

**Pasantía en biología computacional y molecular dentro del
grupo de investigación en Ciencias Básicas, Universidad
CES.**

Estudiante
Juan Pablo Barón Vera

Tutor
Ronald Guillermo Peláez Sánchez

Tutor enlace
Juliana Martínez Garro

Trabajo de Grado
En la modalidad de *Pasantía*

Programa de Biología
Universidad CES
Medellín
Oct 2020

11 de noviembre de 2020.

Se informa que el estudiante **Juan Pablo Barón Vera** identificado con cédula: No. 1152466739 ha concluido de manera satisfactoria su trabajo de grado titulado **Pasantía en biología computacional y molecular dentro del grupo de investigación en Ciencias Básicas, Universidad CES**, en la modalidad de *Pasantía*.

En calidad de **director(es)** del trabajo de grado en mención, y luego de haber revisado con detalle y alto rigor científico y académico el presente documento final, se aprueba este Trabajo de Grado como requisito parcial para optar al título de **Biólogo**.

Ronald G. Peláez Sánchez

Tutor: Ronald Peláez Sánchez
Cédula: 71.352.460
Afiliación: Universidad CES

Tutor enlace: Juliana Martínez
Cédula: 2178862
Afiliación: Universidad CES

Pasantía en biología computacional y molecular dentro del grupo de investigación en Ciencias Básicas, Universidad CES.

Juan Pablo Barón Vera

Resumen

La biología computacional es el uso de herramientas informáticas para la solución de problemas inherentes a la biología, esto se debe al actual volumen de datos disponibles y en proceso de generación que imposibilita la realización manual de los análisis. Haciendo uso de las herramientas bioinformáticas podemos dar respuesta de forma sencilla y eficiente a las diferentes preguntas bases dentro de este proyecto, el cual se enmarca dentro de un macro proyecto financiado por Colciencias que busca la identificación de los diferentes agentes zoonóticos causantes de síndrome febril agudo no palúdico.

Para lograr una eficiente identificación de estos agentes es necesario el uso de técnicas moleculares y bioinformáticas tales como la extracción, purificación, cuantificación y la secuenciación de ADN, para finalmente con los datos generados llenar vacíos en el conocimiento sobre la presencia de los agentes de infecciones zoonóticas en el Urabá antioqueño y adicionalmente proveer de información a las entidades pertinentes para tomar decisiones de salud pública que afecten de una forma positiva a las comunidades cercanas a los sitios de muestreo; todo eso como resultado de un proceso de formación por parte del pasante el cual será entrenado en las diferentes técnicas y softwares necesarias para la realización de este proyecto.

Palabras clave: Bioinformática, Factores de virulencia, Modelación de proteínas, Toxoplasma

TABLA DE CONTENIDO

1.	PRESENTACIÓN.....	5
2.	RESEÑA DE LA INSTITUCIÓN.....	6
3.	OBJETIVOS.....	6
3.1	OBJETIVO GENERAL.....	6
3.2	OBJETIVOS ESPECÍFICOS.....	7
4.	LOGROS ALCANZADOS.....	7
5.	DIFICULTADES.....	7
6.	RESULTADOS.....	8
6.1	DETECCION Y MODELACION DE GENES DE VIRULENCIA POR MEDIOS COMPUTACIONALES EN EL GENERO LEPTOSPIRA	8
6.2	REALIZACION DE PCRs PARA DETECCION DE TRYPANOSOMAS EN SANGRE Y CORAZON.....	8
6.3	ARTICULO CIENTIFICO.....	9
7.	CONCLUSIONES.....	9
9.	ANEXOS	10
10.	BIBLIOGRAFÍA.....	11

Presentación

La biología computacional es el uso de herramientas informáticas para la solución de problemas inherentes a biología, tales como el análisis de secuencias, la creación de cadenas de polipéptidos grandes a partir de secuencias de ADN, el uso de cálculos y estadísticos complejos aplicados a análisis genómicos, además de la ilustración de diferentes objetos como las proteínas; esto debido a que el gran volumen de datos imposibilita la realización manual de los análisis (1), y genera la necesidad de incursionar en la creación de plataformas, aplicaciones y algoritmos que permitan analizar estos datos. Por esta razón y considerando que la biología computacional ha evolucionado de manera acelerada desde el proyecto del genoma humano, se evidencia la necesidad de aprender a usar las herramientas que provee esta área del conocimiento en los diferentes campos de acción a los que puede enfrentarse un biólogo, ya sea agricultura, genética, ecología, biotecnología o incluso veterinaria(2). En pro de mejorar el uso de estas habilidades se planteo un proyecto de pasantía.

