

Vigilancia genómica del virus pandémico SARS-CoV-2

Estudiante
Santiago Toro Casafús

Director(es)
Juan Esteban Gallo Bonilla PhD

Codirector(es)
Isaura Patricia Torres Gómez PhD
Oscar Mauricio Gómez Guzmán M.Sc.

Trabajo de Grado
En la modalidad de *Pasantía*

Programa de Biología
Universidad CES
Medellín
Mayo 2022

Medellín, 21 de junio de 2022

Carta de aprobación del trabajo de grado

Se informa que el estudiante *Santiago Toro Casafús* identificado con cédula: 1152702845 presentó su informe de trabajo de grado en el curso *Seminario de Investigación 3* del semestre 1 del año en curso titulado: "Vigilancia genómica del virus pandémico SARS-CoV-2" y dicho trabajo fue calificado satisfactoriamente en el curso.



Pablo Andrés Guzmán

Docente | Coordinador de Investigación
Biología; Ecología; Maestría en Ciencias Biológicas
Universidad CES

Vigilancia genómica del virus pandémico SARS-CoV-2

Santiago Toro casafús

Resumen

Este informe presenta y describe actividades en el marco de la pasantía desarrollada en Genoma CES por el estudiante Santiago Toro Casafús. Genoma CES es una startup que se dedica a las pruebas diagnósticas, moleculares y genómicas. Se establece como objetivo general, poner en práctica los conocimientos científicos y técnicos adquiridos para optar como profesional en Biología. En la preparación técnica el estudiante estuvo asistiendo diferentes proyectos de investigación y también desarrollando de manera individual sus propios proyectos. Entre los resultados más significativos se destacan: La transformación de células electrocompetentes, purificación de la proteína recombinante Tth, diseño de un marcador de peso molecular y asistencia en el desarrollo del proyecto Vigilancia genómica del virus pandémico SARS-CoV-2. Estos proyectos estuvieron siendo ejecutados bajo la supervisión de Juan Esteban Gallo Bonilla, Isaura Patricia Torres Gómez, Óscar Mauricio Gómez Guzmán y con el acompañamiento de Paula Andrea Pedroza. Con el proyecto Vigilancia genómica del virus pandémico SARS-CoV-2 se logró obtener materiales, reactivos y tecnología que permiten la secuenciación por NGS de aislamientos virales de SARS-CoV-2 en Colombia. Este estudio permitió conocer información genómica valiosa para hacer seguimiento del comportamiento de las cepas de SARS-CoV2 en Colombia y en el mundo.

Palabras clave: ARTIC protocol, COVID-19, Filogenómica, Oxford Nanopore, NGS.

TABLA DE CONTENIDO

1. PRESENTACIÓN.....	5
2. RESEÑA DE LA INSTITUCIÓN.....	5
3. OBJETIVOS.....	5
3.1 OBJETIVO GENERAL.....	5
3.2 OBJETIVOS ESPECÍFICOS.....	6
4. LOGROS ALCANZADOS.....	6
5. RESULTADOS.....	7
5.1 TRANSFORMACIÓN DE CÉLULAS ELECTROCOMOPETENTES <i>E. COLI</i>	7
5.2 PURIFICACIÓN DE LA ENZIMA ADN POLIMERASA THERMUS THERMOPHILLUS.....	9
5.3 DISEÑO Y ESTANDARIZACIÓN DE UN MARCADOR DE PESO MOLECULAR.....	10
5.4 DESARROLLO DE LOS OBJETIVOS DEL PROYECTO VIGILANCIA GENÓMICA DEL VIRUS PANDÉMICO SARS-COV-2.....	12
5.2 CONCLUSIONES.....	13
5.3 BIBLIOGRAFÍA.....	14

1. Presentación

Este informe presenta y describe actividades en el marco de la pasantía que desarrollé en Genoma CES desde el mes de julio hasta finales de noviembre del 2021. Me fue propuesto un rol de asistente de investigación y asistente de producción en los procesos de producción y purificación de enzimas recombinantes.

Esta posibilidad surgió a partir de la necesidad de la institución de contar con un estudiante avanzado en la carrera de biología para desarrollar tareas concernientes en el área de biología molecular y bioinformática.

