

# **Sistema analizador de genoma vacuno para predecir enfermedades endógenas**

*Trabajo Terminal No. \_\_\_\_-\_\_\_\_*

*Alumnos: Alcántara López Óscar Daniel, \*González Cruz Miguel Ángel, Malpica Santes Gonzalo Andrei*

*Directores: Rosas Trigueros Jorge Luis, Palma Orozco Rosaura*

*\*e-mail: mgonzalezc1302@alumno.ipn.mx*

**Resumen** - Se construirá un prototipo de sistema bioinformático para analizar el genoma vacuno, el cual ayudará a predecir posibles enfermedades que se puedan presentar a futuro en el ganado. Para ello se identificarán las similitudes del genoma vacuno con el de otra especie previamente afectada por una enfermedad y mediante una técnica de cómputo evolutivo se implementará el proceso de análisis utilizando algoritmos basados en evolución biológica, dando como resultado la probabilidad que tiene el genoma vacuno de padecer alguna enfermedad endógena. El prototipo presentará una interfaz clara y sencilla para que usuarios de distintas áreas académicas o laborales puedan utilizarlo.

**Palabras Clave** - Bioinformática, Cómputo Evolutivo, Estadística, Genoma.

## **1. Introducción**

Uno de los principales retos de la medicina reconocido por muchos investigadores desde el secuenciamiento del genoma, ha sido identificar a escala genómica la variación genética que puede estar asociada con algunas enfermedades. Líderes mundiales han trabajado en el desarrollo de metodologías experimentales, modelos matemáticos y computacionales para continuar estudiando e identificando las variaciones que pueden detectarse a partir del análisis genómico [1].

La genética animal es un aspecto importante en los análisis del genoma vacuno, el cual ayudará a predecir posibles enfermedades que se puedan presentar a futuro en el ganado. Estudios de diversos investigadores, donde esto permite a criadores y productores, crear estrategias para mantener un control de calidad. Esta herramienta genómica, se basa en los cambios de las frecuencias genéticas de la población de estudio, generalmente determinada por programas de selección y las condiciones económicas de la producción. Estos cambios en el genoma animal nos permiten entender el desarrollo de posibles enfermedades que puedan presentar estas especies a futuro [2].

Entre las posibles enfermedades que se pueden presentar en el ganado vacuno tenemos las enfermedades endógenas, causadas por una alteración en el genoma (mutaciones). Esta alteración puede ser hereditaria o no, causando proteínas defectuosas o alterando la regulación entre genes, lo cual puede causar diversos tipos de cáncer, trastornos metabólicos, malformaciones y otros problemas. Esto da gran valor a la visualización y detección de las enfermedades a tiempo [3].

Con esto se puede decir que debemos de cuidar la genética del ganado vacuno ya que es una de las especies animales con más importancia económica y gastronómica en diversos países. En México no es la excepción, siendo esta especie una de las más abundantes en el país, por lo cual es posible identificar fenotípicamente los mejores ganados vacunos en un programa de mejoramiento genético, teniendo claridad sobre las características cuantitativas que se desean seleccionar por medio de evaluaciones genéticas. Para hallar esas diferencias o polimorfismos genéticos, en la actualidad se dispone de técnicas que permiten evaluar esas diferencias a nivel del DNA desde estadios muy tempranos del individuo. Dichas técnicas se caracterizan por ser no invasivas, actualmente más económicas, fáciles de usar y aportan mucha información acerca de la composición genética de un individuo o población respecto a una característica determinada [2].

Refiriéndose a las técnicas computacionales para evaluar las diferencias respecto al genoma vacuno existe el cómputo evolutivo el cual consta de un apartado de algoritmos que ayudan al desarrollo de este proyecto, estos algoritmos imitan los principios evolutivos naturales para constituir procedimientos de búsqueda y optimización [4].

Para el desarrollo de este proyecto se analizaron diversos trabajos de investigación cercanos al tema principal planteado en este proyecto, una comparativa y breve explicación de los mismos se presenta en la Tabla 1.

