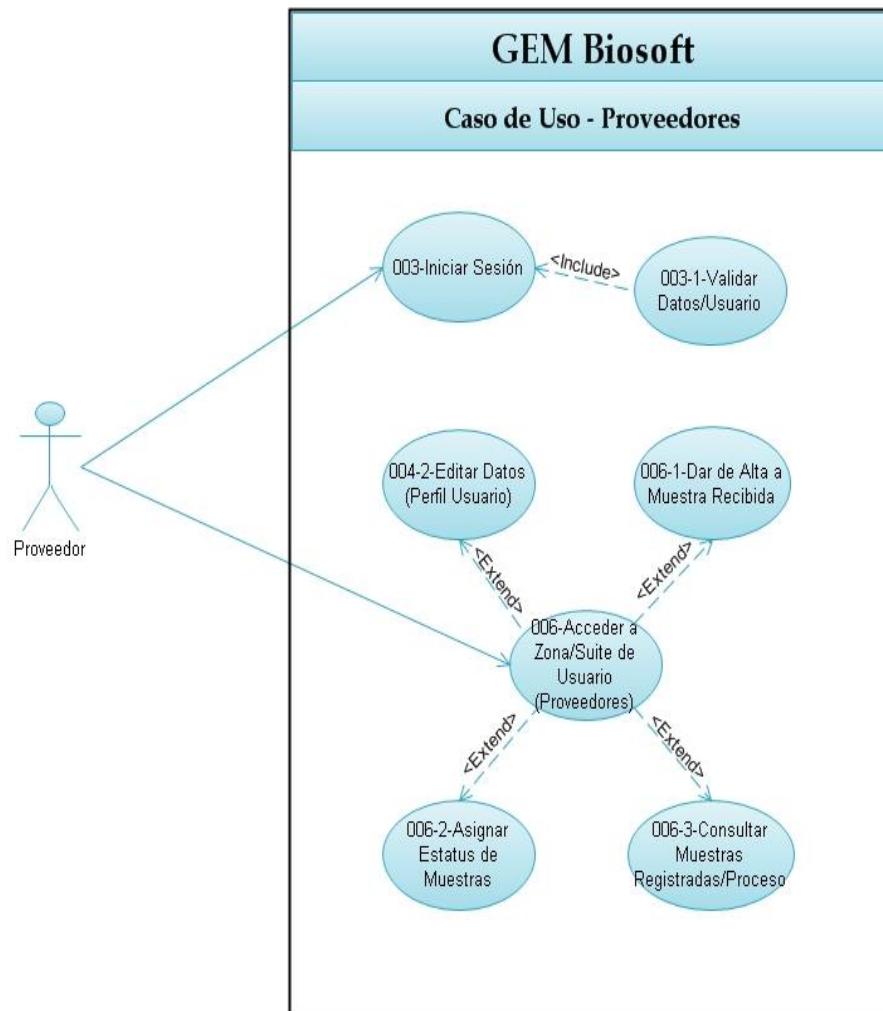


Figura 3.10 Caso de Uso General: Proveedor

En lo adelante los siguientes Casos de Usos serán omitidos puesto que se definen en los puntos anteriores:

- 003- Iniciar Sesión
- 004- Acceder a Zona/Suite de Usuario
- 004-2 Editar Datos (Perfil Usuario)

Nombre CU	006-1 Dar de Alta a Muestra Recibida
Propósito	Reportar las muestras que llegan al proveedor, con el fin de llevar un seguimiento del estado en que se encuentran
Precondiciones	Realizar el inicio de sesión
Postcondiciones	Estado de la muestra registrado/activado y notificado
Casos de Uso que incluye	-

**El caso de uso está formado por los pasos siguientes:**

1. El proveedor accede a su zona de usuario, donde busca el formulario para el registro/activación de muestras.
2. El proveedor inserta el código del colector de muestra, y si es correcto lo activa y notifica a los demás usuarios de que se encuentra en etapa de análisis.
3. Cuando el identificador es incorrecto/invalido le permite introducirlo dos veces más y si falla envía una notificación a la administración.

Nombre CU	006-2 Asignar Estatus de Muestras
Propósito	Notificar sobre el estado en que se encuentra la muestra. Con esta facilidad podemos detectar si la muestra sigue en análisis o si ya está en etapa de finalización
Precondiciones	Realizar el inicio de sesión
Postcondiciones	Estado de la muestra modificado
Casos de Uso que incluye	-

**El caso de uso está formado por los pasos siguientes:**

1. El proveedor accede a su zona de usuario, y ve las muestras que está trabajando en ese momento.
2. El proveedor selecciona una muestra en específica y le asigna un nuevo estado a la muestra.
3. Según la elección del proveedor la muestra mostrará en que etapa del proceso se encuentra, como por ejemplo: en análisis, revisión, envío, etc.
4. El sistema actualiza el estatus y notifica a los demás usuarios sobre la situación actual de la muestra.

Nombre CU	006-3 Consultar Muestras Registradas/Proceso
Propósito	Visualizar todas las muestras que se tienen asignadas
Precondiciones	Realizar el inicio de sesión
Postcondiciones	Información sobre las muestras “obtenida/revisada”
Casos de Uso que incluye	-

**El caso de uso está formado por los pasos siguientes:**

1. El proveedor cuando decide revisar su lista de muestras, se dirige al área de consultas.
2. El proveedor en esta parte puede buscar todas las muestras o algunas en específicas para ver en qué estado se encuentran (pendientes, en proceso, finalizadas, etc.).
3. Al finalizar la consulta el proveedor puede generar informes, realizar cambios u otros.

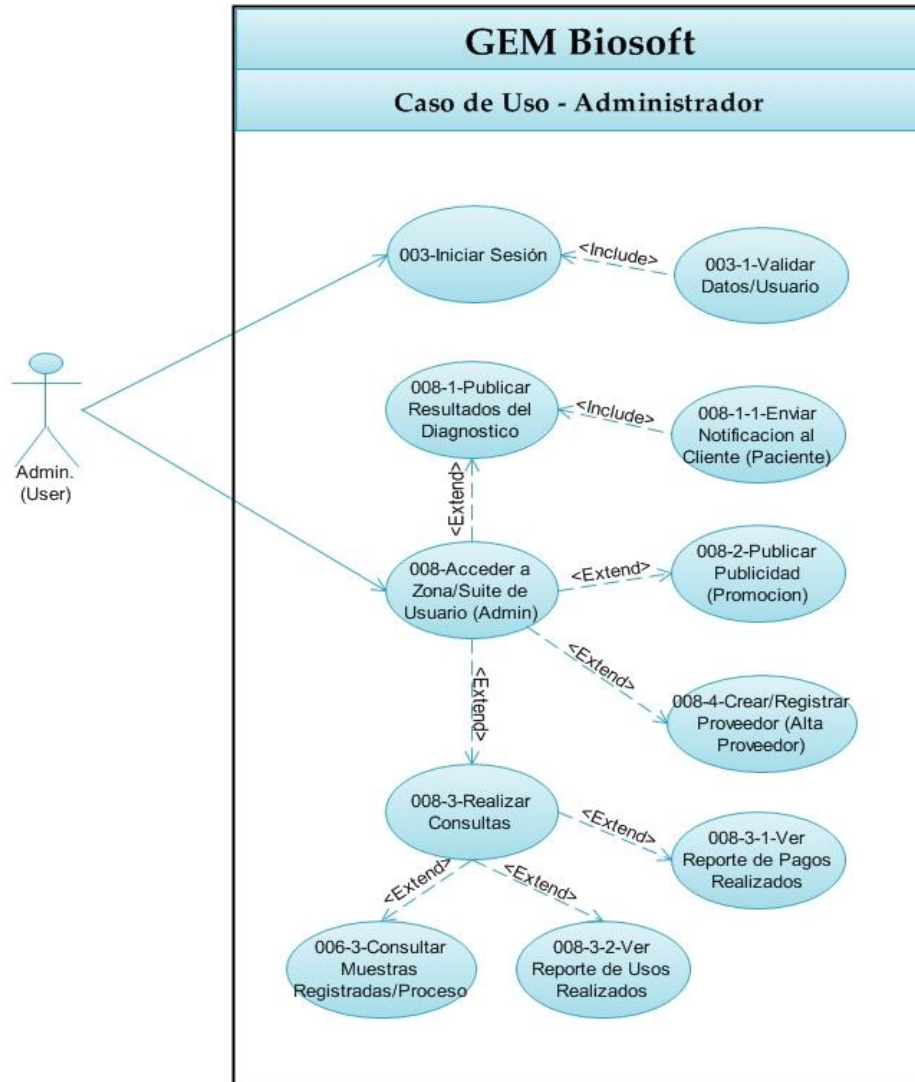


Figura 3.11 Caso de Uso General: Administrador

Nombre CU	008-1 Publicar Resultados del Diagnóstico
Propósito	Colocar en la web los resultados obtenidos luego de realizar el diagnóstico
Precondiciones	Realizar inicio de sesión como administrador
Postcondiciones	Resultado del diagnóstico publicado
Casos de Uso que incluye	-

**El caso de uso está formado por los pasos siguientes:**

1. El administrador tan pronto recibe las secuencias de referencias procede a realizar el diagnóstico.
2. Al momento de finalizar el diagnóstico, el administrador debe colocar en la web los resultados obtenidos para que el usuario (Cliente-paciente) pueda consultarlos.
3. El administrador se dirige a su zona de usuario y selecciona la opción de Publicar Resultados, aquí se coloca el fichero del diagnóstico.

4. El sistema muestra los resultados disponibles en la web y notifica al usuario por vía email que puede acceder y ver su diagnóstico.

Nombre CU	008-2 Publicar Publicidad (Promoción)
Propósito	Agregar en la web promociones sobre nuevos servicios
Precondiciones	Realizar el inicio de sesión
Postcondiciones	Creación de borrador o publicidad publicada en la web
Casos de Uso que incluye	-

**El caso de uso está formado por los pasos siguientes:**

1. El administrador puede agregar en la web publicidad para llevar a cabo la promoción de distintos productos, ofertas, nuevos servicios ofertados y demás.
2. El administrador selecciona la opción de Publicidad (publicar) y aquí coloca los contenidos e imágenes que conformarán el anuncio.
3. Cuando se tienen todos los datos, el administrador puede decidir si guardar la publicidad como borrador (con fines de revisiones y/o retoques) o publicarla directamente en la web.
4. El sistema publica en la web la nueva publicidad.

Nombre CU	008-3 Realizar Consultas
Propósito	Realizar búsquedas de información
Precondiciones	Realizar el inicio de sesión
Postcondiciones	
Casos de Uso que incluye	006-3 Consultar Muestras Registradas/Proceso; 008-3-1 Ver Reporte de Pagos Realizados; 008-3-2 Ver Reporte de Usos Realizados

**El caso de uso está formado por los pasos siguientes:**

1. El administrador accede a su área trabajo y elige la opción de consultas.
2. El administrador selecciona que consulta quiere hacer: 1) Muestras registradas/proceso; 2) Pagos realizados; 3) Usos realizados, los fines de estas opciones son los siguientes:
  - a. Consultar Muestras Registradas/Proceso: listado de muestras que han sido activadas por los usuarios (clientes/proveedores), mostrando el estado en que se encuentran en ese momento.
  - b. Ver Reporte de Pagos Realizados: aquí podemos acceder a la lista de todos los pagos que se han recibido por la compra de servicios ofertados. (Aquí tenemos información sobre los pagos realizados de forma exitosa o errada/abortada).
  - c. Ver Reporte de Usos Realizados: en este vemos todos los movimientos realizados, por medio de este reporte podemos tener un control de todos los cambios hechos en distintos periodos de tiempo (por día, semana, mes, etc.).
3. El administrador puede realizar búsquedas generales o específicas, como por igual generar informes e imprimir los datos encontrados.

Nombre CU	008-4 Crear/Registrar Proveedor (Alta Proveedor)
Propósito	Dar de alta a un nuevo proveedor
Precondiciones	Realizar el inicio de sesión como administrador
Postcondiciones	Proveedor creado (incluye datos de acceso a la web)
Casos de Uso que incluye	-

**El caso de uso está formado por los pasos siguientes:**

1. El administrador accede a su zona y selecciona la opción de Proveedores, al estar aquí accede a Crear Proveedor.
2. En el formulario de creación de proveedores procede a rellenar todos los campos, y luego guardarlos (si algún campo está incorrecto/incompleto lo señalará).
3. Al finalizar, el administrador guarda los datos y el sistema genera un email automático para el proveedor, en el cual le indicará los pasos que debe realizar para la activación de su cuenta en la web.

## Diagrama de clases

A continuación se presenta el diagrama de clases asociado al análisis realizado, en este definimos los usuarios (*administrador, proveedor, clientes, anónimos*) del sistema, así como el tratamiento de: enfermedades, diagnóstico, servicios, pagos, entre otros.

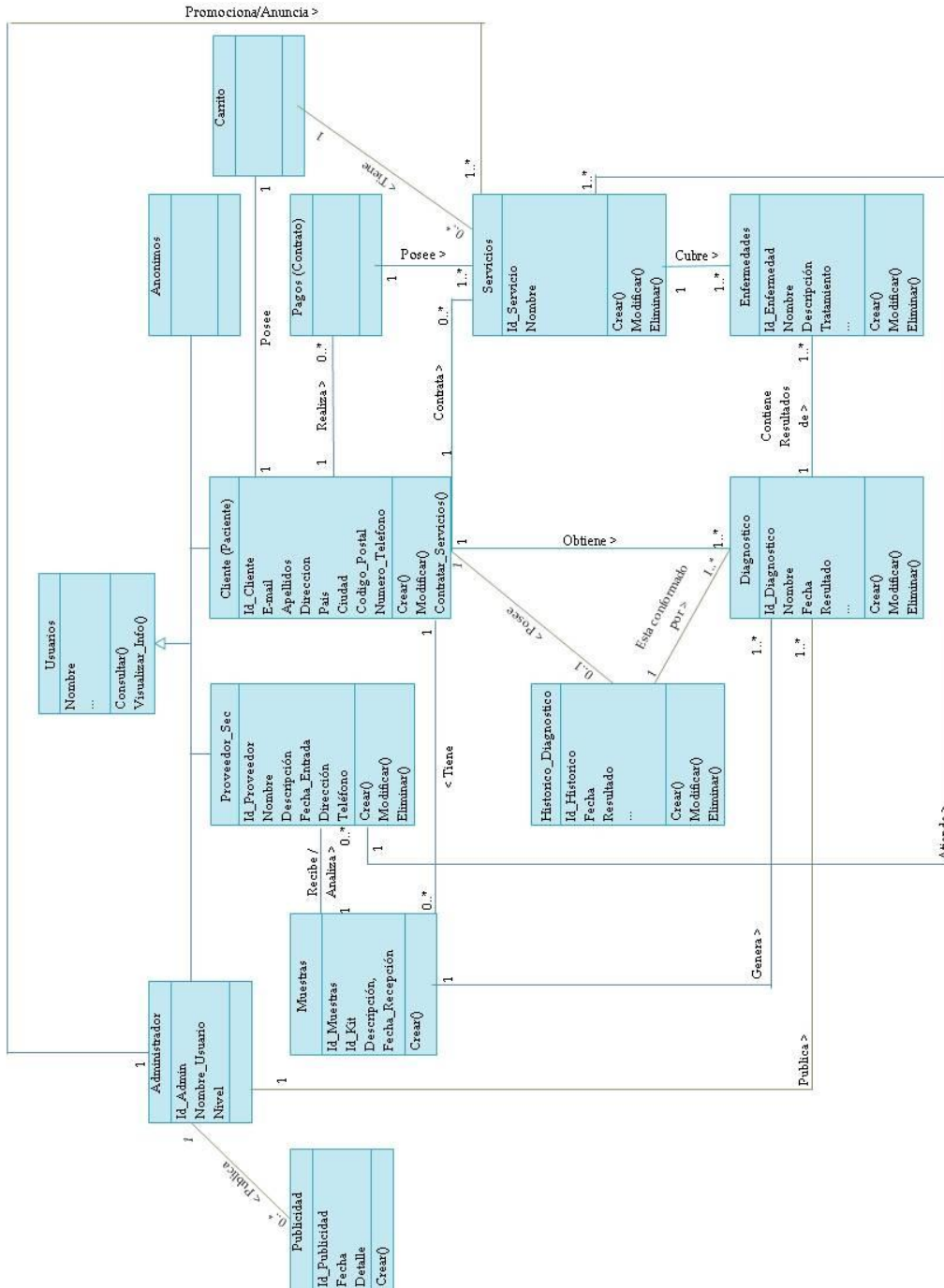


Figura 3.12 Diagrama de Clases de los Servicios Web



## Capítulo 4

# Implementación de SILE para el Diagnóstico de la Sensibilidad al Alcohol

### 4.1 Metodología SILE aplicada a la Sensibilidad al Alcohol

La metodología Search-Identification-Load-Exploitation (SILE) fue desarrollada por el grupo de investigación ProS (Centro de Investigación en Métodos de Producción Software) de la Universidad Politécnica de Valencia (UPV), con el objetivo de mejorar sus procesos para la carga de genes y variaciones de su base de datos del genoma humano.