Este documento se organiza en ocho secciones diferenciadas. Estas secciones se pueden agrupar a su vez en 3 aspectos. Primeramente se describe la reseña de la institución donde fue realizada la pasantía, los objetivos de la pasantía y los logros más significativos que se obtuvieron durante el proceso. En segundo término se encuentran algunos de los resultados de los logros alcanzados, siendo estos los más relevantes para el proceso de formación del estudiante. Por último se encuentran las conclusiones sobre los resultados, las recomendaciones para futuras postulaciones para prácticas similares y la bibliografía.

2. Reseña de la institución

Genoma CES es una startup que se dedica a las pruebas diagnósticas, moleculares y genómicas. Dentro de las pruebas diagnósticas se destacan: ensayos por PCR/PCR en tiempo real, secuenciación Next Generation Sequencing (NGS) de un gen, panel de genes, exoma, exoma trio, genoma completo, genoma mitocondrial, RNA seq, estudios de metilación, estudios prenatales moleculares, estudios oncológicos, estudios farmacogenómicos y estudios de ancestría. Estas pruebas están respaldadas por la Universidad CES y GenomaCES Biotechnologies, pioneros en Colombia en la implementación rutinaria de la medicina genómica y de precisión. Sus instalaciones en Medellín los posicionan como el laboratorio más avanzado del país con uso exclusivo para la implementación de la genómica clínica.

3. Objetivos

3.1 Objetivo general

Se establece como objetivo general, poner en práctica los conocimientos científicos y técnicos adquiridos para optar como profesional en Biología, en el espacio de práctica definido por Genoma CES, dentro de las áreas previamente convenidas entre la Universidad CES, Genoma CES y el estudiante, para su formación integral, que se desarrollan de acuerdo al plan de estudio establecido por la universidad.

3.2 Objetivos específicos

- Revisar bibliografía de temas asignados
- Preparación y presentación de seminarios de revisión de temas
- Entrenamiento y ejecución de técnicas moleculares
- Asistencia en procesos de expresión y purificación de enzimas recombinantes
- Asistencia en la investigación del proyecto titulado “*Vigilancia genómica del virus pandémico SARS-CoV-2*”

4. Logros alcanzados

Previo a la ejecución del proyecto *Vigilancia genómica del virus pandémico SARS-CoV-2*, debí hacer una preparación en conocimientos teóricos como en conocimientos técnicos en el campo de la biología molecular y la bioinformática.

Se realizaron seminarios todos los viernes durante el periodo de la pasantía, en la que se desarrollé los siguientes tópicos: SYBR green VS Eva Green, Fusarium TR4, ¿Cómo hacer un marcador de peso molecular *in house*?, Parámetros de Calidad del ADN, Control de Calidad y Actividad de la enzima Tth polimerasa y adicionalmente se presenté exposiciones de avances semanales.

A partir de los seminarios de Fusarium TR4 se logró generar una propuesta para un proyecto de investigación.

En el área bioinformática, se realizaron capacitaciones semanales todos los miércoles. Estas capacitaciones estuvieron a cargo de la doctora Elizabeth Misas y Óscar Mauricio Gómez.

En la preparación técnica el estuve asistiendo diferentes proyectos de investigación y también desarrollando de manera individual mis propios proyectos. Algunos de estos son: Desarrollo del kit de un qPCR con la enzima Tth como polimerasa, transformación de células electrocompetentes, purificación de la proteína recombinante Tth, diseño de un marcador de peso molecular y asistencia en el desarrollo del proyecto *Vigilancia genómica del virus pandémico SARS-CoV-2*.

5. Resultados

5.1 Transformación de células electrocompetentes *E. coli*

La expresión de proteínas recombinantes en bacterias requiere la inserción de un fragmento de ADN en un vector de expresión, para su posterior transformación en bacterias (1). En la figura 1 se muestran algunas infografías de este proceso. Las células se cultivan y se inducen para expresar la proteína deseada (1).

Los plásmidos pueden llevar uno o más genes de resistencia a los antibióticos, que confieren resistencia a un antibiótico específico, para las bacterias que los portan (2). La presencia de un gen de resistencia a los antibióticos en un plásmido permite el fácil aislamiento de las bacterias que contienen ese plásmido, de las bacterias que no lo contienen por selección artificial (es decir, hacer crecer las bacterias en presencia del antibiótico (2).