**Tabla 1.** Resumen de productos similares.

<b>SOFTWARE</b>	<b>CARACTERÍSTICAS</b>	<b>PAÍS DE ORIGEN</b>
DNA Testing & Sample Submission - Zoetis Genetics Canadá [5].	Análisis de ADN de ganado en busca de optimizar la selección y el manejo de hembras de reemplazo para productos lácteos.	Canadá
BCM-HGSC Bovine Genome Project [6].	Secuenciación y análisis del genoma.	EUA
Introgresión de genes bovino cebú del Sahel [7].	Este sistema se utiliza dentro de una investigación para la detección y evaluación de una enfermedad en ganado bovino que se encuentra en un zona específica.	EUA/Sudáfrica
Análisis del genoma completo de caracteres de crecimiento en ganado charolais de registro [8].	Tesis de análisis del ganado charolais para la identificación de regiones genómicas asociadas al crecimiento de esta raza.	México
Secuenciación completa del genoma de cepas de salmonella enterica obtenidas en diferentes segmentos de la cadena productiva en bovinos de carne [9].	Tesis de análisis referida a cómo afecta en forma genérica la salmonelosis, una enfermedad que ataca a bovinos de carne.	México
Sistema analizador de genoma vacuno para predecir enfermedades endógenas	Software que analizará el genoma vacuno, en donde a diferencia de los trabajos anteriores, se enfocará en analizar el genoma para poder predecir futuras enfermedades que puedan padecer el ganado vacuno, cuidando la especie para evitar futuros problemas en ella.	México

## 2. Objetivo

Construir un sistema bioinformático para comparar el genoma de ganado vacuno Bos Taurus con muestras de su propia especie mediante técnicas de Cómputo Evolutivo para predecir posibles enfermedades endógenas.

## 3. Justificación

En México, 10.7% del gasto en productos de consumo masivo de los hogares, es destinado a la leche líquida. Es la segunda categoría en la que las familias mexicanas gastan más. En el último año a abril el 99.8% de los hogares compraron este producto derivado de la vaca. 217 litros es lo que compra una familia promedio en un año, llevándose 2.4 litros en cada viaje de compra. El desembolso que tiene una familia en un año es de \$2,379

[10].

En México:

- Se destinan alrededor de 109.8 millones de hectáreas al ganado vacuno.
- 842 mil personas alimentan y cuidan al hato ganadero.
- México es el 11° productor mundial de ganadería primaria.
- De los 51.9 millones de mexicanos que trabajan, 739 mil se dedican a la cría y explotación de especies ganaderas.
- En 2017, se produjeron 11,807.5 millones de litros de leche provenientes del bovino [11].

Hablando de productos lácteos provenientes de la vaca, Estados Unidos es el país que más leche produce, con 98.8 millones de toneladas métricas, seguido por India, con 76 y Rusia con 34.1. En este rubro, México aparece en las primeras 10 posiciones y ocupa específicamente el séptimo sitio [12].

Es por eso que se considera importante ser capaces de identificar los genes de estos animales y someterlos a un análisis que compruebe su predisposición a padecer ciertas enfermedades. Esto con el fin de ayudar a extender al máximo su periodo de vida y minimizar el costo causado por las enfermedades. A su vez, el interés de ciertas empresas especializadas en el desarrollo de productos lácteos y el constante crecimiento de esta industria es un indicador de que este estudio puede aportar una mejora para la industria [13].

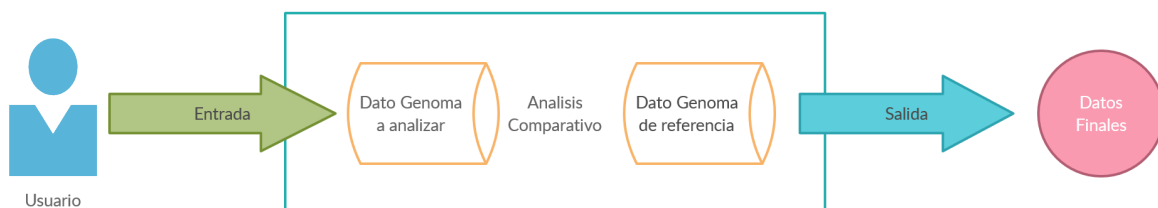
Hemos identificado una base de datos de acceso libre con información genética de bovinos, donde podemos encontrar información acerca de los genes y proteínas que requiere el proyecto. Esta base de datos constituye una base fundamental del proyecto [14]. También podemos aprovechar datos del repositorio Gene Expression Omnibus [15].