Sus iniciales se definen como sigue:

<b>S</b>	Búsqueda - Consiste en la búsqueda exhaustiva de información científica (publicaciones, artículos, etc.) que apoyen la asociación genética con una enfermedad específica.
<b>I</b>	Identificación - es el proceso en el que interviene la parte médica para proporcionar apoyo en el filtrado de los genes y variaciones que tienen una mayor incidencia en la población.
<b>L</b>	Carga - El proceso de carga de la base de datos, es donde se procede a insertar los genes, cromosomas y variaciones a la base de datos con la información tratada (filtrada) y validada, la carga se lleva a cabo a través de diversas tareas y fuentes de datos.
<b>E</b>	Explotación - La explotación se basa en el contenido y presentación del resultado final.

La metodología SILE se lleva a cabo por medio de diversas bases de datos que facilitan la búsqueda y consulta de información biomédica, para las tareas de lugar utilizaremos NCBI. El Centro Nacional de Información Biotecnológica (NCBI) [1] es una base de datos (biblioteca) que recopila información sobre la biomedicina, la biotecnología, la bioquímica, la genética, la genómica, enfermedades genéticas, entre otro más.

NCBI también proporciona algunas herramientas bioinformáticas para el análisis de secuencia de ADN, ARN y proteínas; un ejemplo es BLAST [25], que es una herramienta para la alineación de secuencias de mayor uso actualmente. NCBI fue establecido el 4/noviembre de 1988, con el fin de mantener y distribuir de manera rápida toda la información sobre la biología molecular [1].

## Sensibilidad al Alcohol

Esta enfermedad se manifiesta cuando un individuo ingiere cierta cantidad de alcohol, provocando en el organismo un rechazo inmediato, manifestando malestares, mareos u otros. Por lo que el simple hecho de consumirlo es el causante de molestias y repercusiones futuras en la salud del individuo [22] [26]. Dada la importancia que recae sobre enfermedad la hemos tomado como base para poner en marcha nuestros experimentos, teniendo en cuenta que esta puede afectar de manera directa a la población en general, sin importar estatus sociales, edades o cultura.

### 4.1.1 Search

La etapa inicial de la Metodología SILE se basa en la búsqueda de información asociada a una enfermedad en específica. SILE trabaja de la mano de los avances médicos (comunidad científica) y en este caso hemos realizado una serie de estudios para la Sensibilidad al Alcohol.

En este proceso de búsqueda se ha podido tomar como referencia un grupo de genes que están sumamente ligados con esta enfermedad. Y entre ellos se encuentran:

GEN	OFFICIAL FULL NAME
ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide
ALDH2	aldehyde dehydrogenase 2 family (mitochondrial)
GABRA2	gamma-aminobutyric acid (GABA) A receptor, alpha 2
PECR	peroxisomal trans-2-enoyl-CoA reductase
PKNOX2	PBX/knotted 1 homeobox 2
SLC22A18	solute carrier family 22, member 18
DRD2	DRD2 dopamine receptor D2

Estos genes están relacionados de manera directa o indirecta con la Sensibilidad al Alcohol, los directos influyen de una manera elevada en la predisposición de esta enfermedad y en el caso de los indirectos estos influyen de manera general sobre la misma, ya que algunos son asociados a enfermedades adictivas.

Nuestros estudios se sustentan en la investigación exhaustiva de artículos y publicaciones que validen la asociación de estos genes, tomando referencias de grandes bibliotecas de datos biomédicos como es el caso de NCBI [1] y sus herramientas dbSNP [20] [21], PubMed [27] [28], entre otros.

Nuestro objetivo fundamental en esta tarea es recolectar toda la información posible sobre la enfermedad que estudiamos, para así acceder al siguiente paso de identificación.

#### 4.1.1.1 Haplotipos

En la búsqueda de información encontramos el caso de los haplotipos para un gen, estos repercuten de una manera distinta al momento de manejar las variaciones y de generar el diagnóstico. Un haplotipo es un conjunto de variaciones del ADN, o polimorfismos, que tienden a ser heredados juntos. Esto se puede referir a una combinación de alelos o para un conjunto de polimorfismos de nucleótido sencillo (SNPs) que se encuentran en un mismo cromosoma [29].

Hemos estudiado este caso particular, y vemos la necesidad de utilizarlo puesto que garantizará resultados más eficientes y fiables para los usuarios.

#### 4.1.2 Identification

Nuestro siguiente paso en la implementación de SILE es la Identificación, aquí realizamos un filtrado (depuración) de los genes con sus respectivos SNP's o variaciones, haciéndolo de la mano de médicos y/o especialistas del área que certifiquen/validen los genes de referencia seleccionados.

En esta etapa optamos por seleccionar los genes y variaciones de mayor incidencia, ya que en el proceso previo extraemos todo lo que esté asociado a la enfermedad en estudio, por lo que debemos clasificar muy bien los genes que han sido probados científicamente de los que aun se encuentran en periodo de pruebas. Entonces, basándonos en las referencias tomadas y el conocimiento médico realizamos esta validación.

En la tabla 4.1 podemos visualizar todos los genes y variaciones encontrados en la etapa inicial de SILE, vemos también sus nombres oficiales y sinónimos.

En la tabla 4.2 encontramos la tabla anterior luego de pasar por el proceso de identificación. Aquí hemos delimitado los genes y variaciones, y nos hemos quedado con los que poseen una mayor incidencia en la población. Los SNP/variaciones en rojo fueron descartados para el proceso de carga, por no poseer un peso representativo o validez oficial en el campo médico; los amarillos son los identificados como Haplotipos; los verdes son los que pasaron el filtrado y se tienen contemplados para el proceso siguiente de SILE (carga de la base de datos).

GEN	Official Full Name	Also known as...	SNP/Variations (opensnp.org)			
ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide	ADH1B; ADH2	rs671	rs1229982	rs1229984	rs1230025
ALDH2	aldehyde dehydrogenase 2 family (mitochondrial)	ALDM; ALDHI; ALDH-E2	rs671	rs7590720	rs1800497	-
GABRA2	gamma-aminobutyric acid (GABA) A receptor, alpha 2	-	rs279836	rs279858	rs279871	-
PECR	peroxisomal trans-2-enoyl-CoA reductase	TERP; DCRRP; PVIARL; HPDHASE; SDR29C1; HSA250303	rs7590720	-	-	-
PKNOX2	PBX/knotted 1 homeobox 2	PREP2	rs1426153	rs750338	rs585977	rs10893366
SLC22A18	solute carrier family 22, member 18	HET; ITM; BWR1A; IMPT1; TSSC5; ORCTL2; BWSCR1A; SLC22A1L; p45-BWR1A	rs16928809	-	-	-
DRD2	DRD2 dopamine receptor D2	D2R; D2DR	rs1076560	rs1800497	rs6276	-

Tabla 4.1 Identificación de Genes & SNP's

GEN	Official Full Name	Also known as...	SNP/Variations (opensnp.org)			
ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide	ADH1B; ADH2	rs671	rs1229982	rs1229984	rs1230025
ALDH2	aldehyde dehydrogenase 2 family (mitochondrial)	ALDM; ALDHI; ALDH-E2	rs671	rs7590720	rs1800497	-
GABRA2	gamma-aminobutyric acid (GABA) A receptor, alpha 2	-	rs279836	rs279858	rs279871	-

Tabla 4.2 Filtrado de Genes & SNP's



Summary NCBI SNP's in GenesCard							
Gen	GeneLoc location for						SNP
	chromosome	Start	End	Size	Orientation	RefSeq DNA sequence	
ADHIB	4	100,226,121 bp from pter	100,242,572 bp from pter	16,452 bases	minus strand	NC_000004.11 NT_016354.19	rs671
							rs1229982
							rs1229984
							rs1230025
ALDH2	12	112,204,346 bp from pter	112,247,784 bp from pter	43,439 bases	plus strand	NC_000012.11 NT_009775.17	rs671
							rs7590720
							rs1800497
GABRA2	4	46,250,444 bp from pter	46,477,247 bp from pter	226,804 bases	minus strand	NC_000004.11 NT_006238.11	rs279836
							rs279858
							rs279871

Tabla 4.4 Resumen GeneCards - Localización en Genes: inicio-final, tamaño, orientación, etc.

Summary NCBI SNP's in GenesCard											
Gen	SNP	Population								# Total Samples	
		Not Specified (NS)	East Asia (EA)	Central Asia (CA)	North America (NA)	Central /South Africa (CSA)	West Africa (WA)	Multi-National (MN)	Europe (EU)		
ADHIB	rs671	-	-	-	-	-	-	-	-	0	11,857
	rs1229982	Yes	Yes	-	Yes	Yes	Yes	-	-	3,379	
	rs1229984	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	8,478	
	rs1230025	-	-	-	-	-	-	-	-	0	
ALDH2	rs671	Yes	Yes	Yes	Yes	-	-	Yes	-	2,426	2,426
	rs7590720	-	-	-	-	-	-	-	-	0	
	rs1800497	-	-	-	-	-	-	-	-	0	
GABRA2	rs279836	Yes	Yes	-	Yes	Yes	Yes	-	-	770	11,939
	rs279858	Yes	Yes	-	Yes	-	Yes	Yes	Yes	9,675	
	rs279871	-	Yes	-	Yes	Yes	Yes	-	-	1,494	

Tabla 4.5 Resumen GeneCards II - Asociación de SNP's con Poblaciones (Estudiadas)

GeneCards es una base de datos integrada, de los genes humanos que proporciona información concisa relacionada con la genómica [36].

En las tablas 4.4 y 4.5 podemos ver toda la información de los genes detectados con sus respectivos SNP's o variaciones. De esta forma conocemos la ubicación del gen, así como su tamaño, orientación, entre otros. También se presenta las muestras analizadas dentro de la comunidad científica para las distintas poblaciones. Con estas informaciones vemos factores que nos ayudan en el proceso de identificación.

Cuando finalizamos la etapa de identificación debemos de ser conscientes de la calidad y validez de los datos tratados, puesto que es en esta etapa donde definimos los parámetros para la futura generación de los diagnósticos (resultados).

#### 4.1.2.1 Haplotipo detectado en la Sensibilidad al Alcohol

Con nuestro caso de estudio de la Sensibilidad al Alcohol, realizamos una búsqueda intensiva de genes y variaciones que estuvieran relacionadas para determinar la predisposición de padecer esta enfermedad.

Dentro del conjunto de genes y variaciones detectamos un caso especial, el cual era sobre el tratamiento de los haplotipos. Como sabemos los haplotipos son un conjunto de SNP's que se heredan y encuentran juntos en un cromosoma. Tomar en cuenta los casos de haplotipos permite mejorar circunstancialmente los resultados, ya que las probabilidades de certeza aumentarían en grandes porcentajes.

En el gen GABRA2 encontramos el haplotipo formado por la variación rs279871, rs279836 y rs279845; donde anteriormente no considerábamos la variación rs279845 [22] [23].

El rs279871 forma un haplotipo junto con rs279836 y rs279845			
<b>rs279871</b>	RefSNP Alleles:	A/G	Ancestral Allele: A
<b>rs279836</b>	RefSNP Alleles:	A/T	Ancestral Allele: T
<b>rs279845</b>	RefSNP Alleles:	A/T	Ancestral Allele: A

Información sobre la variación que no teníamos en cuenta:

**rs279845:** (SNV: single nucleotide variation)  
 RefSNP Alleles: A/T  
 Ancestral Allele: A  
 HGVS Names:  
 - NC\_000004.11:g.46329723T>A  
 - NG\_012835.1:g.67334A>T  
 - NM\_000807.2:c.255+4909A>T  
 - NM\_001114175.1:c.255+4909A>T  
 - NT\_006238.11:g.6032627T>A

En lo adelante veremos la solución a esta problemática, puesto que debemos manejar los haplotipos de manera distinta que las variaciones, porque en este caso es el conjunto de alelos que caracterizan el riesgo. Cada uno por separado no es suficiente, por ello se deben tratar como haplotipos.

#### 4.1.2.2 Estudio HapMap (Frequencies)

Con el objetivo de conocer más sobre las variaciones detectadas para la Sensibilidad al Alcohol, se realizó una investigación/estudio en HapMap, con fines de explotar este proyecto

que es de gran repercusión en el estudio de este tipo de variaciones. Con esto pudimos observar las distintas frecuencias de cada variación para distintas poblaciones.

El proyecto HapMap consiste en un multi-país con la colaboración de científicos y organismos de financiación de: Japón, el Reino Unido, Canadá, China, Nigeria y los Estados Unidos [19].

Las poblaciones estudiadas por el Proyecto HapMap son las siguientes [19] [24]:

- ASW: African ancestry in Southwest USA
- CEU: Utah residents with Northern and Western European ancestry from the CEPH collection
- CHB: Han Chinese in Beijing, China
- CHD: Chinese in Metropolitan Denver, Colorado
- GIH: Gujarati Indians in Houston, Texas
- JPT: Japanese in Tokyo, Japan
- LWK: Luhya in Webuye, Kenya
- MXL: Mexican ancestry in Los Angeles, California
- MKK: Maasai in Kinyawa, Kenya
- TSI: Toscani in Italia
- YRI: Yoruba in Ibadan, Nigeria

A continuación se presentará el resumen de los datos obtenidos en HapMap luego de realizar la descarga de los ficheros de cada variación (en su respectivo cromosoma); en el proceso de extracción de estos datos identificamos cuales variaciones han sido más tratadas y cuales aun no poseen un alto nivel de pruebas pero son consideradas porque tienen cierto nivel de probabilidad dentro de la población.