La introducción de ADN en bacterias por transformación es un paso esencial en la construcción de cepas recombinantes (3). Recientemente, la electroporación, o electropermeabilización, en la que se usa una breve descarga eléctrica de alto voltaje para hacer que las células sean permeables al ADN, ha revolucionado la transformación de las bacterias (3). La técnica también puede ofrecer ventajas para bacterias transformables como *E. coli* (3).

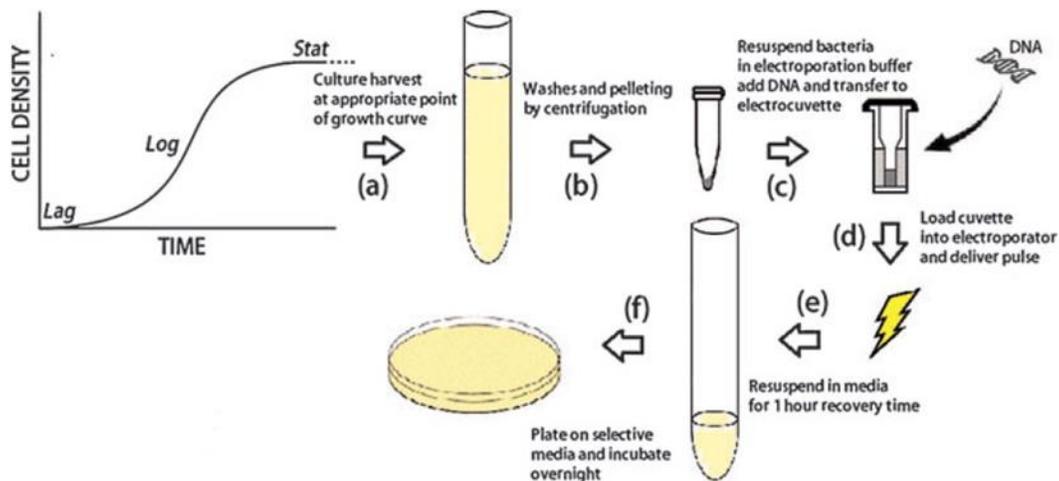


Figura 1. Proceso de electroporación. Provenzano D., Trevino V., Ermolinsky B. (2018)

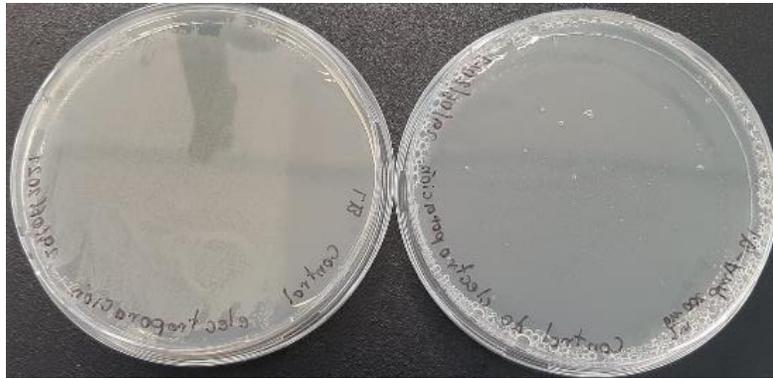


Figura 2: Control de electroporación.

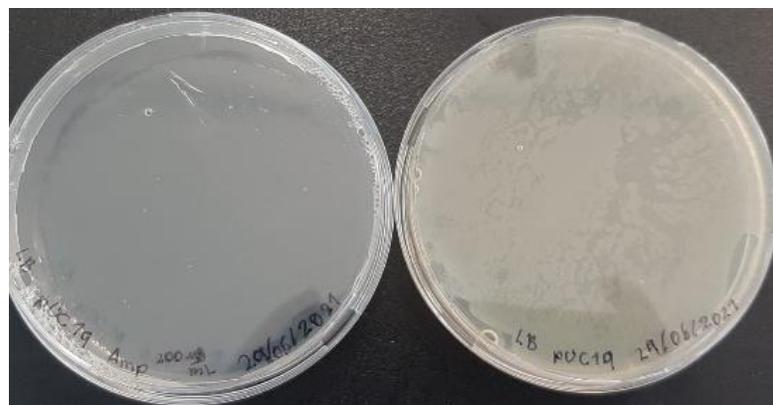


Figura 3: Control negativo y positivo.