El proyecto de pasantía fue realizado en el Grupo de Investigación en Ciencias Básicas de la Universidad CES, fue posible gracias a Ronaldo Guillermo Peláez Sánchez, el cual, realizó las funciones de tutor, siendo un gran apoyo en todas las labores que debí realizar dentro del laboratorio, y en los cursos que tomé para adquirir los conocimientos teóricos necesarios para el buen desarrollo de la pasantía, y a Juliana Martínez Garro, la cual, realizó las funciones de tutora de enlace, siendo siempre un apoyo incondicional en todos los trámites necesarios para el desarrollo de los diferentes cursos de investigación y además al grupo de investigación en Ciencias Básicas el cual aportó los recursos necesarios para todas las labores de laboratorio. Este proyecto de pasantía esta enmarcado dentro de un macro proyecto de Colciencias desarrollado por el grupo de investigación, el cual busca identificar agentes zoonóticos en diferentes poblaciones de murciélagos en el Urabá antioqueño, los cuales podrían afectar a las poblaciones humanas cercanas a los lugares de captura. Si bien, esta pasantía se inició el 21 de febrero de 2020, las labores de capacitación y el acompañamiento con miembros del Grupo de Investigación iniciaron desde 2019-2, esto en pro de tener más tiempo para fortalecer habilidades y conocimientos pertinentes del investigador en el área de la salud. Los dos procesos principales en la pasantía fueron: la detección de organismos generadores de enfermedades zoonóticas por medios moleculares, y la detección bioinformática de factores de virulencia en el género *Leptospira*. Con lo anterior, se buscaba fortalecer las diferentes habilidades necesarias de un biólogo que se dedique a la investigación en el área de la salud tales como la realización de diferentes técnicas moleculares, la realización de informes y reuniones semanales, además de ampliar el conocimiento sobre las enfermedades infecciosas, generando datos que permitan tomar decisiones oportunas para mejorar la salud de las diferentes comunidades.

Reseña de la institución

La Universidad CES es una entidad autónoma, de carácter privado no confesional, sin ánimo de lucro, en busca de la formación de profesionales en pregrado y posgrado en todas las áreas del conocimiento, con las más altas cualidades humanas, éticas y científicas (3).

Dentro de su estructuración, la investigación constituye una de las funciones dentro del quehacer universitario, en busca de consolidar una cultura científica armónica y concordante con la formación académica en todas sus dimensiones. Por esta razón, la conformación de grupos de investigación multidisciplinarios, en los cuales, la suma de las experiencias y trabajo, tanto de docentes y estudiantes, se convierten en el escenario perfecto para la generación de ideas de investigación, la consolidación de semilleros de investigación, la producción y divulgación científica. Es por esta razón que la Universidad CES posee un total de 33 grupos avalados por Colciencias, dentro de los cuales, se encuentra el Grupo de investigación en ciencias básicas (4).

Este grupo hace parte del área de conocimiento de ciencias médicas y de la salud, pertenece a la Facultad de Medicina, y es categorizado como A1 en las clasificaciones de Colciencias. Está liderado por la doctora Piedad Matilde Agudelo Flórez, y fue conformado en el mes de enero del año 2002, en la ciudad de Medellín – Antioquia. El grupo de investigación posee varios objetivos generales dentro de su organización, entre los cuales están: Desarrollar proyectos de investigación que impulsen el adelanto científico de las comunidad universitaria y el país, formular proyectos de investigación que busquen el avance del conocimiento de las patologías que aquejan las poblaciones desde un punto de vista tanto fisiopatológico como epidemiológico, impulsar trabajos de investigación en colaboración con otras instituciones tanto a nivel nacional como internacional, propiciar la creación de nuevas líneas terapéuticas para diferentes patologías e incentivar el pensamiento científico de los estudiantes de pregrado con un enfoque en sus líneas de investigación, dentro de los cuales se destacan algunos temas como: Epidemiología clínica y molecular, genómica de agentes infecciosos y biología computacional (5).

Objetivos

1.1 Objetivo general

Potencializar las diferentes habilidades necesarias para el quehacer de un biólogo en el campo laboral con un enfoque bioinformático y molecular.

1.2 Objetivos específicos

- Aprender a implementar las diferentes herramientas útiles en el desarrollo de las labores de un biólogo en el ambiente profesional dentro de un grupo de investigación.
- Revisar las técnicas y protocolos necesarios para el desarrollo del trabajo bioinformático y en biología molecular.
- Desarrollar habilidades útiles para las diferentes actividades a las que un biólogo debe enfrentarse en la vida laboral en el campo de la biología molecular y las ciencias de la salud.