Con los estudios que se obtuvieron se nos facilitó el entendimiento de ciertos factores de incidencia en las variaciones o SNP's. Además de recopilar información que nos permitió profundizar aun más y ver cuáles son las proyecciones futuras con el estudio de los haplotipos.

GENE	Population	rs#	chrom	pos	strand	refallele	refallele_ freq	refallele_ count	otherallele	otherallele_ freq	otherallele_ count	totalcount	refallele_freq + otherallele_freq
ALDH2	CEU	rs671	chr12	110726149	+	G	1.000	130	A	0	0	130	1
ALDH2	CHB	rs671	chr12	110726149	+	G	0.838	228	A	0.162	44	272	1
ALDH2	CHD	rs671	chr12	110726149	+	G	0.739	161	A	0.261	57	218	1
ALDH2	JPT	rs671	chr12	110726149	+	G	0.779	176	A	0.221	50	226	1
ALDH2	LWK	rs671	chr12	110726149	+	G	0.995	219	A	0.005	1	220	1
ALDH2	MEX	rs671	chr12	110726149	+	G	0.983	114	A	0.017	2	116	1
ALDH2	YRI	rs671	chr12	110726149	+	G	1.000	124	A	0	0	124	1
ADH1B	ASW	rs1229982	chr4	100462955	+	T	0.482	55	G	0.518	59	114	1
ADH1B	CEU	rs1229982	chr4	100462955	+	T	0.190	43	G	0.810	183	226	1
ADH1B	CHB	rs1229982	chr4	100462955	+	T	0.040	11	G	0.960	263	274	1
ADH1B	CHD	rs1229982	chr4	100462955	+	T	0.069	15	G	0.931	203	218	1
ADH1B	GIH	rs1229982	chr4	100462955	+	T	0.139	28	G	0.861	174	202	1
ADH1B	JPT	rs1229982	chr4	100462955	+	T	0.053	12	G	0.947	214	226	1
ADH1B	LWK	rs1229982	chr4	100462955	+	T	0.464	102	G	0.536	118	220	1
ADH1B	MEX	rs1229982	chr4	100462955	+	T	0.181	21	G	0.819	95	116	1
ADH1B	MKK	rs1229982	chr4	100462955	+	T	0.403	125	G	0.597	185	310	1
ADH1B	TSI	rs1229982	chr4	100462955	+	T	0.186	38	G	0.814	166	204	1
ADH1B	YRI	rs1229982	chr4	100462955	+	T	0.561	165	G	0.439	129	294	1
ADH1B	CHB	rs1229984	chr4	100458342	+	T	0.767	69	C	0.233	21	90	1
ADH1B	JPT	rs1229984	chr4	100458342	+	T	0.739	65	C	0.261	23	88	1
GABRA2	CEU	rs279836	chr4	46033827	+	T	0.508	64	A	0.492	62	126	1
GABRA2	YRI	rs279836	chr4	46033827	+	T	0.842	101	A	0.158	19	120	1
GABRA2	ASW	rs279845	chr4	46024480	+	T	0.737	84	A	0.263	30	114	1
GABRA2	CEU	rs279845	chr4	46024480	+	T	0.527	119	A	0.473	107	226	1
GABRA2	CHB	rs279845	chr4	46024480	+	T	0.485	133	A	0.515	141	274	1
GABRA2	CHD	rs279845	chr4	46024480	+	T	0.519	112	A	0.481	104	216	1
GABRA2	GIH	rs279845	chr4	46024480	+	T	0.609	123	A	0.391	79	202	1
GABRA2	JPT	rs279845	chr4	46024480	+	T	0.540	122	A	0.460	104	226	1
GABRA2	LWK	rs279845	chr4	46024480	+	T	0.714	157	A	0.286	63	220	1
GABRA2	MEX	rs279845	chr4	46024480	+	T	0.466	54	A	0.534	62	116	1
GABRA2	MKK	rs279845	chr4	46024480	+	T	0.615	192	A	0.385	120	312	1
GABRA2	TSI	rs279845	chr4	46024480	+	T	0.644	130	A	0.356	72	202	1
GABRA2	YRI	rs279845	chr4	46024480	+	T	0.752	221	A	0.248	73	294	1

Tabla 4.6 Estudio de Frecuencias en HapMap para las Variaciones de la Sensibilidad al Alcohol.



Los datos mostrados en la tabla anterior corresponden a lo siguiente:

- Gene: se refiere al gen que estamos buscando.
- Population: población para la cual se aplica el estudio.
- Rs#: se refiere a la variación dentro del gen.
- Chrom: cromosoma donde se encuentra el gen.
- Pos: posición de la variación dentro del gen.
- Strand: esto indica en que hebra se encuentra la variación, esta puede ser hebra positiva o negativa.
- Refallele: representa el alelo de referencia.
- Refallele\_freq: indica la frecuencia del alelo de referencia.
- Refallele\_count: indica el conteo dentro de la población del alelo de referencia.
- Otherallele: estudios detectados con otros alelos, cambio en el alelo.
- Otherallele\_freq: indica la frecuencia del otro alelo encontrado.
- Otherallele\_count: conteo del otro alelo encontrado dentro una población.
- Totalcount: número total de estudios realizados en la población.
- Refallele\_freq + otherallele\_freq: es la sumatoria de las frecuencias de los dos alelos estudiados (refallele\_freq y otherallele\_fre).

### 4.1.3 Load

El proceso de carga es el encargado de introducir la información identificada en los procesos anteriores a la base de datos del genoma humano. Este proceso está subdividido en 3 subprocesos: extracción, transformación y carga [30].

- ✓ Extracción: consiste en obtener la información de las variaciones de un repositorio externo en un formato que se pueda almacenar en un servidor, para luego poder manejarlo localmente.
- ✓ Transformación: luego de la extracción selectiva de la información de los distintos repositorios, se debe tratar la información obtenida para adecuarla a la estructura de la base de datos del genoma humano.
- ✓ Carga: es el último subproceso del proceso completo de carga. En este se introduce la información ya transformada desde un fichero local a la base de datos.

Para asegurar que la base de datos posea sólo datos de calidad, se paso de la carga masiva a la selectiva. Tomando como referencia la calidad de alineamiento de la secuencia, tomando solo aquellas que contaban con un 100% en alineación [30].

#### 4.1.3.1 Modelo Conceptual del Genoma Humano (v3) e Integración de los Haplotipos

En la figura 4.7 se presenta el modelo conceptual del genoma humano utilizado actualmente (sección señalada en rojo fue la parte estudiada, revisada y modificada), y más adelante vemos la vista de la base de datos (figura 4.9). También indicaremos los cambios recomendados para el tratamiento/solución del caso de los haplotipos (figura 4.8).

El caso de los haplotipos no estaba tratado, pero al momento de trabajar con la Sensibilidad al Alcohol y encontrarlo, nos vimos en la necesidad de implementarlos de forma que garantizaran diagnósticos correctos y fiables.

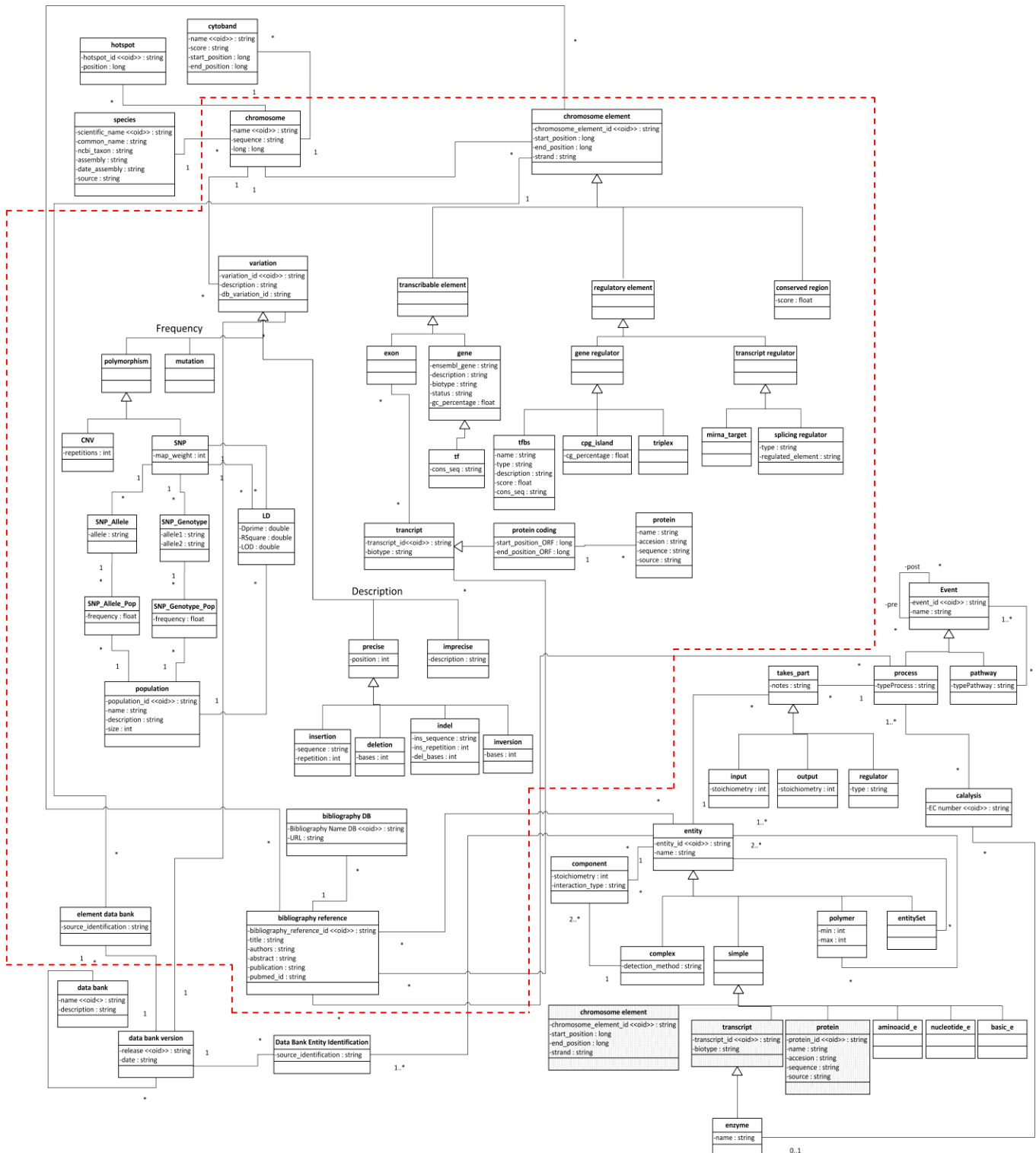


Figura 4.7 Modelo Conceptual del Genoma Humano

El modelo conceptual del genoma humano por el grupo de investigación ProS surge en el año 2008, donde se inicia el desarrollo y evolución del genoma, esto con fines de potenciar los Sistemas de Información en el área de la bioinformática [31].

Muchas han sido las versiones generadas, y con el paso del tiempo y la ayuda de personas especializadas se lograron crear versiones posteriores más eficientes, mejorando principalmente en los aspectos de mutaciones y la relación genotipo-fenotipo [31].

En esta sección presentamos una mejora al modelo, el cual fue estudiado en varias sesiones y validado con expertos del grupo Genoma del ProS. El conocimiento adquirido con el paso de los años permitirá cada vez más buscar puntos de mejoras y crecimiento para el modelo.

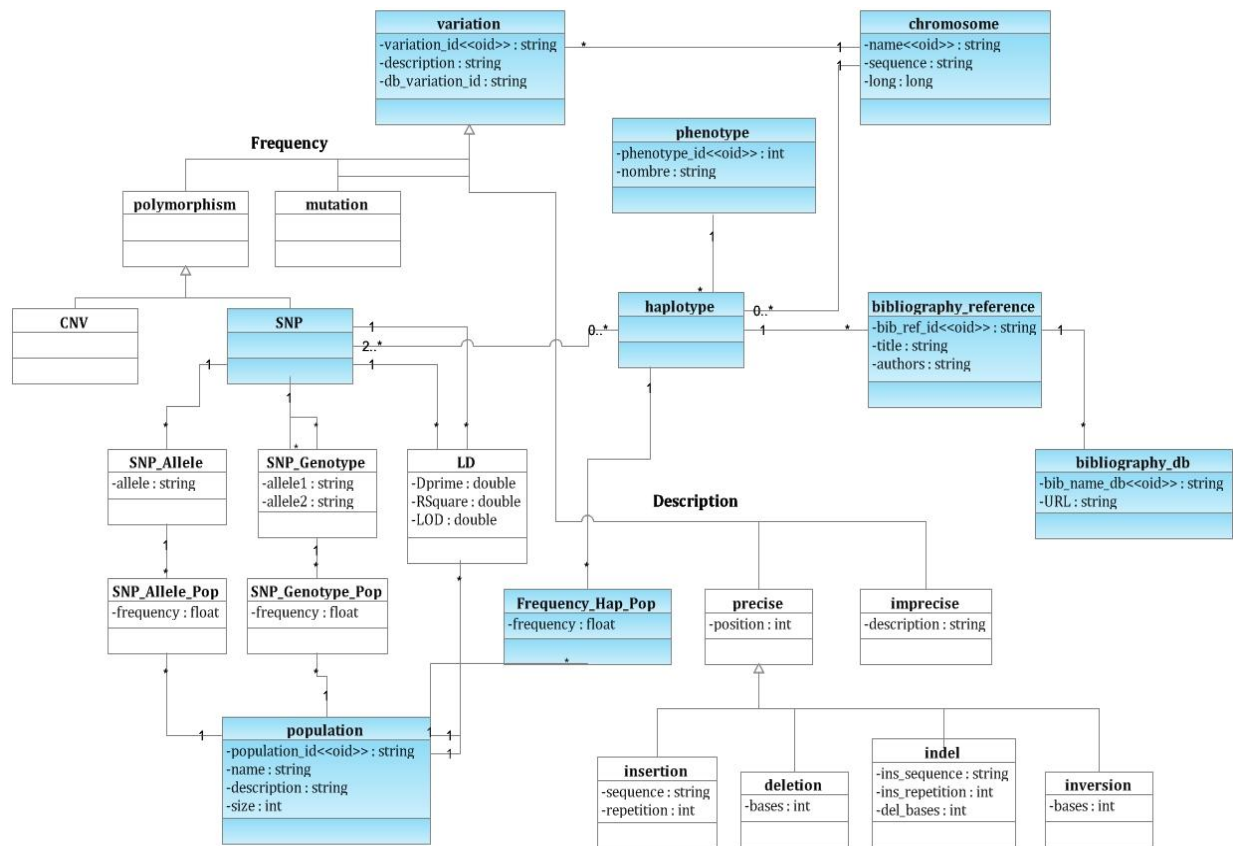


Figura 4.8 Modificación al Modelo Conceptual del Genoma Humano para la Integración de los Haplotipos

La figura 4.6 presenta la modificación realizada al modelo conceptual del genoma humano para realizar la integración de los haplotipos. Las clases sombreadas en azul son las que se ven afectadas de manera directa o indirecta con esta implementación, a continuación explicamos los detalles/cambios:

- Variation: esta clase posee todas las variaciones existentes en la cadena de ADN.
- Phenotype: representa los fenotipos asociados a una o varias variaciones del ADN.
- Chromosome: representa los elementos que forman la secuencia de ADN, los cromosomas.
- SNP: Polimorfismo de nucleótido simple, es una especialización de la clase variation.