Figura 4: Resultados del experimento.

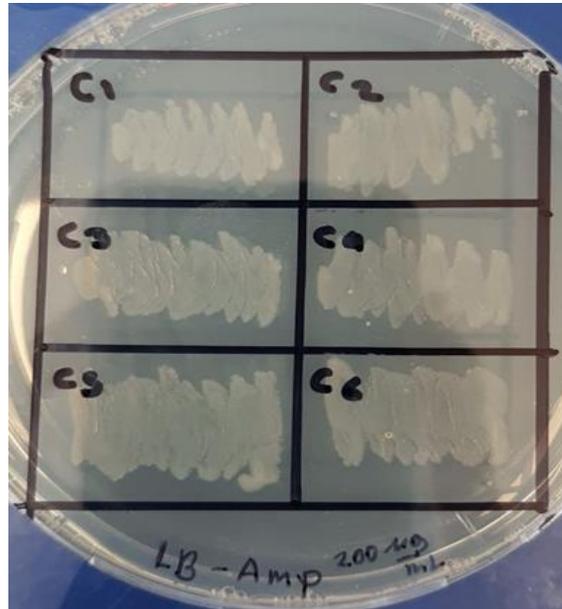


Figura 5: Aislamiento de los clones.

5.2 Purificación de la enzima ADN polimerasa *Thermus*

thermophilus

La expresión de proteínas recombinantes en bacterias requiere la inserción de un fragmento de ADN en un vector de expresión, para su posterior transformación en bacterias (1). Las células se cultivan y se inducen para expresar la proteína deseada (1).

El sistema de expresión *E. coli* no produce únicamente la proteína de interés Tth polimerasa, sino también otras ciertas de miles de proteínas que constituyen su función metabólica normal como se muestra en la figura 5 los resultados en un SDS-PAGE después de purificar un lote de Tth polimerasa. Además de que en este residen moléculas de ADN y ARN, entre otras. Por consiguiente, se hace necesaria la purificación de la proteína Tth polimerasa, evitando que otras sustancias puedan interferir con su actividad.

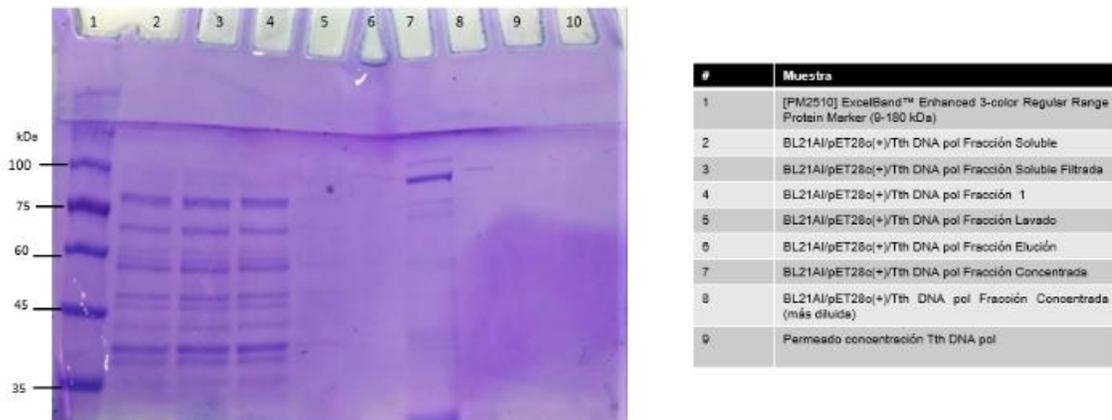


Figura 5: Resultados del proceso de purificación de un lote de Tth pol. En el pozo número 7 se evidencia una banda de un color azul oscuro. Esto indica que está fracción se encuentra en una mayor concentración a las demás proteínas presentes.