Dicho lo anterior y teniendo en cuenta el alcance del sistema, se prevé que este proyecto puede ser de utilidad para las personas que se dedican al sector ganadero, puesto que podrán evitar muchos gastos en un futuro y poder reducir la posibilidad de que sus negocios se vean afectados.

Se tiene contemplado, para la elaboración de este proyecto, un tiempo aproximado de 6 meses académicos, teniendo en cuenta que la investigación y el desarrollo de dicho proyecto, estén en tiempo y forma para su correcto uso al finalizar este período establecido de tiempo.

#### 4. Productos o Resultados esperados

El sistema mediante un algoritmo evolutivo comparará dos datos genómicos, uno que ingresa el usuario y otro que está de referencia en el sistema con los datos de alguna enfermedad endógena. Como resultado indicará si el genoma ingresado por el usuario presenta indicios de predisposición hacia alguna enfermedad.



**Figura 1.** Arquitectura del sistema.

Los productos finales del sistema serán

1. Software del sistema.
2. La documentación técnica del sistema.
3. Código fuente del Sistema.
4. El manual de usuario.
5. Publicación.

## **5. Metodología**

Para el desarrollo del proyecto se ha decidido usar una metodología estructurada orientada al flujo de datos, dentro de estas especificaciones usaremos la metodología denominada “Yourdon”.

Se realizará el enfoque en 7 partes específicas para el trabajo en el proyecto. llevando a cabo primero las pruebas modulares, luego las pruebas del subsistema y finalmente las pruebas del sistema mismo. este enfoque también se conoce como el ciclo de vida en cascada [16].

Cada una de las 7 partes se muestran a continuación:

### **1.- Requerimientos del sistema**

Se evalúan las especificaciones necesarias para que el sistema pueda funcionar adecuadamente

### **2.- Requerimientos del software**

Especificamos la descripción del sistema que se va a desarrollar además de definir las herramientas asociadas a la tecnología como lo son

- Diagramas de flujo
- Diagramas entidad - relación
- Diagramas Estructurados
- Descripciones funcionales

### **3.- Análisis**

Implica el desarrollo de un modelo ambiental y el desarrollo de un modelo de comportamiento. Estos dos modelos se combinan para formar el modelo esencial que representa una descripción formal de lo que el nuevo sistema debe hacer, independientemente de la naturaleza de la tecnología que se use para cubrir los requerimientos [16].

### **4.- Diseño del Programa**

Se proponen 4 pasos para el diseño del programa

- Trazar diagrama de flujo de datos
- Trazar diagrama de estructura
- Evaluación del diseño
- Preparación del diseño para la implantación

### **5.- Codificación**

Se programará el código fuente del sistema con base en los requerimientos previamente definidos, para poder llevar a cabo el funcionamiento correcto del sistema.

### **6.- Pruebas**

Se realizarán pruebas al sistema, para comprobar que esté operado de manera óptima y corregir errores en caso de que existan.

## 7.- Operaciones

Se analizan que los procesos funcionen adecuadamente y con calidad para dar por hecho el resultado final del proyecto

## 6. Cronograma

Alumno : Miguel Angel Gonzàlez Cruzot

<b>Actividad</b>	<b>FEB</b>	<b>MAR</b>	<b>ABR</b>	<b>MAY</b>	<b>JUN</b>	<b>AGO</b>	<b>SEP</b>	<b>OCT</b>	<b>NOV</b>	<b>DIC</b>
Identificar y definir requisitos										
Estudiar alternativas de construcción.										
Realizar el análisis detallado del nuevo sistema.										
Completar especificaciones de entrega.										
Diseñar la estructura física de datos del sistema.										
Desarrollar y probar componentes del sistema.										
Determinar necesidades especiales para el funcionamiento del sistema.										
Presentación TT1										
Diseñar y realizar las pruebas del sistema										
Establecer procedimientos de producción										