- Haplotype: representa la asociación de dos o más SNP's que componen el haplotipo.
- Frequency\_Hap\_Pop: representa la frecuencia de los haplotipos dentro de una población.
- Population: contiene las poblaciones y tamaños de pruebas realizadas para ciertas variaciones.
- Bibliography\_reference: contiene las referencias bibliográficas de la validación médica/científica de los haplotipos encontrados.
- Bibliography\_db: contiene los enlaces y datos de la fuente de información de los haplotipos.

Relaciones:

- a) Un **Cromosoma** puede contener muchas **Variaciones**.
- b) Un **Cromosoma** puede tener de cero a muchos **Haplotipos**.
- c) Un **Fenotipo** puede tener asociado muchos **Haplotipos**.
- d) Dos o más **SNP** pueden conformar cero o muchos **Haplotipos**.
- e) Un **Haplotipo** puede tener muchas **Frecuencias** dentro de la población
- f) Una **Población** puede tener muchas **Frecuencias** de haplotipos.
- g) Un **Haplotipo** puede contener muchas **Referencias\_Bibliograficas**.
- h) Una **Referencia\_Bibliografica** puede tener muchas **db\_Bibliograficas**.

Estas relaciones y restricciones son las que permiten el manejo correcto de las informaciones a cargar en la base de datos. Debemos estar claros de los alcances de este modelo, y entender por ejemplo casos como que: un haplotipo siempre se encuentra en un mismo cromosoma, un haplotipo podría afectar a varias enfermedades, entre otros.

La integración de los haplotipos permite que la base de datos contenga información más especializada y de gran valor. En el caso de estudio de la Sensibilidad al Alcohol se hace obligatoria esta observación, porque hay casos en que los SNP's tratados de manera independiente no nos ayudarían a contar con un diagnóstico real.

Es por ello que realizamos esta asociación de SNP's o variaciones; esta unión es que garantizará una mayor fiabilidad en el diagnóstico final.

#### 4.1.3.2 Carga de Base de Datos

En el proceso de carga de la base de datos procedemos a introducir todas las informaciones que fueron encontradas y tratadas para la enfermedad estudiada.

Luego de aplicar los pasos previos a la carga, procedemos a insertar todas las informaciones en la base de datos. La figura 4.7 presenta una vista de la base de datos v3 [31]:

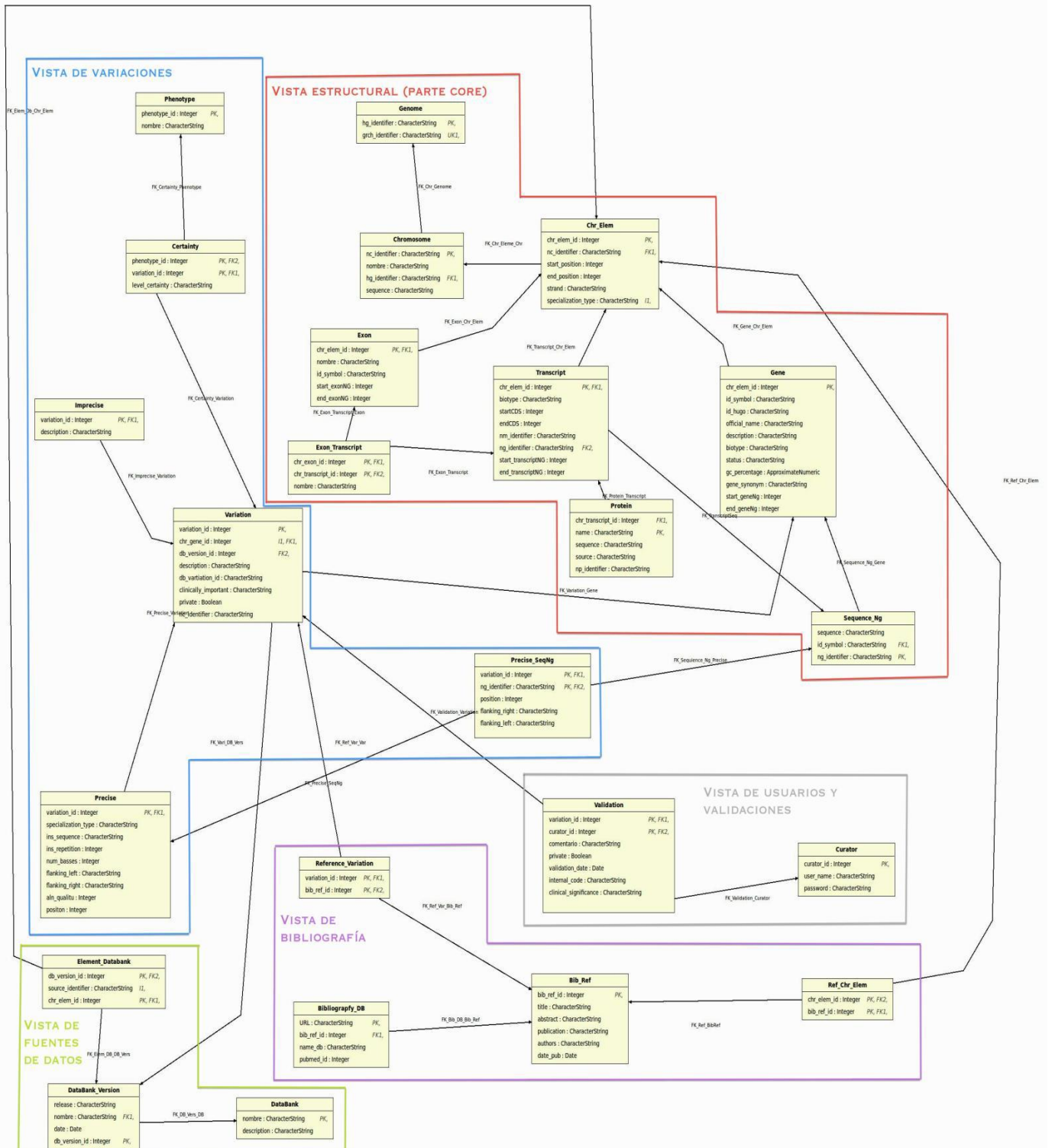


Figura 4.9 Vista de la Base de Datos v3

- **Vista Estructural:** modela la estructura del genoma de los organismos.
- **Vista de Variaciones:** esta vista representa el conocimiento relacionado con las diferencias encontradas en la secuencia de ADN de diversos individuos.
- **Vista de Fuentes de Datos & Bibliografía:** esta vista proporciona información sobre las fuentes de datos de las que se ha extraído la información que se va a almacenar en el modelo, así como una serie de documentos bibliográficos de consulta para quien desee obtener más información con respecto a algún aspecto aquí definido.
- **Vista de Usuarios y Validaciones**

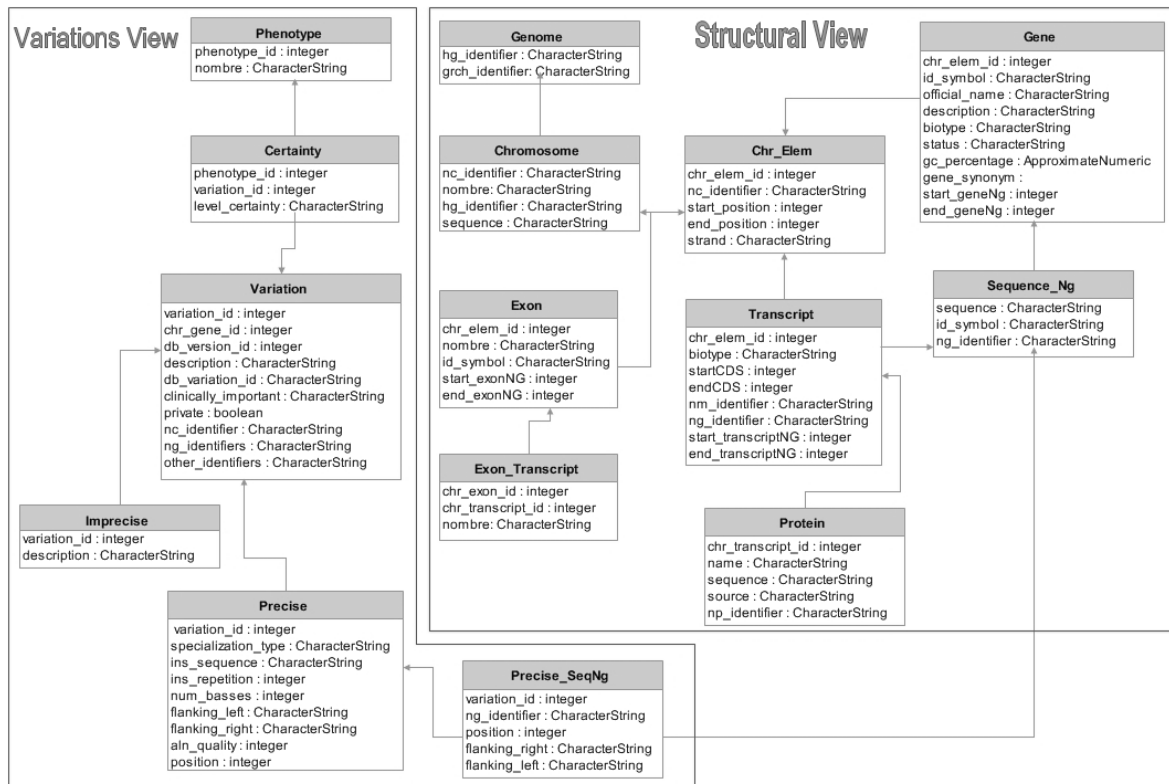


Figura 4.10 Vista de Variaciones & Estructural (Base de Datos v3)

La figura 4.8 nos muestra las vistas de variaciones y estructural, las mismas son de gran importancia porque en ellas se encuentra almacenada toda la información relacionada a las enfermedades, variaciones, cromosomas, entre otras.

#### 4.1.3.2.1 Carga Selectiva

Son muchos los beneficios que brinda la carga selectiva, y esto se debe a los procesos previos a la inserción en la base de datos.

Con la carga selectiva mejoramos la eficiencia en el momento de generar los resultados, obteniendo así diagnósticos más exactos. Esto se debe a que en la carga masiva tenemos una base de datos que contiene todas las variaciones sin tomar en cuenta que muchos de estos se encuentran en procesos de estudio o que por el momento no han sido validados por la comunidad médica. Por lo que tenemos una mayor imprecisión en los resultados.

La carga selectiva se enfoca solamente en las variaciones asociadas a una enfermedad y que las mismas hayan sido comprobadas por diversos estudios científicos. El tratamiento de los datos con la carga selectiva, nos permite crear bases de datos limpias y fiables. Porque este tipo de carga pasa por procesos que garantizan un filtrado de los datos, evitando almacenar datos irrelevantes y que no agregan ningún valor en la generación de los diagnósticos. La carga selectiva nos ayuda a mantener sistemas óptimos y de calidad.

### Consideración de Aspectos Poblacionales

La integración de los estudios poblacionales dentro de la base de datos sirve para la mejora de los resultados. El conocer sobre genes y variaciones que se desarrollan más en una población que en otra, genera grandes implicaciones dentro de los resultados finales.

Cuando manejamos variaciones validadas científicamente y le añadimos los aspectos poblacionales de mayor incidencia podemos generar diagnósticos más exactos, garantizando así un diagnóstico precoz y más fiable.

#### 4.1.4 Exploitation

La explotación es la última etapa de la metodología SILE, y su objetivo consiste en la generación y/o desarrollo de los contenidos desde el punto de vista de empresa y/o de la web.

Los contenidos pueden ser: texto, imágenes u otros. Estos deben ser claros y concisos, para que los usuarios puedan entender y manipular la información de forma efectiva.

Realizar una buena explotación, repercutirá de manera directa en el éxito o fracaso de los productos; desde que iniciamos con la búsqueda y seguimos los pasos de SILE podemos ir acumulando información muy valiosa que nos podrá servir de apoyo para la explotación.

##### 4.1.4.1 Diseño e Implementación del Servicio Web

A continuación veremos el resultado de la puesta en marcha de los servicios web, los diseños realizados y como fueron implementados. Podremos ver como se realiza el tratamiento de los usuarios y cuáles son las interfaces disponibles para acceder de manera fácil y sencilla a los test genéticos.





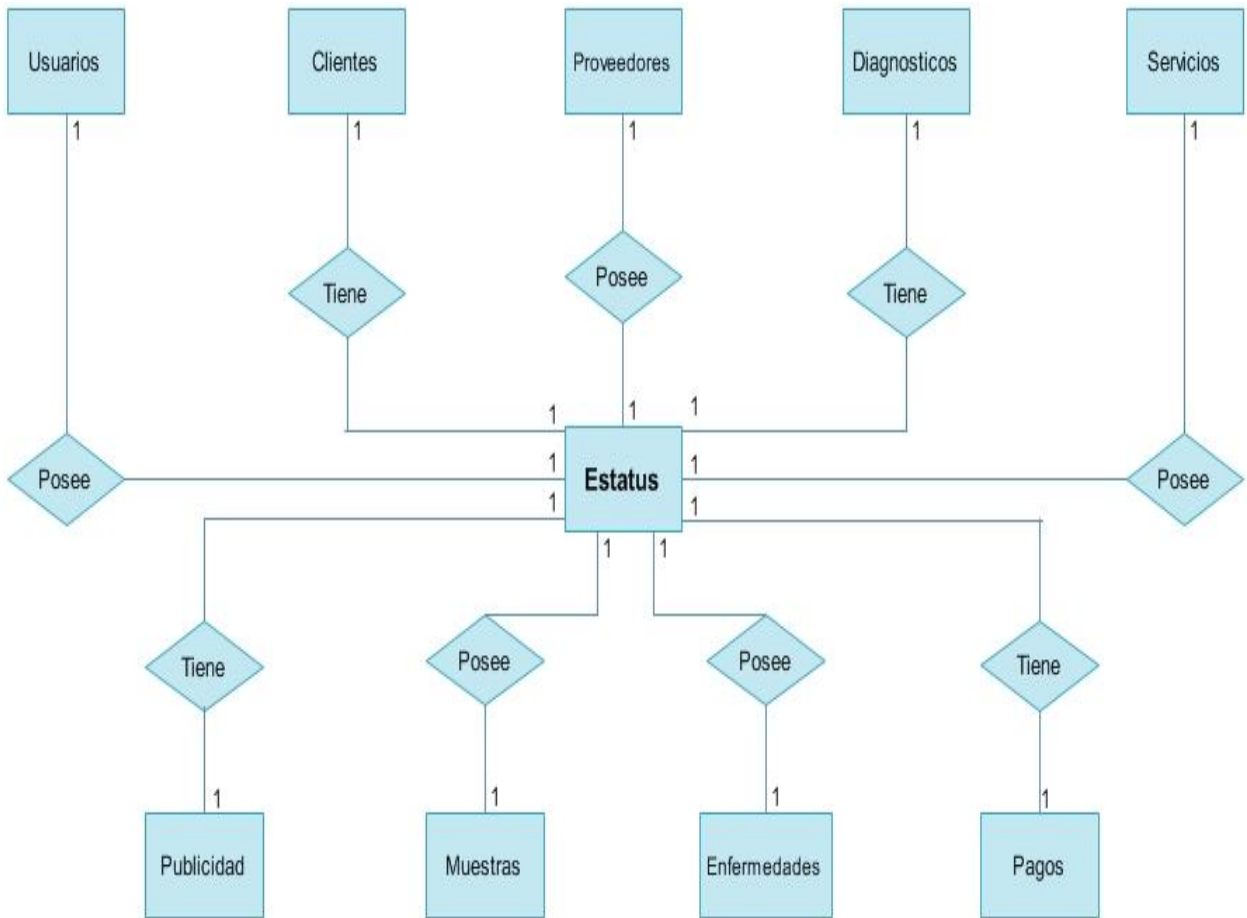


Figura 4.11 Diagrama Entidad-Relación (2)