5.3 Diseño y estandarización de un marcador de peso molecular

La determinación del peso molecular (mw) o la longitud del par de bases (pb) de los ácidos nucleicos es una urgencia esencial en la biología molecular. Esta necesidad comprende una amplia gama de pesos o longitudes de ácidos nucleicos que van desde mega pb hasta oligonucleótidos muy cortos de solo unos pocos pb (4). Los estándares de mw de ácido nucleico son herramientas útiles para estimar la calidad, el tamaño y la cantidad de la muestra de ácido nucleico (4).

El propósito de este diseño y estandarización fue desarrollar un mw rentable para la producción de una escalera de ADN que sea aplicable a los procesos de rutina en el laboratorio.

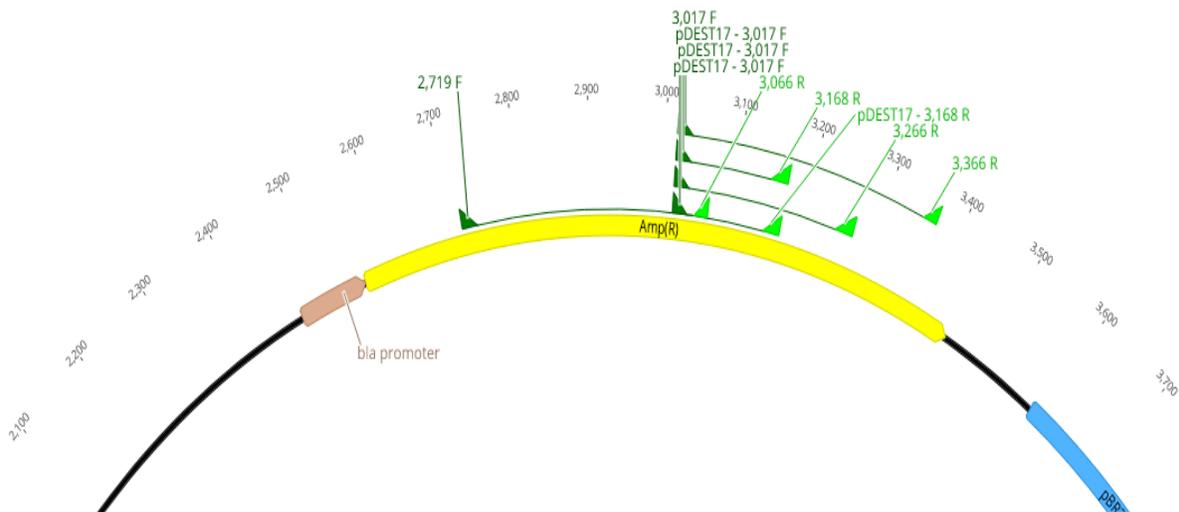


Figura 6: Diseño *in silico* de los amplicones para la construcción del mw en un plásmido.

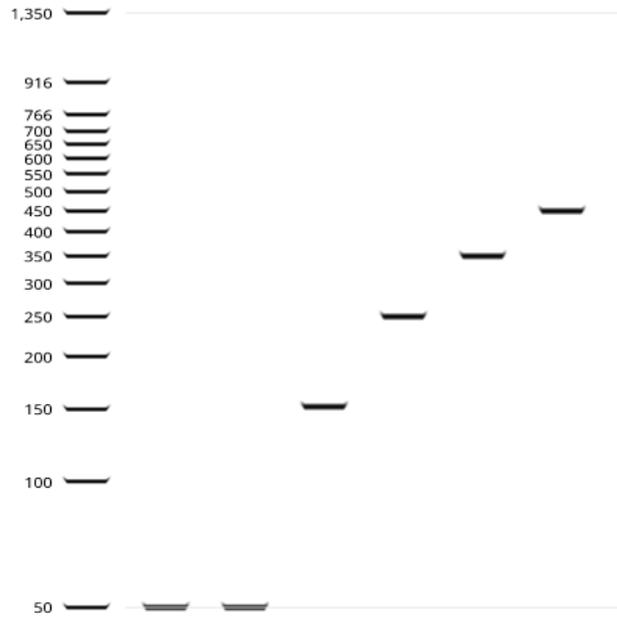


Figura 7: Resultados esperados para un gel de agarosa después de haber aplicado la técnica de electroforesis.