Logros alcanzados

- Asistir y cursar el proyecto curricular de biología computacional 1 en la Universidad de Antioquia, con un desempeño sobresaliente.
- Realizar prácticas de investigación en biología molecular donde se generaron resultados para la escritura de un artículo de investigación, en revista A1.
- Realizar múltiples MOOCs propuestos con el tutor, en pro de aumentar las habilidades en biología computacional y aumentar el entendimiento en estos temas.
- Adquirir y potencializar competencias en el trabajo en equipo, generación de informes y ejecución de labores en el laboratorio.

Dificultades

1.3 Pandemia covid-19

El año 2020 ha sido inusual para todos los procesos formativos, debido a los diferentes aislamientos preventivos impulsados por las entidades gubernamentales, durante los cuales, nos vimos enfrentados a nuevos retos, en el caso específico de la Pasantía en biología computacional y molecular, dentro del grupo de investigación en Ciencias Básicas de la Universidad CES, nos vimos enfrentados a cambiar los cronogramas y los productos propuestos, ya que el trabajo en laboratorio se vio interrumpido durante varios meses, retrasando en gran medida las actividades propuestas.

Resultados

1.4 Detección y modelación de genes de virulencia por medios computacionales en el género *Leptospira*

La realización de las modelaciones para las proteínas fue un proceso que requirió el uso de diferentes herramientas computacionales. Lo primero, fue descargar los genomas de referencia para las 67 especies del género *Leptospira* del Centro Nacional para la Información Biotecnológica o NCBI por sus siglas en inglés. Posteriormente, por medio de la terminal se unieron todos los genomas en un solo archivo FASTA, el cual, se configuró para ser el archivo correspondiente a la base de datos de un BLAST local. Una vez finalizada la base de datos, se procedió a descargar las secuencias de los factores de virulencia detectados in vivo en *Leptospira interrogans* y a realizar el BLAST a cada una de las secuencias, con un límite mínimo de 70% de identidad (**fig1**). Posteriormente, con estas secuencias seleccionadas se realizaron modelaciones por medio del software I-TASSER (6-8) (fig 2-4), y, adicionalmente, las secuencias fueron pasadas por el software interpro (9) obteniendo información sobre la familia protéica a la que pertenece la secuencia enviada en el proceso biológico en el que participa, su función molecular y si se encuentran presentes en algún componente celular. Esto se realizó para cada una de las proteínas que pasaron el umbral de selección.

De los 37 factores de virulencia que se analizaron, sólo 21 de ellos se encontraron en al menos una de las 66 especies restantes, con una identidad mínima del 70 %, habiendo especies que tienen más de una copia de genes específicos en el genoma, como los son *Leptospira noguchii*, con el gen de la hemolisina con identidades del 87% y 73%, además de otras especies con genes que poseen identidad del 100%.

1.5 Realización de PCRs para detección de *Trypanosoma* en sangre y corazón

En la búsqueda de tripanosomatidos en los diferentes órganos de los murciélagos pertenecientes a el macroproyecto de Colciencias realizamos la estandarización de una prueba PCR que amplificara el gen 18s de los tripanosomatidos ya que este gen posee suficiente variabilidad para permitir análisis que nos ayuden a determinar la especie que

infecta a los murciélagos, sin embargo en ninguna de las 204 PCRs realizadas no obtuvimos ningún amplicon de el tamaño esperado de 700pb.

1.6 Realización de PCRs para detección de *Toxoplasma gondii* en riñones

Con la aparición de nueva literatura, se encontraron artículos que describen la identificación de *Toxoplasma gondii* en riñones de murciélagos, aun cuando estos mismos murciélagos aparecen negativos para infección en los órganos como el corazón, similar a la situación presentada en nuestros resultados. Por lo cuál, con esta nueva información se decidió iniciar de nuevo la búsqueda de toxoplasma en las 192 muestras que previamente habían arrojado resultados negativos a la amplificación en ADN extraído de corazón. Con esto, se realizó la estandarización de una PCR *nested* para el gen B1 presente en toxoplasma, utilizando dos set de primers que comparten el mismo tm de 58 °C y generan amplicones de 594pb para el amplicon externo y 530pb para el amplicon interno; hasta la fecha de entrega de este manuscrito no se han concluido las labores de laboratorio para la escritura de un artículo con estos resultados, se estima que para la fecha de sustentación se encuentren finalizados estos análisis aun así podemos observar resultados presuntamente positivos(fig5-6).