## 4.1.4.1.2 Descripción de Tablas Normalizadas

USUARIOS	Tabla para registrar todos los usuarios del sistema, con sus respectivos Username y claves de acceso.
@Id_Usuario	{Numérico}
Nombre_User	1{Carácter}50
Rol	1{Carácter}50
Contraseña	1{Carácter}16
Confirmacion	1{Carácter}16
Id_Estatus	{Numérico}

ESTATUS	Registro de todos los posibles estatus dentro del servicio, y según surjan nuevos serán añadidos.
@Id_Estatus	{Numérico}
Status	1{Carácter}50
Status_es	1{Carácter}50
Descripcion	1{Carácter}100
Nota	1{Carácter}500

CLIENTES	Registro de los datos del cliente (paciente).
@Id_Cliente	{Numérico}
DNI	1{Carácter}10
Id_Usuario	{Numérico}
Fecha_Alta	[dd/mm/aaaa; hh:mm:ss: tt]
Email	1{Carácter}50
Nombre	1{Carácter}50
Apellidos	1{Carácter}50
Direccion	1{Carácter}75
Ciudad	1{Carácter}75
Pais	1{Carácter}75
Codigo_Postal	{Numérico}
Numero_Telefono	1{Carácter}12
Movil	1{Carácter}12
Id_Estatus	{Numérico}
Nota	1{Carácter}500

PROVEEDORES	Registro de los proveedores, datos necesarios para conceder el acceso a su zona de usuario.
@Id_Proveedor	{Numérico}
Nombre	1{Carácter}75
Descripcion	1{Carácter}100
Especialización	1{Carácter}100
Fecha_Alta	[dd/mm/aaaa; hh:mm:ss: tt]
Dirección	1{Carácter}75
Teléfono	1{Carácter}12
Representante	1{Carácter}50
Email	1{Carácter}50
Id_Estatus	{Numérico}
Id_Usuario	{Numérico}
Nota	1{Carácter}500

SERVICIOS	Aquí se registrarán todos los servicios ofertados.
@Id_Servicio	{Numérico}
Id_Enfermedad	{Numérico}
Nombre	1{Carácter}75
Descripcion	1{Carácter}100
Costo	{Moneda}
IVA	{Moneda}
Id_Estatus	{Numérico}
Nota	1{Carácter}500

DIAGNOSTICOS	Colocaremos todo lo concerniente a los diagnósticos obtenidos para cada cliente.
@Id_Diagnostico	{Numérico}
Id_Muestra	{Numérico}
Fecha	[dd/mm/aaaa; hh:mm:ss: tt]
Confianza	{Numérico}
Resultado	{Si   No   ...}
Porcentaje_Riesgo	{Numérico}
Id_Estatus	{Numérico}
Id_Cliente	{Numérico}

HISTORICO_DIAGNOSTICO	Tabla para asociar los diagnósticos y así facilitar un histórico a los usuarios.
@Id_Historico	{Numérico}
Id_Diagnostico	{Numérico}
Nota	1{Carácter}500

TIPO_MUESTRAS	Se crearán los distintos tipos de muestras que se puedan utilizar para realizar los diagnósticos (sangre, saliva).
@Id_Tipo_Muestra	{Numérico}
Descripcion	1{Carácter}100
Nota	1{Carácter}500

<b>MUESTRAS</b>	<i>Manejo de las muestras y su tipo, así como el estatus en que se encuentren las mismas.</i>
@Id_Muestra	{Numérico}
Id_Tipo_Muestra	{Numérico}
Fecha_Recepcion	[dd/mm/aaaa; hh:mm:ss: tt]
Fecha_EntregaResul	[dd/mm/aaaa; hh:mm:ss: tt]
Id_Cliente	{Numérico}
Id_Servicio	{Numérico}
Id_Estatus	{Numérico}
Nota	1{Caracter}500

<b>ENFERMEDADES</b>	<i>Registro de las enfermedades, contando con la información detallada de la misma.</i>
@Id_Enfermedad	{Numérico}
Nombre	1{Caracter}75
Descripción	1{Caracter}100
Características	1{Caracter}100
Tratamiento	1{Caracter}75
Id_Tipo_Muestra	{Numérico}
Id_Estatus	{Numérico}
Nota	1{Caracter}500

<b>PUBLICIDAD</b>	<i>Se registraran las publicidades que se lleven a cabo por la empresa.</i>
@Id_Publicidad	{Numérico}
Id_Tipo_Publicidad	{Numérico}
Fecha_Inicio	[dd/mm/aaaa; hh:mm:ss: tt]
Fecha_Fin	[dd/mm/aaaa; hh:mm:ss: tt]
Descripcion	1{Caracter}100
	{Numérico}
Id_Servicio	{Numérico}
Id_Usuario	{Numérico}
Id_Estatus	{Numérico}
Precio_Inicial	{Moneda}
Porciento_Descuento	{Numérico}

<b>TIPO_PUBLICIDAD</b>	<i>Se crearán los posibles tipos de publicidad que se vayan a desarrollar o implementar en la empresa.</i>
@Id_Tipo_Publicidad	{Numérico}
Nombre	1{Caracter}75
Nota	1{Caracter}500

<b>CARRITO</b>	<i>El carrito de compra se irá incrementando según los servicios que decida contratar el usuario.</i>
@Id_Carrito	{Numérico}
Id_Servicio	{Numérico}
Id_Cliente	{Numérico}

<b>PAGOS</b>	<i>Tabla con la descripción del pago (factura), la fecha y el estado.</i>
@Id_Pago	{Numérico}
Id_Cliente	{Numérico}
Descripcion	1{Caracter}100
Fecha	[dd/mm/aaaa; hh:mm:ss: tt]
Tipo_Pago	[Paypal   Transferencia Bancaria   Tarjeta de Credito   ...]
Uso_Personal	[Si   No]
Nota	1{Caracter}500

<b>DETALLE_PAGO</b>	<i>Se creará un detalle de los pagos realizados, donde se tendrán todos los productos que el cliente (paciente) ha seleccionado en el carrito de compra.</i>
@Id_Detalle_Pago	{Numérico}
Id_Carrito	{Numérico}
Importe	{Moneda}
Id_Pago	{Numérico}

#### 4.1.4.1.3 Diseño de Interfaces de Usuario

En esta sección se presentan las distintas interfaces del servicio web, para acceder a ello solo debemos ir a la siguiente dirección web: [www.geneslove.me](http://www.geneslove.me)

Por medio de estas interfaces el usuario accede a la web y realiza sus peticiones, desde consultar información, comprar productos y acceder a sus diagnósticos.

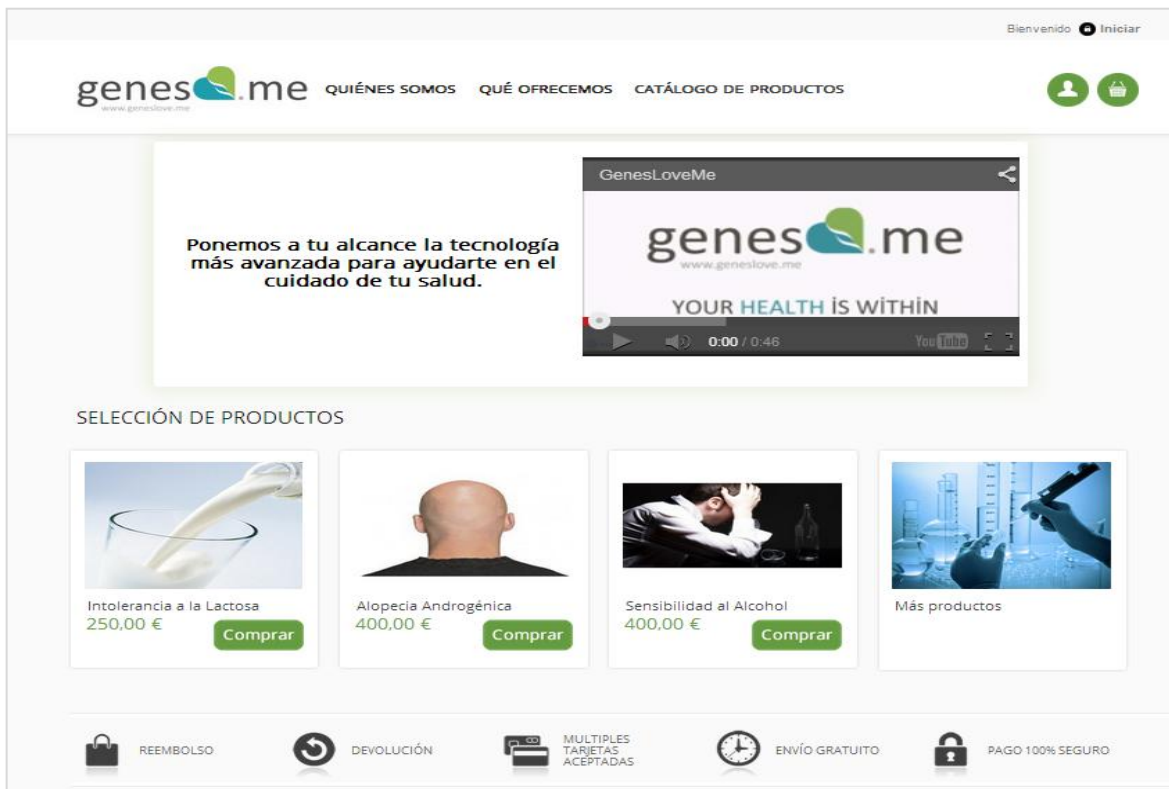


Figura 4.12 Página Principal de GenesLove.me

En la Figura 4.12 observamos un video explicativo sobre qué es lo que hacen y que le facilitan al público en general. Además de una lista de los productos ofertados e información de contacto y/o ayuda.

Bienvenido [Inicio](#)

geneslove.me QUIÉNES SOMOS QUÉ OFRECEMOS CATÁLOGO DE PRODUCTOS

> Catálogo de productos > Sensibilidad al Alcohol

### SENSIBILIDAD AL ALCOHOL

El alcohol es parte intrínseca de nuestra cultura. Ambientes, reuniones o fiestas de trabajo son lugares donde habitualmente se establecen beneficiosas relaciones laborales y sociales.

El Test de Sensibilidad al alcohol te permitirá conocer genéticamente tu grado de tolerancia al alcohol. Así, de una manera sencilla, obtendrás a la vez toda esta necesaria información directamente de tus genes.

Cantidad:  [AÑADIR A LA CESTA](#)

Más información

El metabolismo del alcohol en Humanos tiene un importante componente genético, haciendo que la sensibilidad al alcohol sea un factor poblacional fácilmente identificable mediante técnicas moleculares. Se sabe que esta sensibilidad está directamente relacionada con la enzima ADH2, la cual se encarga de metabolizar el Acetaldehído, metabolito intermedio de la ruta de degradación del alcohol (etanol) a acetato dentro de las células hepáticas del organismo. Este enzima presenta dos formas diferenciadas que genera por el cambio de un péptido en la enzima. Siendo la forma ADH2\*1 la que genera una menor sensibilidad al alcohol al producir una enzima más efectiva que la que producida por ADH2\*2. También se ha visto que la presencia de una de las dos formas está directamente relacionado con la generación de dependencia por el alcohol.

Más información

Figura 4.13 Página de información sobre la enfermedad

En esta sección (Figura 4.13) se presenta una breve descripción de que es y las implicaciones de las enfermedades ofertadas por la empresa.

geneslove.me QUIÉNES SOMOS QUÉ OFRECEMOS CATÁLOGO DE PRODUCTOS

> Qué ofrecemos > Testes genéticos

Geneslove.me te ofrece un surtido de test genéticos capaces de determinar tu predisposición a padecer ciertas enfermedades basándose en la información que proporcionan tus genes.

Respaldados por un grupo de biólogos e informáticos y utilizando la tecnología más puntera en el campo de la genética, nuestros expertos son capaces de proporcionarte tu perfil genético para que puedas personalizar un plan de salud con tu doctor.

En Geneslove.me nos preocupamos por ti y por tanto te ayudamos a que conozcas cómo tus genes pueden impactar en tu vida y en tu salud.

MI CUENTA

Registra tu kit

Mis informes

Mis pedidos

Mis direcciones

Mis datos personales

CONTACTE

Geneslove.me

Camino de Vera s/n

Edificio 6D planta 4a

46022 Valencia – España

Tel: +34 961 042 993

Email: info@geneslove.me

NEWSLETTER

Inscríbete en el boletín

Tu email

Suscribirse

[ENVIAR](#)

Figura 4.14 Info. Test Genéticos: Aquí vemos la página que habla sobre los test genéticos que ofrece la empresa a través de Geneslove.me y cuáles son los fines de los mismos.



Figura 4.15 Diagnósticos Genéticos (Pasos)

Esta página (Figura 4.15) nos muestra en seis sencillos pasos como obtener tu diagnóstico genético, de forma que el usuario entienda el funcionamiento que llevamos a cabo para facilitarle sus resultados.

## Diseño de Reportes y Formularios

A continuación veremos algunos de los reportes y formularios utilizados para gestión de los servicios web. Entre estos se encuentran las siguientes figuras:

- Figura 4.16: Resumen de tu Cesta. En esta página podemos visualizar los productos seleccionados para la compra, así con sus respectivos detalles y opciones de edición de la cesta.
- Figura 4.17: Cuenta de Usuario. En esta sección se nos pregunta si ya estamos registrados, en caso afirmativo nos envía a realizar el inicio de sesión, de lo contrario te solicitará el registro como nuevo cliente (datos personales, entre otros).
- Figura 4.18: Privacidad de los Datos de los Clientes, Procesos de Verificación & Elección del Método de Pago. En esta parte se realizan las verificaciones de lugar y se guardan los datos. El siguiente paso es elegir el método de pago, y para ello se deberá hacer el registro o inicio de sesión.