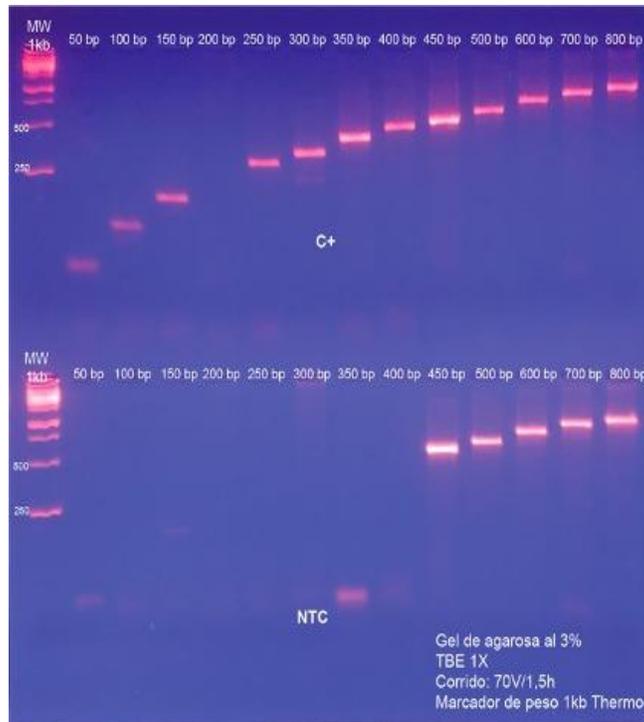


Figura 8: Resultados reales posteriores a la estandarización de las concentraciones de los amplicones y revelado en un gel de agarosa. Nótese la escala de 50pb – 450pb.

5.4 Desarrollo de los objetivos del proyecto *Vigilancia genómica del virus pandémico SARS-CoV-2*

La taxonomía de virus es la clasificación de virus en categorías llamadas taxones y el desarrollo e implementación de una nomenclatura estandarizada para taxones. Mientras que la taxonomía de virus ha sido una subespecialidad de nicho durante muchas décadas, el campo ha cobrado importancia recientemente debido al número exponencialmente creciente de nuevos virus descubiertos por métodos de secuenciación de próxima generación (5).

SARS-CoV-2 es un miembro del orden Nidovirales, familia Coronaviridae, subfamilia Orthcoronavirinae, el cual se subdivide en 4 géneros *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* y *Deltacoronavirus*. El género Alpha y Beta se originaron en los murciélagos, mientras que los Gamma y Delta han evolucionado de las aves y reservas genéticas porcinas

Basados en su caracterización molecular, SARS-CoV-2 es considerado como un nuevo *Betacoronavirus* perteneciente al subgénero *Sarbecovirus*. Los genomas obtenidos de pacientes infectados durante la pandemia han sido analizados, mostrando un 79.5% de similitud en su secuencia con la de SARS-CoV-2.

Para secuenciar el SARS-CoV-2, se ha establecido un protocolo con el cual se amplifica el material genético del virus a partir de muestras clínicas, conocido como el protocolo ARTIC. Este protocolo utiliza un conjunto de iniciadores específicos de SARS-CoV, los cuales amplifican fragmentos del genoma de SARS-CoV-2 los cuales son secuenciados de forma eficiente en el equipo menos complejos y más económico que ofrece la tecnología nanopore, el MinION. Con el MinION y el protocolo ARTIC se pueden obtener hasta 50 genomas de SARS-CoV-2 por corrida (6).

Para el desarrollar el presente proyecto se requirieron las siguientes actividades:

Evaluar el comportamiento epidemiológico y clínico de aislamientos de SARS-CoV-2 utilizando tecnología de secuenciación de nueva generación, en una cohorte de pacientes del estudio PC-COVID-19 utilizando tecnología Oxford Nanopore con el protocolo ARTIC.

En total se obtuvieron 46 genomas, 17 de la ciudad de Medellín y 29 de la ciudad de Bogotá. Esta información se puede ver ampliada en la tabla de Resultados-copia en los archivos anexos.

Posterior a la secuenciación y ensamblaje de los genomas virales se evaluarón los linajes específicos de cada aislamiento de SARS-CoV-2 secuenciado comparandolo con un lote de 1513 genomas de SARS-CoV-2 de la base de datos de GISAID. Se identificaron un total de 7 linajes diferentes. A continuación se diferencian por localidad.