Alumno : Alcántara López Oscar Daniel

[illegible]

## 7. Referencias

- [1] C. I. Bermudez Santana, "Aplicaciones de la bioinformática en la medicina: el genoma humano. ¿Cómo podemos ver tanto detalle?" *Acta Biologica Colombiana*, no. 19. pp. 351-255.
- [2] Ángel-Marín, Paula, et al. "GENÓMICA EN LA PRODUCCIÓN ANIMAL." *Rev. Colombiana Cienc. Anim*, vol. 5, no. 2, 2013, pp. 497–518.
- [3] "Enfermedades Endógenas y Exógenas," prezi.com, 2020. [Online] Disponible en: <https://prezi.com/h8y3q2eeyeha/enfermedades-endogenas-y-exogenas/>.
- [4] "COMPUTACIÓN EVOLUTIVA." [Online]. Disponible en: [http://www.dia.fi.upm.es/~ajimenez/Docu\\_Metaheurísticas/Transparencias/Tema5Evolutivos.pdf](http://www.dia.fi.upm.es/~ajimenez/Docu_Metaheurísticas/Transparencias/Tema5Evolutivos.pdf).
- [5] "Animal Genetics | Zoetis," Zoetis.com, 2013. [Online] disponible en: <http://www.zoetis.com/products-services/animal-genetics.aspx>.
- [6] "Bovine Genome Project," BCM-HGSC, Mar. 04, 2016. [Online] Disponible en: <https://www.hgsc.bcm.edu/other-mammals/bovine-genome-project>.
- [7] "El SERIDA participa en el proyecto: Introgresión de genes de bovino cebú del Sahel en las poblaciones bovinas tripanotolerantes del oeste de África. Servicio Regional de Investigación y Desarrollo Agroalimentario," [Online] Disponible en: [Serida.org](http://www.serida.org/vernoticia.php?id_noticia=792), 2019.
- [8] Biól. Francisco Joel Jahuey Martínez, "Análisis del genoma completo de caracteres de crecimiento en ganado charolais de registro.", Tesis de maestría, Centro de biotecnología genómica, IPN, Reynosa, Tamps, 2014.
- [9] E. J. Delgado Suárez, "Secuenciación completa del genoma de cepas de salmonella enterica obtenidas en diferentes segmentos de la cadena productiva de bovinos carne," Tesis Doctoral, Ciencias Biológicas, UNAM, Ciudad de México 2019.
- [10] M. L Ayala, "México: ¿Cuánta leche consumen los Mexicanos? | América Retail," *America-retail.com*, Jun. 2017. <https://www.america-retail.com/mexico/mexico-cuanta-leche-consumen-los-mexicanos/>.
- [11] Servicio de Información Agroalimentaria y Pesquera, "La ganadería: símbolo de fortaleza del campo mexicano," *gob.mx*, 2018. <https://www.gob.mx/siap/articulos/la-ganaderia-simbolo-de-fortaleza-del-campo-mexicano>.
- [12] Énfasis Alimentación, "Señalan a India como el mayor consumidor de leche," *énfasis*, 2020. <http://www.alimentacion.enfasis.com/notas/83165-senalan-india-como-el-mayor-consumidor-leche>.
- [13] "Portal Lechero - Mundo: Estos son los países con mayor consumo de leche," *Portalechero.com*, 2019. <https://www.portalechero.com/innovaportal/v/14091/1/innova.front/mundo:-estos-son-los-paises-con-mayor-consumo-de-leche.html>.
- [14] "U.S Cattle Genome Coordination Program," *Animalgenome.org*, 2014. <https://www.animalgenome.org/cattle/>.
- [15] Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* 2013 Jan;41(Database issue):D991-5.
- [16] Yourdon, E., 1996. *Análisis Estructurado Moderno*. México: Prentice-Hall Hispanoamericano, pp.87-94.