Bienvenido [Iniciar](#)

genes.me QUIÉNES SOMOS QUÉ OFRECEMOS CATÁLOGO DE PRODUCTOS

Tu cesta

RESUMEN DE TU CESTA Tu cesta contiene: 1 producto

PRODUCTO	DESCRIPCIÓN	PRECIO UNITARIO	CANT.	TOTAL	QUITAR
	Sensibilidad al Alcohol	400,00 €	1	400,00 €	

Envío Envío gratuito!

Subtotal: 330,58 €

IVA 21%: 69,42 €

CUPÓN DESCUENTO  [CANJEAR](#)

TOTAL: 400,00 €

Figura 4.16 Resumen de tu Cesta.

**1 CUENTA**

¿YA ESTÁS REGISTRADO?

» CLIC AQUÍ

---

NUEVO CLIENTE

\* Campo obligatorio

Nombre \*      Apellidos \*      Correo electrónico \*

.....      .....      DNI / NIF / NIE

Contraseña \*      Confirmar contraseña \*      Teléfono personal

Dirección \*      Población \*

Código postal \*      España      Provincia

Usar otra dirección para la facturación


PRIVACIDAD DE LOS DATOS DE LOS CLIENTES

El dato personal que da esta utilizado para responder a sus búsquedas, procesa sus ordenes o autorizase el acceso específico a la información. Tiene derecho de modificar todas las informaciones personales que tenemos sobre usted en pagina "my account".

Figura 4.17 Cuenta de Usuario

PRIVACIDAD DE LOS DATOS DE LOS CLIENTES

El dato personal que da esta utilizado para responder a sus búsquedas, procesa sus ordenes o autorizase el acceso específico a la información. Tiene derecho de modificar todas las informaciones personales que tenemos sobre usted en pagina "my account".



Código de verificación

Para continuar su compra, por favor, haga clic en el botón GUARDAR

**GUARDAR**

---

**2 ELIGE EL MÉTODO DE PAGO**

Por favor, regístrate para ver los métodos de pago

Figura 4.18 Privacidad de los Datos de los Clientes, Procesos de Verificación



### Mapa Navegacional (GenesLove.me)

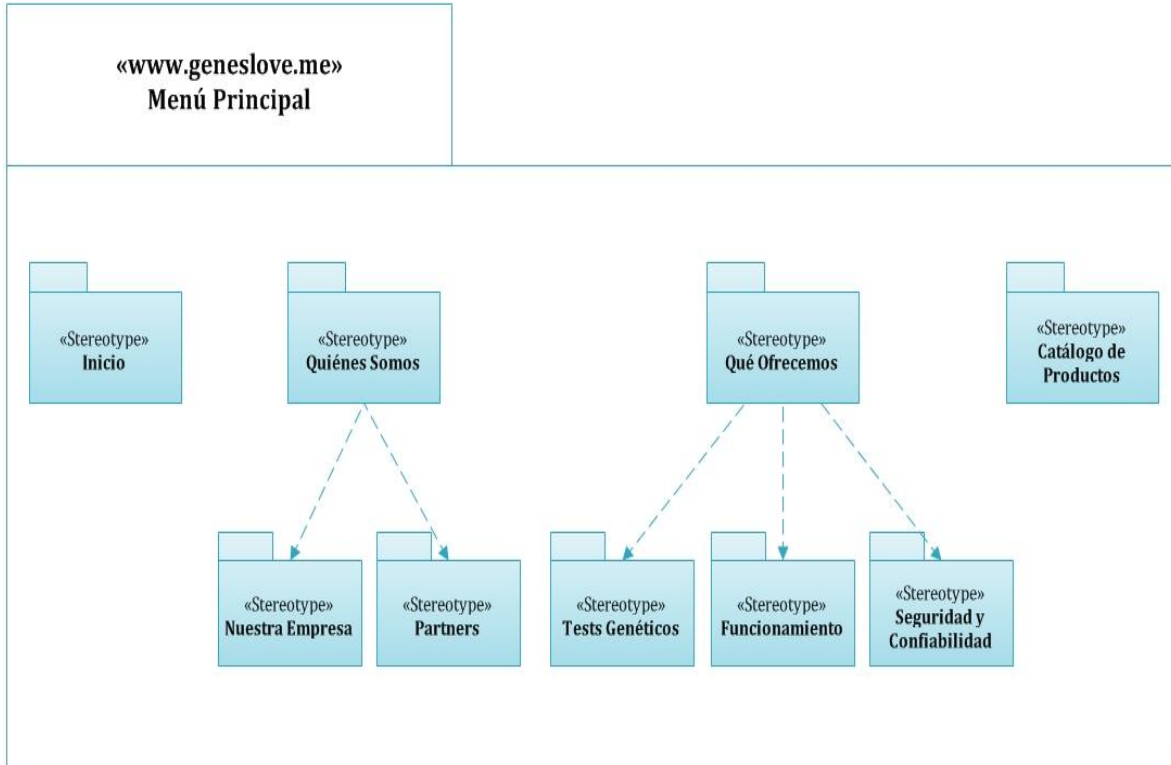


Figura 4.19 GenesLove.me – Menú Principal (1)

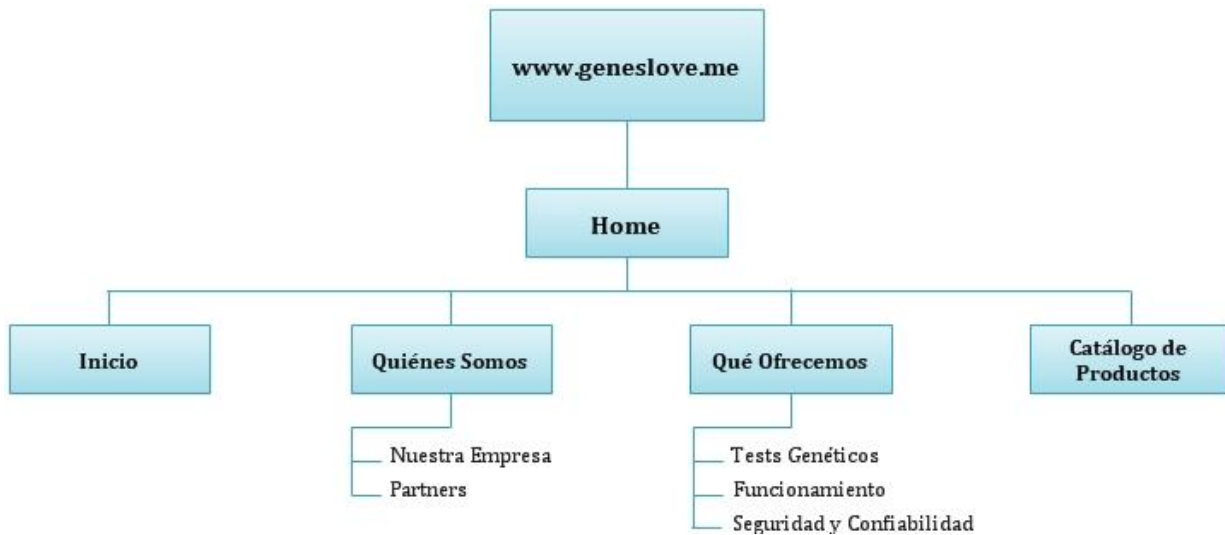


Figura 4.20 Mapa de Navegación de GenesLove.me (2)

#### 4.1.4.2 Contenidos GenesLove.me

##### **Sensibilidad al Alcohol**

Sinónimos: Tolerancia al alcohol, resistencia a la ingesta de alcohol

##### ***¿Qué es la sensibilidad al alcohol?***

La Sensibilidad al Alcohol es el grado de metabolización el alcohol ingerido. Provocando que a mismas concentraciones de alcohol el efecto sobre el organismo tengas efectos claramente diferentes. Este grado de sensibilidad está claramente diferenciado poblacionalmente e individualmente. Teniendo asociado un componente genético y metabólico directo.

La sensibilidad al alcohol difiere del alcoholismo en que está última se trata trastorno psicológico que implica un consumo excesivo y prologando del alcohol generando dependencia y trastornos fisiológicos irreversibles a largo plazo. Este puede está Influenciado por componentes tanto ambientales y genéticos Por lo tanto, cabe distinguir entre sensibilidad y alcoholismo, ya que aunque los componentes ambientales y genéticos están relacionados, describen fenotipos diferentes.

##### ***¿Cuál es el origen de esta sensibilidad?***

El origen de esta enfermedad es genético y ambiental. Identificamos variaciones específicas que se relacionan directamente con la sensibilidad de ingesta del alcohol. Los genes que estudiamos son: ALDH1 y ALDH2, los dos genes que codifican las enzimas de la ruta de metabolismo de alcohol en humanos.

##### ***¿Qué genes intervienen en la determinación del grado de sensibilidad?***

###### **ALDH1B**

Gen situado en el Cromosoma 4, codifica para la enzima alcohol deshidrogenasa II, también llamada alcohol deshidrogenasa 1B, o fragmento beta de la alcohol deshidrogenasa. Esta enzima se encarga del primer paso del catabolismo del alcohol en las células hepáticas del organismo Corresponde al paso de oxidación del etanol a acetaldehído. Asociadas a este gen existen dos variaciones (SNP's), posiblemente relacionadas con el consumo de alcohol.

###### **ALDH2**

Gen situado en el cromosoma 12 que codifica para la enzima Aldehído Deshidrogenasa, segunda enzima en la ruta de eliminación del alcohol. Esta enzima utiliza como sustrato el acetaldehído generado por la enzima alcohol deshidrogenasa, y lo reduce a acetato.

##### ***Más información en el siguiente enlace:***

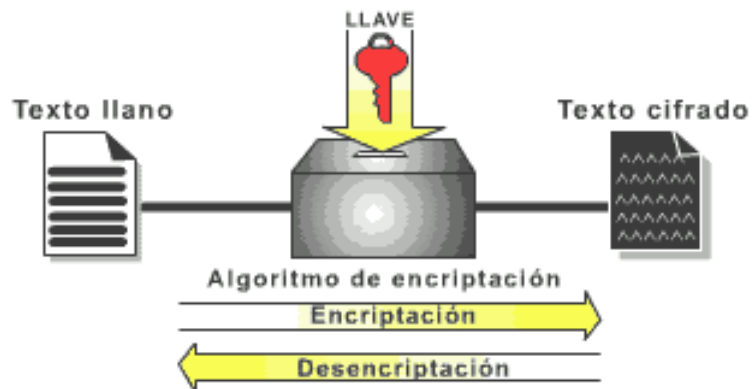
[http://www.geneslove.me/img/cms/prod\\_info/AlcoholSensitivity.pdf](http://www.geneslove.me/img/cms/prod_info/AlcoholSensitivity.pdf)

### 4.1.4.3 Seguridad

Cabe destacar que para el manejo de la información genómica y de los procesos que requieren los diagnósticos se cuentan con altos niveles de seguridad.

Los servidores se encuentran uno dentro del campus universitario y el otro contratado en un servicio externo. Estos cuentan con backups programados de manera semanal, los cuales garantizan un respaldo seguro por si ocurriese algún fallo.

Por igual se cuentan con mecanismos de encriptación de los datos y la utilización de certificados (SSL - Secure Sockets Layer) que permitan realizar los procesos del negocio con altos niveles de seguridad para los usuarios [50] [51].



## Capítulo 5

# Conclusiones & Trabajo Futuro

En esta tesis de máster se ha presentado un estudio y análisis para la puesta en marcha de test genéticos directos al consumidor que permitan conocer la predisposición a padecer a ciertas enfermedades. Con el desarrollo de este trabajo se presentan las necesidades detectadas por los usuarios, y para la puesta en marcha tomamos como caso de estudio la Sensibilidad al Alcohol.

La realización de los BPMN para esta empresa de orientación genómica nos han sido de gran beneficio para poder entender y mejorar los procesos que se llevaban a cabo, esto nos permitió medir la calidad de las tareas realizadas y detectar los procesos que se podían reestructurar para lograr ser eficientes y eficaces al momento de brindar los test genéticos directos al consumidor.

Para la comprobación y validación de los estudios realizados nos fue necesario utilizar la Metodología SILE con un caso real “Sensibilidad al Alcohol”. Con la misma logramos obtener grandes cantidades de datos que permitieron detectar las variaciones/SNP’s que se asociaban con esta enfermedad.

De esta exhaustiva búsqueda surge la problemática de los Haplotipos, existía la necesidad de tratar las variaciones de una manera distinta para poder brindar el diagnóstico. Es por ello que se estudió el Modelo Conceptual del Genoma Humano, y se presentó la solución de cómo asociar las variaciones para tratar los haplotipos. Esto lo integramos al modelo conceptual y a la base de datos, logrando con esto alcanzar un nivel más alto en la generación de los diagnósticos.

Los servicios web creados permiten la óptima gestión de los test genéticos directos al consumidor, facilitando un canal de distribución muy atractivo y de fácil manejo para todos los usuarios.

La bioinformática es una rama que seguirá dando pasos agigantados con el transcurso de los años, porque los beneficios que ha generado y que seguirán surgiendo nos permitirán ver algo que en unos años atrás tal vez era imposible: “Conocer nuestro genoma de una manera tan sencilla.”

La aplicación de los puntos tratados en esta tesis nos permite generar y proporcionar a los usuarios un diagnóstico precoz de calidad para cualquier enfermedad de origen genético.

## 5.1 Trabajo Futuro

Dentro de los trabajos futuros se proyecta el tratamiento de más enfermedades de origen genético. Permittiéndonos por medio del uso de bases de datos cargadas de forma selectiva generar resultados con altos niveles de calidad para los usuarios.

Seguir extendiendo el modelo conceptual del genoma humano para tratar aspectos que sigan mejorando la precisión de los diagnósticos. También implementar mecanismos que brinden un mayor soporte al momento de aplicar metodologías para el tratamiento de nuevas enfermedades, favoreciendo y aumentando los niveles de calidad de la Medicina Personalizada.

Internacionalización, este trabajo podemos extenderlo a otros países y/o naciones, como por ejemplo a la República Dominicana. Los estudios realizados se pueden establecer en empresas y explotarse por igual a todas las enfermedades que posean un origen genético.