Linjae	Localidad
B	Medellín
B.1	Medellín y Bogotá
B.1.111	Medellín y Bogotá
B.1.113	Bogotá
B.1.420	Bogotá
B.1.466	Bogotá
B.1.579	Bogotá

Tabla 1: Linajes del virus SARS-CoV-2 identificados en la localidad de Medellín y Bogotá.

6. Conclusiones

En este trabajo se hace una mención en la sección de resultados de 4 proyectos en los que estuve participando de manera activa. Estos proyectos estuvieron siendo ejecutados bajo la supervisión de Juan Esteban Gallo Bonilla, Isaura Patricia Torres Gómez, Óscar Mauricio Gómez Guzmán y con el acompañamiento de Paula Andrea Pedroza.

El proyecto de transformación de células electrocompetentes se realizó de manera exitosa, se pudo generar un banco de células electrocompetentes y transformar genéticamente una cepa wildtype de *E. coli* con un plasmido de interés para obtener un clon de la cepa productora o propagadora de la proteína de interés. Como se aprecia en la figura 5 las cepas transformadas pudieron crecer en presencia del antibiótico (ampicilina). Posteriormente pasaron a un proceso de plateado y almacenado.

Tras el proceso de purificación de la Tth podemos observar que el experimento no fue exitoso en su totalidad. Esto se puede evidenciar porque en el pozo 7 del SDS-PAGE solo debía observarse una banda a la altura de 100 kDa. Esto quiere decir que efectivamente tenemos proteína concentrada pero no está completamente purificada. Sin embargo a pesar de las impurezas sigue siendo un lote utilizable y funcional para los experimentos de prueba del laboratorio pero no para su comercialización.

El análisis holístico de diferentes parámetros en el diseño de cebadores dio como resultado la amplificación de fragmentos en un solo programa de PCR sin subproductos ni pasos de

purificación. Esto disminuye los costos al eliminar la necesidad de comprar mw de marcas comerciales y los tiempos de demora que están sujetos a los envíos de los mismos.

Con el proyecto *Vigilancia genómica del virus pandémico SARS-CoV-2* se logró obtener materiales, reactivos y tecnología que permiten la secuenciación por NGS de aislamientos virales de SARS-CoV-2 en Colombia. Este estudio permitió conocer información genómica valiosa para hacer seguimiento del comportamiento de las cepas de SARS-CoV2 en Colombia y en el mundo.

7. Recomendaciones

Se recomienda en el caso de los aspirantes al título de biología y que a su vez estén interesados en la línea de profundización de bioinformática, que cuenten con un conocimiento previo y profundo en las dos técnicas moleculares que son la base del conocimiento en bioinformática estas son la end point PCR y qPCR. A su vez, también se recomienda que los estudiantes que hayan sido formados en el manejo básico de sistemas operativos Linux, terminal y línea de comandos.

8. Bibliografía

1. Langlais C, Korn B. Recombinant Protein Expression in Bacteria. En: Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine [Internet]. Springer Berlin Heidelberg; 2006 [citado 9 de mayo de 2022]. p. 1609-16. Disponible en: http://link.springer.com/10.1007/3-540-29623-9_4800
2. Addgene: Protocol - How to Inoculate a Bacterial Culture [Internet]. [citado 9 de mayo de 2022]. Disponible en: <https://www.addgene.org/protocols/inoculate-bacterial-culture/>
3. Mercenier A, Chassy BM. Strategies for the development of bacterial transformation systems. *Biochimie*. abril de 1988;70(4):503-17.
4. Mofid M, Abbasian M, Seyedi HA, Boroujeni Z. Easy method for production of a home-made DNA ladder in every laboratory. *Adv Biomed Res*. 2015;4(1):70.
5. Kuhn JH. Virus Taxonomy. En: Encyclopedia of Virology [Internet]. Elsevier; 2021 [citado 9 de mayo de 2022]. p. 28-37. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/B9780128096338212314>
6. Itokawa K, Sekizuka T, Hashino M, Tanaka R, Kuroda M. Disentangling primer interactions improves SARS-CoV-2 genome sequencing by the ARTIC Network's multiplex PCR [Internet]. *Microbiology*; 2020 mar [citado 9 de mayo de 2022]. Disponible en: <http://biorxiv.org/lookup/doi/10.1101/2020.03.10.985150>