1.7 Artículo científico

Hasta el momento de la entrega de este manuscrito no se tienen los resultados totales de los diferentes proyectos realizados en esta pasantía por lo cual no se han escrito los borradores de los dos artículos esperados como resultado de la pasantía, esto debido en gran medida al retraso que presentamos por el covid 19, se espera poder adjuntarlos en la sustentación oral de este trabajo el día 23 de noviembre del 2020.

Conclusiones

El proceso de pasantía es altamente recomendable en la formación de un biólogo, ya que esta le permite desarrollar habilidades técnicas, además de reforzar sus conocimientos teóricos, todo bajo un ambiente laboral real, y este trae consigo nuevos retos, los cuales son algo complicados de presentar en un ambiente académico. Adicionalmente, áreas como la biología molecular, requieren de un buen manejo en elementos del laboratorio, el correcto protocolo para el buen funcionamiento de los equipos, las buenas prácticas de bioseguridad o la asepsia, las cuales son fundamentales para asegurar un resultado confiable y son claves en la hora de los análisis.

Otro aspecto fundamental a la hora de realizar proyectos de investigación es seguir constantemente revisando literatura, de esta forma resultados que aparentemente son negativos pueden volverse a analizar utilizando métodos diferentes que nos arrojen los resultados esperados, el caso de toxoplasma es un claro ejemplo de la necesidad constante de la revisión de literatura y la necesidad de apoyo entre las diferentes instituciones que se dedican a realizar ciencia.

Anexos

Figura 1: Genes sobre los cuales se realiza el blast, NE: Numero de especies en el que se supera el limite de identidad del 70%, Ide Max: Mayor porcentaje de identidad encontrado en el blast, Ide Min: Menor porcentaje de identidad encontrado en el blast.

Gen	NE	Ide Max	Ide Min
alaS	31	89%	76%
clpA	22	94%	72%
cysT	4	91%	75%
hisD	5	93%	76%
lpxB	9	92%	76%
pdxH	5	92%	79%
rpoN	14	92%	74%
LipL32	32	98%	75%
MCE	6	94%	79%
TonB	4	88%	81%
Sphingomyelinase C1-C2	3	88%	72%
Putative lipoprotein	8	91%	81%
OmpL1	2	94%	91%
LipL71	14	93%	77%
LipL45	10	100%	82%
Hemolysin related protein	2	92%	73%
Heme oxygenase	8	92%	81%
Flagellar motor switch protein FliN	16	95%	79%
Flagellar filament sheath protein	15	100%	74%
Microbial collagenase	6	91%	82%
Hemolysin	2	92%	73%

Figura 2,3,4: Modelación computacional de los factores de virulencia ClpA, Hemo-oxigenasa y Hemolisina por medio del software I-TASSER

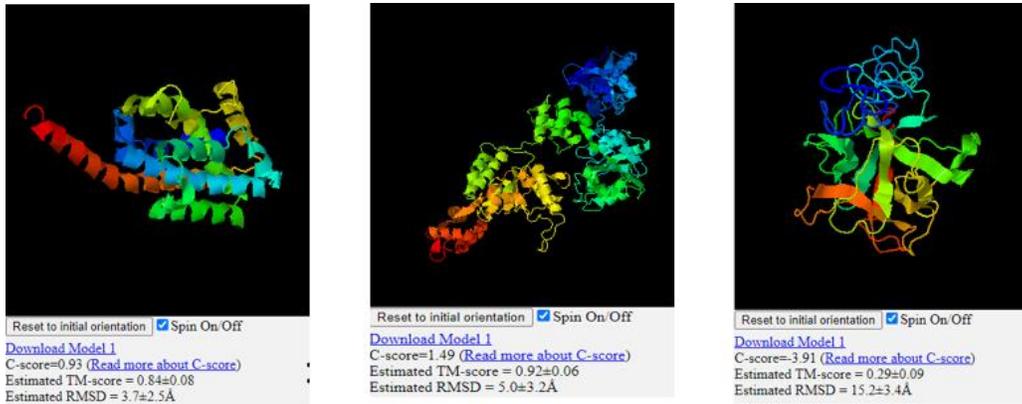


Figura 4: Estandarización de el protocolo para la PCR *nested* en la detección de *Toxoplasma gondii* usando 3 muestras positivas de cultivo