## Referencias Bibliográficas

1. Viewing and using NCBI. NCBI National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov/> (2013).
2. Viewing and using NCBI. NCBI National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov/About/index.html> (2013).
3. OMIM (Online Mendelian Inheritance in Man), reviewed in NCBI. NCBI OMIM. <http://www.ncbi.nlm.nih.gov/omim> (2013).
4. About the Ensembl Project. Ensembl. <http://www.ensembl.org/info/about/index.html> (2013)
5. Centro de Investigación en Métodos de Producción de Software ProS. <http://www.pros.upv.es/index.php/es/lineas/77-lineagenoma> (2013).
6. GemBiosoft. <http://www.gembiosoft.com/> (2013)
7. María Liliana Franco, Juan Fernando Cediél, MD, César Payán, MD. Colomb. Med. vol.56 no.1 Cali Jan./Mar. 2008 - On-line version ISSN 1657-9534. [http://www.scielo.org.co/scielo.php?pid=S1657-95342008000100015&script=sci\\_arttext](http://www.scielo.org.co/scielo.php?pid=S1657-95342008000100015&script=sci_arttext)
8. Test Genéticos. [http://www.fisiotraining.com/cordoba/analisisgeneticos\\_quees.htm](http://www.fisiotraining.com/cordoba/analisisgeneticos_quees.htm) (2013)
9. Genetic Test (Información). Grupo Reto Hermosillo A.C. <http://gruporetohermosilloac.com/index.php> (2013).
10. Contributions of Bioinformatics. Bioingeniería. (Argentina). [http://www.bioingenieria.edu.ar/extension/la\\_facultad/carreras/qesbioinf.htm](http://www.bioingenieria.edu.ar/extension/la_facultad/carreras/qesbioinf.htm) (2005).
11. Metzker, M. L. (January 2010). Sequencing Technologies — The Next Generation. Nature Reviews Genetics 11, 31-46.
12. Voelkerding, K. V.; Dames, S. A.; Durtschi, J. D. (2009). Next-Generation Sequencing: From Basic Research to Diagnostics. The American Association for Clinical Chemistry.
13. 23andMe. <https://www.23andme.com/> (2013)
14. Funcionamiento de 23andMe. <https://www.23andme.com/howitworks/> (2013)

15. Escarabajal, M.D.: Alteraciones Genéticas Relacionadas con el Alcoholismo, Área de Psicobiología. Universidad de Jaén. Jaén, España, Email: descara@ujaen.es (2003).
16. Jean Bierut, L.: Genetic Vulnerability and Susceptibility to Substance Dependence. February 2011.
17. Wang, J.; Li, M. D.: Genes and Pathways Co-associated with the Exposure to Multiple Drugs of Abuse, Including Alcohol, Amphetamine/Methamphetamine, Cocaine, Marijuana, Morphine, and/or Nicotine: a Review of Proteomics Analyses, Universidad de Virginia. September 2011.
18. Drogodependencias: Farmacología. Patología Psicología. Legislación, Pedro Lorenzo Fernández, Edición 3, Editor Médica Panamericana, 30/06/2009 - 725 páginas, ISBN: 8498354684, 9788498354683
19. Proyecto HapMap. <http://hapmap.ncbi.nlm.nih.gov/thehapmap.html>.en
20. Sherry, M. H.; Kholodov, W. M.; Baker, J.; Phan, L.; Smigielski, E. M.; Sirotkin, K. (October 4, 2000.). dbSNP: The NCBI Database of Genetic Variation. <http://nar.oxfordjournals.org/content/29/1/308.short>
21. Resumen de dbSNP. [http://www.ncbi.nlm.nih.gov/SNP/snp\\_summary.cgi](http://www.ncbi.nlm.nih.gov/SNP/snp_summary.cgi) (2013)
22. The Effects of Alcohol on the Brain (2013). [http://www.scripps.edu/newsandviews/e\\_20020225/koob2.html](http://www.scripps.edu/newsandviews/e_20020225/koob2.html)
23. Edenberg, H. J.; Dick, D. M.; Xuei, X.; Tian, H.; Almasy, L.; Bauer, L. O.; Crowe, R. R.; Goate, A.; Hesselbrock, V.; Jones, K.; Kwon, J.; Li, T. K.; Nurnberger, JI. Jr.; O'Connor, S. J.; Reich, T.; Rice, J.; Schuckit, M. A.; Porjesz, B.; Foroud, T.; Begleiter, H. (Epub, Mar. 2012). Variations in GABRA2, encoding the alpha 2 subunit of the GABA(A) receptor, are associated with alcohol dependence and with brain oscillations. Indiana University School of Medicine, Indianapolis 46202, USA. [edenberg@iupui.edu](mailto:edenberg@iupui.edu). <http://www.ncbi.nlm.nih.gov/pubmed/15024690>
24. Poblaciones estudiadas en el Proyecto HapMap. <http://www.sanger.ac.uk/resources/downloads/human/hapmap3.html> (2013)
25. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T. L. (March 18, 2008). NCBI BLAST: A Better Web Interface. [http://nar.oxfordjournals.org/content/36/suppl\\_2/W5.short](http://nar.oxfordjournals.org/content/36/suppl_2/W5.short)
26. García Gutiérrez, M. E.; Lima Mompó, G.; Aldana Vilas, L.; Casanova Carrillo, P.; Feliciano Álvarez, V. Alcoholismo y Sociedad, Tendencias Actuales. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0138-65572004000300007](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0138-65572004000300007)
27. Tutoriales PubMed. <http://www.nlm.nih.gov/bsd/disted/pubmed.html> (2013)
28. Página PubMed en NCBI. <http://www.ncbi.nlm.nih.gov/pubmed> (2013)

29. Mapa de Haplotipos – HapMap (conceptos).  
<http://www.genome.gov/GlossaryS/index.cfm?id=98> (2013)
30. Incorporación del Fenotipo Alzheimer Early-Onset A Hgdb, Explicación del BICE – Centro de Investigación en Métodos de Producción de Software (ProS), Universidad Politécnica de Valencia (UPV), Valencia, España.
31. Martín Mayordomo, A. (2011). Integración de Bases de Datos Genómicas: Una Aproximación Basada en Modelado Conceptual. Tesis de Máster en Ing. Software, Métodos Formales & SI, Univ. Politécnica de Valencia, Valencia, España.
32. Martin, A.; Celma, M.: Integrating Human Genome Variation Data: An Information System Approach, International Workshop on Database and Expert Systems Applications, DEXA, pp.65-69. (2011).
33. Villanueva del Pozo, Ma. J. (2011). Diagen: Modelado e Implementación de un Framework para el Análisis Personalizado del ADN. Tesis de Máster en Ing. Software, Métodos Formales & SI, Univ. Politécnica de Valencia, Valencia, España.
34. Pastor, O.; Levin, A. M.; Casamayor, J. C.; Celma, M.; Eraso, L. E.; Villanueva, Ma. J.; Perez, M. (RCIS, 2010, 85-92). Enforcing Conceptual Modeling to Improve the Understanding of Human Genome. Centro de Investigación en Métodos de Producción de Software (ProS), Universidad Politécnica de Valencia (UPV) & El Departamento de Genética de la Universidad de Valencia (UV); Valencia, España.
35. Pastor, O; Casamayor, J. C.; Celma, M.; Mota, L.; Pastor, M. A.; Levin, A. (2012). Conceptual Modeling of Human Genome: Integration Challenges. Centro de Investigación en Métodos de Producción de Software (ProS), Universidad Politécnica de Valencia (UPV), Valencia, España.
36. GeneCards Consulta/Información. <http://www.genecards.org/> (2013).
37. Sánchez Turet, M.; C. Clemente, I. Genética del Alcoholismo: Asociación con Marcadores Biológicos. Depto. Psiquiatría y Psicobiología Clínica. Unidad Psicobiología Clínica. Facultad de Psicología. Universidad de Barcelona, España.
38. Sánchez Pérez, A. M. (Junio, 2000). Sensibilidad al Alcohol y la Predisposición a beber. Ciencia al Día Internacional, Número 2, Volumen 3, ISSN 0717-3849. <http://www.ciencia.cl/CienciaALDia/volumen3/numero2/articulos/articulo3.html>
39. Aragón, C.; Miquel, M.; Correa, M.; Sanchis-Segura, C. (2002). Alcohol y Metabolismo Humano. Área de Psicobiología. Universitat Jaume I. Castelló. Adicciones (2002), Vol. 14, Supl. 1.
40. Hoenicka, J., Ampuero, I., Y Ramos Atance J. A. (2003). Aspectos Genéticos del Alcoholismo. Departamento de Bioquímica y Biología Molecular. Facultad de Medicina. Universidad Complutense. Madrid. España. (Trastornos Adictivos 2003; 5(3):213-22).
41. M. Dick, D.; Foroud, T. (Mayo, 2003). Candidate Genes for Alcohol Dependence: A Review of Genetic Evidence From Human Studies. Vol. 27, No. 5. Alcoholism: Clinical & Experimental Research.



42. J. Duell, E.; Sala, N.; Travier, N.; Muñoz, X.; Boutron-Ruault, M. C.; Clavel-Chapelon, F.; Barricarte, A. (2012). Genetic variation in alcohol dehydrogenase (ADH1A, ADH1B, ADH1C, ADH7) and aldehyde dehydrogenase (ALDH2), alcohol consumption and gastric cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. *Carcinogenesis* vol.33 no.2 pp.361–367, 2012.
43. Alcohol Research: Current Review Update on the Genetics of Alcoholism. National Institute on Alcohol Abuse, Alcoholism (NIAAA) (2013). <http://www.niaaa.nih.gov/publications/journals-and-reports/alcohol-research>
44. Y. Wong, C. C.; Mill, J.; Fernandes, C. (2010). *Drugs and Addiction: An Introduction to Epigenetics*. Institute of Psychiatry, SGDP Research Centre and Department of Psychosis Studies, King's College London, De Crespigny Park, Denmark Hill, London, UK.
45. Jean Bierut, L. (Febrero, 2011). Genetic Vulnerability and Susceptibility to Substance Dependence. Volume 69, Issue 4, 24 February 2011, Pages 618–627. Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, 63110, USA.
46. Fonseca, D.; Silva, C.; Mateus, H.; Restrepo, C. M. (2008). Deletions Identification in Female Carriers of Duchenne's Muscular Distrophy. *Laboratorio de Biología Molecular y Celular. Instituto de Ciencias Básicas. Facultad de Medicina. Universidad del Rosario. Acta Med Colomb vol.33 no.2 Bogotá Apr./June 2008.*
47. E. Bornberg-Bauer and N. W. Paton, "Conceptual data modelling for bioinformatics," *Briefings in bioinformatics*, vol. 3, p. 166, 2002.
48. N. W. Paton, S. A. Khan, A. Hayes, F. Moussouni, A. Brass, K. Eilbeck, C. A. Goble, S. J. Hubbard, and S. G. Oliver, "Conceptual modelling of genomic information," *Bioinformatics*, vol. 16, p. 548, 2000.
49. S. Ram and W. Wei, "Modeling the semantics of 3D protein structures," *Conceptual Modeling-ER 2004*, pp. 696-708, 2004.
50. *Encriptación de datos (Imagen: 2007).*  
<http://encriptaciondedatos.blogspot.com.es/2007/09/encriptacion-de-datos.html>
51. *Certificados SSL (Imagen: 2013).*  
<http://www.profesionalhosting.com/certificados-ssl/>

# Apéndice A

## A.1 Índice de Figuras

### TABLAS

TABLA 2.1 COMPARATIVA DE EMPRESAS QUE OFRECEN TEST GENÉTICOS DIRECTOS AL CONSUMIDOR .....	11
TABLA 3.6 LISTA DEFINITIVA DE REQUISITOS DEL NUEVO SISTEMA (LDR) (1) .....	21
TABLA 3.6 LISTA DEFINITIVA DE REQUISITOS DEL NUEVO SISTEMA (LDR) (2) .....	22
TABLA 4.1 IDENTIFICACIÓN DE GENES & SNP'S .....	35
TABLA 4.2 FILTRADO DE GENES & SNP'S .....	35
TABLA 4.4 RESUMEN GENE CARDS - LOCALIZACIÓN EN GENES .....	37
TABLA 4.5 RESUMEN GENE CARDS II - ASOCIACIÓN DE SNP'S CON POBLACIONES .....	37
TABLA 4.6 ESTUDIO DE FRECUENCIAS EN HAPMAP PARA LAS VARIACIONES DE LA SENSIBILIDAD AL ALCOHOL .....	40

### FIGURAS

FIGURA 3.1 ORGANIGRAMA DE LA EMPRESA .....	15
FIGURA 3.2 LOGO DE LA EMPRESA .....	15
FIGURA 3.3 EL PROCESO DEL DIAGNÓSTICO GENÉTICO .....	18
FIGURA 3.4 EL PROCESO DE REPOSICIÓN DE EQUIPOS (COLECTOR DE MUESTRAS) .....	19
FIGURA 3.5 PROCESOS DEL SERVICIO WEB .....	20
FIGURA 3.7 CASO DE USO GENERAL DEL SISTEMA .....	23
FIGURA 3.8 CASO DE USO GENERAL: USUARIO ANÓNIMO .....	24
FIGURA 3.9: CASO DE USO GENERAL: CLIENTE (PACIENTE) .....	25
FIGURA 3.10 CASO DE USO GENERAL: PROVEEDOR .....	27
FIGURA 3.11 CASO DE USO GENERAL: ADMINISTRADOR .....	29
FIGURA 3.12 DIAGRAMA DE CLASES DE LOS SERVICIOS WEB .....	32
FIGURA 4.3 PROCESOS DE BÚSQUEDA E IDENTIFICACIÓN (METODOLOGÍA SILE) .....	36
FIGURA 4.7 MODELO CONCEPTUAL DEL GENOMA HUMANO .....	42
FIGURA 4.8 MODIFICACIÓN AL MODELO CONCEPTUAL DEL GENOMA HUMANO PARA LA INTEGRACIÓN DE LOS HAPLOTIPOS .....	43
FIGURA 4.9 VISTA DE LA BASE DE DATOS V3	45
FIGURA 4.10 VISTA DE VARIACIONES & ESTRUCTURAL (BASE DE DATOS V3)	46
FIGURA 4.11 DIAGRAMA ENTIDAD-RELACIÓN (1) .....	48
FIGURA 4.11 DIAGRAMA ENTIDAD-RELACIÓN (2) .....	49
FIGURA 4.12 PÁGINA PRINCIPAL DE GENESLOVE.ME .....	52
FIGURA 4.13 PÁGINA DE INFORMACIÓN SOBRE LA ENFERMEDAD .....	53
FIGURA 4.14 INFO. TEST GENÉTICOS .....	53
FIGURA 4.15 DIAGNÓSTICOS GENÉTICOS (PASOS) .....	54
FIGURA 4.16 RESUMEN DE TU CESTA .....	55

FIGURA 4.17 CUENTA DE USUARIO .....	56
FIGURA 4.18 PRIVACIDAD DE LOS DATOS DE LOS CLIENTES, PROCESOS DE VERIFICACIÓN .....	56
FIGURA 4.19 GENESLOVE.ME – MENÚ PRINCIPAL (1) .....	57
FIGURA 4.20 MAPA DE NAVEGACIÓN DE GENESLOVE.ME (2) .....	57
FIGURA C1.1 INTERFAZ GEMBIOSOFT (PÁGINA PRINCIPAL, PARTE SUPERIOR) .....	70
FIGURA C1.2 INTERFAZ GEMBIOSOFT (PÁGINA PRINCIPAL, PARTE INFERIOR) .....	70
FIGURA C1.3 PÁGINA DE INFORMACIÓN SOBRE DIAGEN (SOFTWARE DE ANÁLISIS DE SECUENCIAS GENÓMICAS) .....	71
FIGURA C3.1 INFRAESTRUCTURA DE GEMBIOSOFT .....	74
ESTRUCTURA DE LA TESIS DE MÁSTER .....	75

## Apéndice B

### B.1 Glosario

- **Ácido Desoxirribonucleico (ADN):** El ácido desoxirribonucleico (frecuentemente abreviado como ADN, y también como DNA, del inglés deoxyribonucleic acid) es un ácido nucleico que contiene instrucciones genéticas usadas en el desarrollo y funcionamiento de todos los organismos vivos conocidos y algunos virus, y es responsable de su transmisión hereditaria.
- **Acuerdo de Nivel de Servicio:** Un acuerdo de nivel de servicio o Service Level Agreement, también conocido por las siglas ANS o SLA, es un contrato escrito entre un proveedor de servicio y su cliente con objeto de fijar el nivel acordado para la calidad de dicho servicio.
- **Alelo:** Cada uno de los valores de los nucleótidos de cada par de cromosomas del ADN que situados en la misma posición dentro del genotipo.
- **Bioinformática:** es una disciplina científica emergente que utiliza tecnología de la información para organizar, analizar y distribuir información biológica con la finalidad de responder preguntas complejas en biología.
- **Diagnóstico Genético:** permite conocer la base genética de una enfermedad hereditaria, y la información que puede obtenerse es necesaria para la toma de decisiones en varios aspectos.
- **Examen Genético:** permite el diagnóstico de la vulnerabilidad hacia determinadas enfermedades hereditarias basándose en la genética, y puede ser también utilizado para determinar la ascendencia de una persona.
- **Exámenes Genéticos Directos al Consumidor (DTC):** Los exámenes genéticos DTC son un tipo de pruebas accesibles directamente al consumidor, sin tener que tratarse con un profesional de la salud. Usualmente, para obtener un examen genético, profesionales de la salud tales como médicos consiguen el permiso del paciente y ordenan la prueba deseada.
- **Fenotipo:** es una propiedad observada en el organismo, como la morfología, el desarrollo, o el comportamiento.
- **Formato FASTA:** En bioinformática, el formato FASTA es un formato de fichero informático basado en texto, utilizado para representar secuencias bien de ácidos nucleicos, bien de péptido, y en el que los pares de bases o los aminoácidos se representan usando códigos de una única letra. El formato también permite incluir nombres de secuencias y comentarios que preceden a las secuencias en sí.