## 8. Alumnos y Directores

Miguel Angel González Cruz.- Alumno de la carrera de Ing. en Sistemas Computacionales ESCOM, Especialidad Sistemas, Boleta:2014120531, Tel.5585654641.

Correo electrónico: [mgonzalezc1302@alumno.ipn.mx](mailto:mgonzalezc1302@alumno.ipn.mx)

Gonzalo Andrei Malpica Santes.- Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta:2014030771, Tel.5539038103.

Correo electrónico: [gmalpicas1300@alumno.ipn.mx](mailto:gmalpicas1300@alumno.ipn.mx)

Oscar Daniel Alcántara López.- Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta: 2017630021, Tel. 5546926242.

Correo electrónico: [ocalcantaral1601@ipn.mx](mailto:ocalcantaral1601@ipn.mx).

Rosaura Palma Orozco. Dra. En Tecnología Avanzada por el CICATA del IPN (2012), M. en C. en Matemáticas por el CINVESTAV del IPN (2004), Ing. en Sistemas Computacionales por la Escuela Superior de Cómputo del IPN (1998). Actualmente es profesora Titular en ESCOM y sus áreas de interés son: Optimización Combinatoria, Sistemas Evolutivos, Teoría de Filtrado Lineal Estocástico y Biología Sintética.

Correo electrónico: [rpalma@ipn.mx](mailto:rpalma@ipn.mx)

Jorge Luis Rosas Trigueros. Dr. en Biotecnología por el IPN, (2012), M. en C. en Ing. Eléctrica por la Universidad de Texas A & M en College Station, Estados Unidos (2002), Ing. en Sistemas Computacionales por la Escuela Superior de Cómputo del IPN (1998). Actualmente es profesor Titular en ESCOM y sus áreas de interés son: Sistemas Evolutivos, Bioinformática y Graficación.

Correo electrónico: [jlrosas@ipn.mx](mailto:jlrosas@ipn.mx)

CARÁCTER: Confidencial


FUNDAMENTO LEGAL: Artículo 11 Fracc. V y Artículos 108, 113 y 117 de la Ley Federal de Transparencia y Acceso a la Información Pública.

PARTES CONFIDENCIALES: Número de boleta y teléfono

## Formato de acuses Directores e Integrantes

Protocolo: Sistema analizador de genoma vacuno para predecir enfermedades endógenas

### Director 1: Jorge Luis Rosas Trigueros



Jorge Luis Rosas Trigueros <jlrosas@ipn.mx>

Lun 09/11/2020 08:26 PM

Para: Usted

↩ ↶ ➜ ...

Buenas noches.

Acuso de recibido.

Dr. Jorge Luis Rosas Trigueros


Obtener [Outlook para iOS](#)

...

Responder

Reenviar

### Director 2: Rosaura Palma Orozco



Rosaura Palma Orozco <rpalma@ipn.mx>

Lun 09/11/2020 08:03 PM

Para: Usted

↩ ↶ ➜ ...

Estimado Estudiante,

Acuso de recibido.

Saludos cordiales  
Rosaura Palma Orozco

...

Responder

Reenviar

### Integrante 1: Miguel Angel González Cruz



Miguel Angel Gonzalez Cruz

Lun 09/11/2020 20:13

Para: Oscar Daniel Alcantara Lopez

👍 ↩ ↶ ➜ ...

Acuso de recibido  
Miguel Angel González Cruz

...

Responder

Reenviar

## Integrante 2: Gonzalo Andrei Malpica Santes



Gonzalo Andrei Malpica Santes <gmalpicas1300@alumno.ipn.mx>

Lun 09/11/2020 08:18 PM

Para: Usted

Acuso de Recibido

Malpica Santes Gonzalo Andrei


...

Responder

Reenviar



## Integrante 3: Oscar Daniel Alcántara López



Oscar Daniel Alcántara López <oalcantara1601@alumno.ipn.mx>

Lun 09/11/2020 07:56 PM

Para: Usted

Acuso de recibido.

Oscar Daniel Alcántara López.

...

Responder

Reenviar