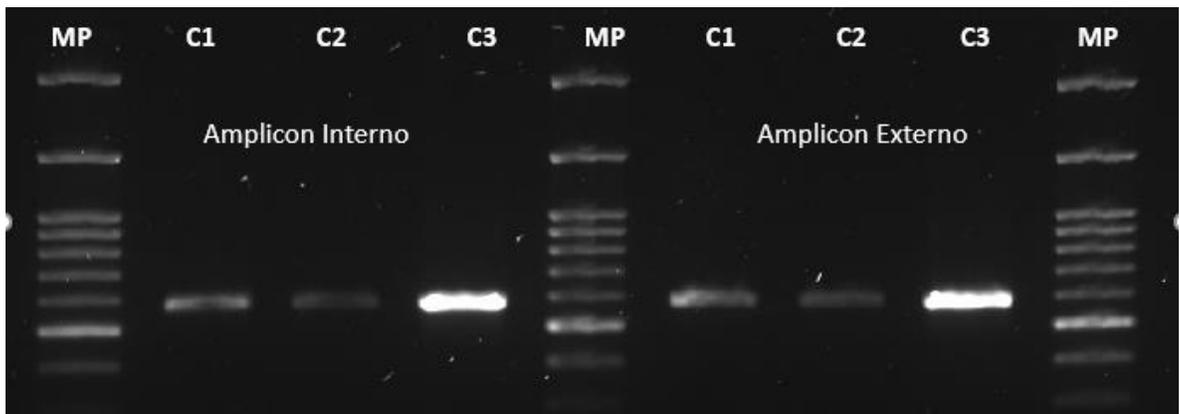
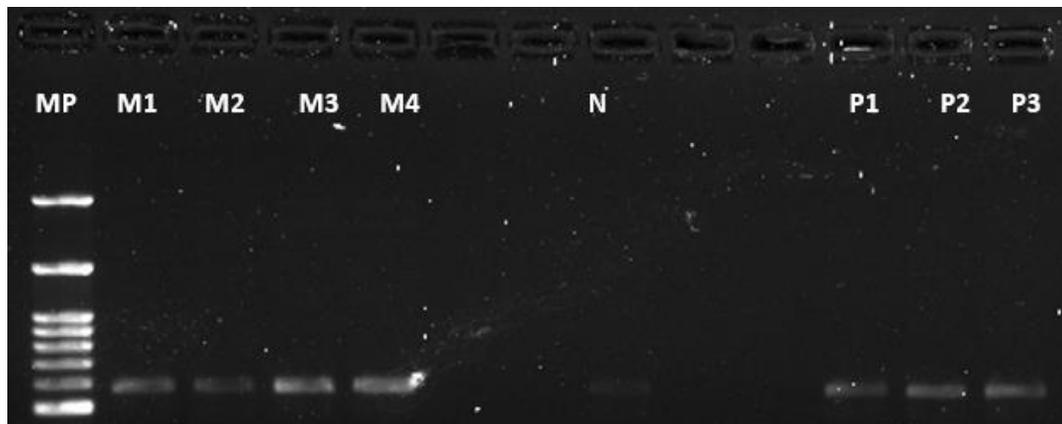


Figura 5: Resultados de cuatro muestras positivas para toxoplasma en riñones



Bibliografía

- Ramos, M. (2019). Sistemas y Computación - UNIANDES - ¿Qué es la Biología Computacional?. [online] Sistemas.uniandes.edu.co. Available at: <https://sistemas.uniandes.edu.co/es/inicio/noticias/item/426-biologiacomputaciona> [Accessed 17 Aug 2019]. (1)
- Rahul Kumar Sharma. Role of Bioinformatics in Various Aspects of Biological Research: A Mini Review. Research and Reviews: Research Journal of Biology (2015).
- Acerca de la Universidad CES - Universidad CES - Medellín. (2019). Retrieved 24 September 2019, from <https://www.ces.edu.co/acerca-de-la-universidad-ces/> (3)
- Grupos de investigación - Universidad CES. (2019). Retrieved 24 September 2019, from <https://www.ces.edu.co/investigacion-e-innovacion/grupos-de-investigacion/> (4)
- GrupLAC - Plataforma SCienti - Colombia. (2019). Retrieved 24 September 2019, from <https://scienti.colciencias.gov.co/gruplac/jsp/visualiza/visualizagr.jsp?nro=00000000003534> (5)
- A Roy, A Kucukural, Y Zhang. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols, 5: 725-738 (2010). (6)
- J Yang, R Yan, A Roy, D Xu, J Poisson, Y Zhang. The I-TASSER Suite: Protein structure and function prediction. Nature Methods, 12: 7-8 (2015). (7)
- .J Yang, Y Zhang. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Research, 43: W174-W181 (2015). (8)
- Mitchell AL, Attwood TK, Babbitt PC, et al. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Research. 2019 Jan;47(D1):D351-D360. DOI: 10.1093/nar/gky1100.(9)