- **Gen:** es una secuencia ordenada de nucleótidos en la molécula de ADN (o ARN, en el caso de algunos virus) que contiene la información necesaria para la síntesis de una macromolécula con función celular específica, habitualmente proteínas pero también ARNm, ARNr y ARNt.
- **Genotipo:** es la información hereditaria completa de un organismo, incluso si no se expresa.
- **Genoma:** Conjunto de todo el material genético contenido en cada cromosoma. Está formado por la sucesión de todos los nucleótidos que forman el ADN de una célula. El genoma humano contiene aproximadamente 3.000 millones de pares de nucleótidos.
- **Genoma Humano:** es el genoma del Homo sapiens, es decir, la secuencia de ADN contenida en 23 pares de cromosomas en el núcleo de cada célula humana diploide. De los 23 pares, 22 son cromosomas autosómicos y un par determinante del sexo (dos cromosomas X en mujeres y uno X y uno Y en hombres). El Proyecto Genoma Humano produjo una secuencia de referencia del genoma humano euromático, usado en todo el mundo en las ciencias biomédicas.
- **Medicina Genómica:** La Medicina Genómica es, según la revista científica New England Journal of Medicine, la aplicación del conocimiento del Genoma Humano a la práctica de la Medicina.
- **Mutación:** Variación que tiene un efecto en la creación de proteínas y tiene como consecuencia la aparición de una enfermedad genética.
- **Secuencia genómica:** Secuencia lineal de todos los nucleótidos del genoma.
- **Secuencia genética:** Secuencia lineal de todos los nucleótidos de un gen.
- **Secuenciación:** Métodos y técnicas bioquímicas utilizadas para obtener el genotipo de un individuo a partir de una muestra de ADN.
- **Secuencia de referencia:** Sucesión de nucleótidos ficticia que pretende representar la secuencia de nucleótidos original de todos los individuos de una especie.
- **Variación:** Cambio en un nucleótido de una secuencia genética o genómica de un individuo respecto a la secuencia de referencia de su especie.

# Apéndice C

## C.1 Interfaz GemBiosoft



Figura C1.1 Interfaz GemBiosoft (Página Principal, parte superior)



Figura C1.2 Interfaz GemBiosoft (Página Principal, parte inferior)

The screenshot shows the website for Diagen Soft Plus, a genomic analysis software. The header includes the Gembiosoft logo (Genome Knowledge Software) and navigation links for 'Inicio', 'Oferta Comercial', and 'SOBRE NOSOTROS'. There are also links for 'Iniciar Sesión' and 'Nuevo Usuario' with flags for UK and Spain, and social media icons for Facebook, Twitter, and LinkedIn.

The main content area features the title 'Diagen Soft Plus' and the subtitle 'Herramienta Software para diagnósticos genéticos.' Below this is a large 'Diagen' logo and a paragraph describing the software's capabilities: 'Diagen Soft Plus es un software capaz de analizar secuencias genómicas usando una base de datos donde están relacionados los genes y las mutaciones con las enfermedades. Esta relación se extrae de la información relacionada con las características genéticas del paciente y permite una temprana detección de posibles fenotipos, decisivo para elegir el tratamiento adecuado.'

Another paragraph states: 'Con Diagen Soft Plus los especialistas en laboratorios genéticos pueden alinear secuencias con la secuencia de referencia, detectar mutaciones o buscar en la fuente de datos la información necesaria para realizar un diagnóstico automáticamente. Diagen Soft Plus ofrece un informe para el diagnóstico, el cual después de ser analizado en un laboratorio y confirmado por el doctor, este será informado al paciente.'

A final paragraph says: 'Si estas interesado en comprar o probar Diagen Soft Plus puede ponerse en contacto con nosotros por teléfono o por email'

The right sidebar contains a search bar, a 'Oferta Comercial' section with a link to 'Diagen Soft Plus', and a 'Servicios' section with buttons for 'TUTORIAL', 'FICHA TÉCNICA', 'ÚSELO', and 'VIDEO'. Below this is a 'Noticias' section with a link to 'Emprender en Biomedicina' dated 12 Jun 2013, and another link to 'La "fotografía" genética de la leucemia mieloide aguda' dated 02 May 2013.

Figura C1.3 Página de Información sobre Diagen (Software de Análisis de Secuencias Genómicas)

## C.2 Consentimiento Informado



### Consentimiento Informado

Doy mi consentimiento a GemBiosoft, al igual que al centro o centros designados por el mismo, para que usen esta muestra con el fin de realizar el estudio genético arriba indicado.

#### *Descripción del test genético*

1. El test genético de la Sensibilidad al Alcohol basado en análisis del ADN por secuenciación está indicado para desórdenes en los que se sospecha una anomalía genética.
2. Este test examina el genoma del paciente en búsqueda de variaciones genéticas causantes de la Sensibilidad al Alcohol.
3. Un resultado positivo de la prueba indica que se ha encontrado una alteración genética con significación clínica. Un resultado negativo indica que o bien no se ha encontrado ninguna alteración, o las alteraciones encontradas no tienen significación clínica.
4. El análisis de ADN por secuenciación no es la única técnica para detectar alteraciones genéticas, por lo que mi médico puede recomendar el test antes o después de realizar otras pruebas genéticas.
5. Aunque los métodos de este test son altamente sensibles y específicos, existe un pequeño riesgo de fracaso de la técnica o error de interpretación.

#### *He sido informado y he entendido que:*

- Puedo ser portador o estar afectado de un trastorno hereditario cuyo diagnóstico puede ser confirmado mediante el análisis del ADN obtenido a partir de una muestra biológica.
- En caso de no encontrar alteración genética no quiere decir que no tenga una enfermedad genética.
- Los resultados de esta prueba pueden ser determinantes para mi familia.
- Es aconsejable recibir consejo genético antes y después de la realización de la prueba.
- Ocasionalmente podría producirse un fracaso técnico o una deficiencia de la calidad de la muestra biológica en cuyo caso podría solicitarse una nueva toma de muestra.



- En el caso de estudios indirectos, el resultado se expresará en términos de probabilidad y su fiabilidad será dependiente de la certeza diagnóstica y de la informatividad familiar de los marcadores utilizados. Esta informatividad varía entre pacientes y poblaciones.
- Una vez completado el estudio genético, una alícuota de ADN, debidamente codificada, quedará en custodia del laboratorio de acuerdo a la legislación vigente.
- Toda la información personal (clínica, genética, etc.) será recogida y tratada de forma confidencial respetando en todo momento los principios éticos básicos de la investigación con muestras biológicas y según lo establecido por la legislación vigente (LGS 14/1986, LOPD 15/1999, LBRAP 41/2002, y LIB 14/2007).
- Este consentimiento informado podrá ser revocado en cualquier momento mediante notificación por escrito dirigida a GemBiosoft S.L.
- En el caso de menores, el firmante debidamente identificado (nombre, D.N.I y relación con el paciente), autoriza la participación del menor en el estudio genético.

Nombre:

Apellidos:

DNI:

He leído y acepto los Términos y Condiciones de Uso

## C.3 Infraestructura de la Empresa

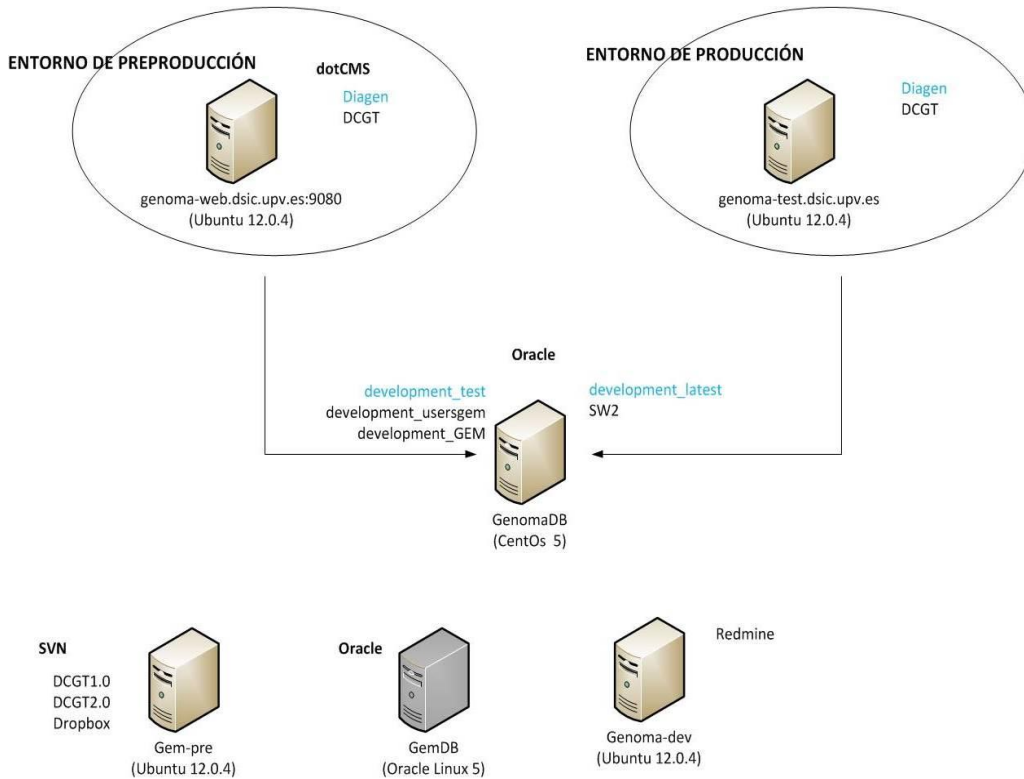
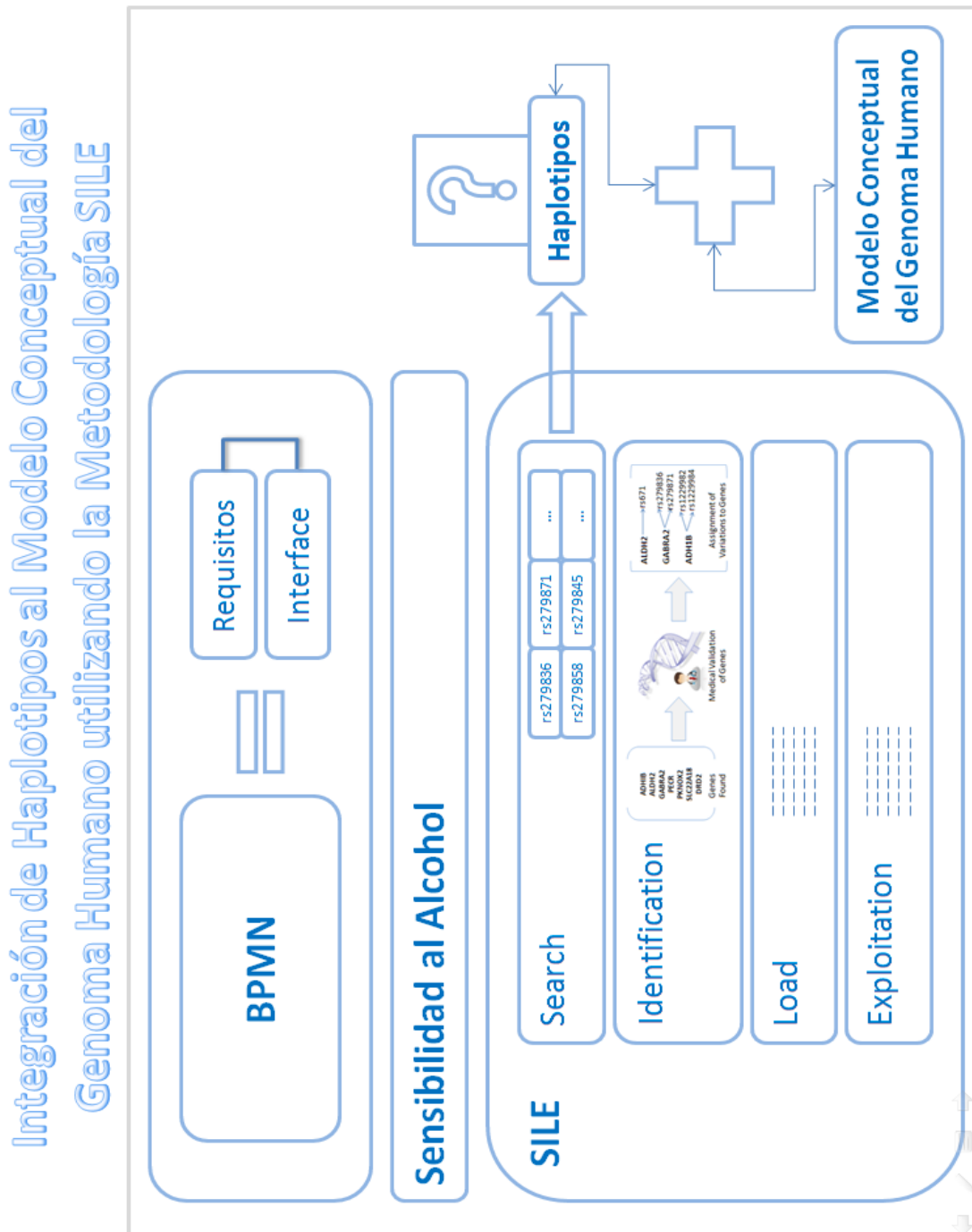


Figura C3.1 Infraestructura de GemBiosoft

## C.4 Estructura Tesis de Máster



Visión General - Estructura de la Tesis de Máster

