

# Integración de herramientas bioinformáticas y métodos en biología molecular para el diseño de un kit diagnóstico del COVID-19: un ejemplo de aprendizaje significativo

## Integration of Bioinformatics Tools and Molecular Biology Methods for the design of a COVID-19 Test Kit: An Example of Significant Learning

**Catalina Rivera Forero**  
**Luis Alejandro Yáñez Dukon**  
**Carolina Herrera Khenayzir**  
**Juan Carlos Arias**  
**Jonathan Niño Vargas**

Estudiantes de Biología Marina y Biología Ambiental, Departamento de Ciencias Naturales y Ambientales, Facultad de Ciencias e Ingeniería, Grupo de Investigación Genética, Biología Molecular y Bioinformática, GENBIMOL, Universidad de Bogotá Jorge Tadeo Lozano, Colombia

**Pilar Rodríguez Becerra**

Docente, Departamento de Ciencias Naturales y Ambientales, Facultad de Ciencias e Ingeniería, Grupo de Investigación Genética, Biología Molecular y Bioinformática, GENBIMOL, Universidad de Bogotá Jorge Tadeo Lozano, Colombia  
<https://orcid.org/0000-0003-3843-1490>

**Leslie Leal Mejía**

Docente, Departamento de Ciencias Naturales y Ambientales, Facultad de Ciencias e Ingeniería, Grupo de Investigación Genética, Biología Molecular y Bioinformática, GENBIMOL, Universidad de Bogotá Jorge Tadeo Lozano, Colombia  
<https://orcid.org/0000-0001-7940-0020>

**Javier Hernández Fernández**

Docente, Departamento de Ciencias Naturales y Ambientales, Facultad de Ciencias e Ingeniería, Grupo de Investigación Genética, Biología Molecular y Bioinformática, GENBIMOL, Universidad de Bogotá Jorge Tadeo Lozano, Colombia  
<https://orcid.org/0000-0001-8442-9266>  
[javier.hernandez@utadeo.edu.co](mailto:javier.hernandez@utadeo.edu.co)

**Fecha de recepción: 10 de marzo de 2020**  
**Fecha de aceptación: 30 de marzo de 2020**

**Sugerencia de citación:** Rivera-Forero, C., Yáñez-Dukon, L. A., Herrera-Khenayzir, C., Arias, J. K., Niño-Vargas, J., Rodríguez-Becerra, P. ... y Hernández-Fernández, J. (2019). Integración de herramientas bioinformáticas y métodos en biología molecular para el diseño de un kit diagnóstico del COVID-19: un ejemplo de aprendizaje significativo.

*Mutis*, 9(2), 62-80. | [doi: 10.21789/22561498.1599](https://doi.org/10.21789/22561498.1599)

## RESUMEN

Como una forma pedagógica de aprendizaje activo, se relacionaron temas desarrollados en clase de biología molecular con la pandemia del coronavirus COVID-19 que en este momento avanza en casi la totalidad del planeta. Contar con un método eficiente para diagnosticar la infección durante el inicio de esta infección respiratoria aguda es importante para detectar tanto los pacientes sintomáticos como los asintomáticos. Por este motivo, se retó a los estudiantes de séptimo semestre de las carreras de biología marina y ambiental que cursan la materia biología molecular en la Universidad de Bogotá Jorge Tadeo Lozano a diseñar un método diagnóstico para el virus COVID-19 basado en RT PCR y qPCR. El objetivo de este ejercicio académico consistió en incentivar a los estudiantes para que propusieran un método efectivo que permita detectar el virus COVID-19 a través de los conocimientos adquiridos durante el curso. Se realizó una revisión de la literatura previa para identificar los métodos existentes para la detección del virus y con ayuda de herramientas bioinformáticas se realizó el análisis de las secuencias de los genomas del COVID-19, el SARS humano y el SARS de murciélago disponibles en GenBank.

**Palabras claves:** RT-qPCR, SARS COV 2, pandemia, coronavirus, educación contextualizada.

## ABSTRACT

As a pedagogical form of active learning, topics developed in molecular biology class were related to the coronavirus COVID-19 pandemic that is currently advancing in almost the entire planet. Having an efficient method of diagnosing infection during the onset of this acute respiratory infection is important in detecting both symptomatic and asymptomatic patients. For this reason, seventh semester students from marine and environmental biology courses studying molecular biology at the University of Bogotá Jorge Tadeo Lozano were challenged to design a diagnostic method for the COVID-19 virus based on RT PCR and qPCR. The objective of this academic exercise was to encourage students to propose an effective method to detect the COVID-19 virus using the knowledge acquired during the course. A review of the previous literature was carried out to identify the existing methods for the detection of the virus and, with the help of bioinformatic tools, the analysis of the sequences of the genomes of COVID-19, human SARS and bat SARS available at GenBank was carried out.

**Keywords:** RT-qPCR, SARS COV 2, pandemic, coronavirus, contextualized education.

## INTRODUCCIÓN

La biología molecular y la genética en la sociedad del conocimiento del siglo XXI se mezclan cada vez más con disciplinas tradicionalmente separadas pero relacionadas, como la física, la informática y la matemática (Furge, Stevens-Truss, Moore, & Langeland, 2009). El proyecto genoma humano, iniciado en 1990, produjo el desarrollo de una ciencia sin la cual la gran cantidad de datos del genoma no hubiera podido ser analizada: la bioinformática. Esta ciencia –que cada día toma más fuerza–, unida al uso de métodos computacionales, puede resolver problemas biológicos

que requieren nuevos enfoques. Un plan de estudio bien estructurado e integral en biología molecular debe incluir esta ciencia. De otra parte, el aprendizaje basado en problemas, orientado a la investigación y el desarrollado en equipos, y dirigido de manera adecuada por profesores, se convierte en una estrategia centrada en el estudiante que apoyan el aprendizaje activo y que son muy efectivas en la educación científica (Prieto, 2006).

La bioinformática se refiere a: “La investigación, desarrollo o aplicación de herramientas y enfoques computacionales para expandir el uso de datos

biológicos, médicos, conductuales o de salud, incluidos aquellos para adquirir, almacenar, organizar, archivar, analizar o visualizar datos” (BISTIC, 2000). Teniendo esta definición en mente, y siguiendo nuestro programa durante la semana siete de clases, se desarrolló el tema de métodos en biología molecular, específicamente los métodos de RT-PCR, PCR tiempo real, sondas, hibridaciones de Southern, Northern y Western, así como la utilización de vectores (plásmidos, cósmidos, fásmidos, cromosomas bacterianos y cromosomas de levaduras) para la clonación de genes. Paralelamente, se realizaron aplicaciones bioinformáticas para aprender a descargar secuencias de GenBank y utilizar programas para el corte con enzimas de restricción [Nebcutter V2.0, <http://nc2.neb.com/NEBcutter2/>] (Vincze, Posfai, & Roberts, 2003) y construir los perfiles electroforéticos. Adicionalmente, los estudiantes aprendieron a realizar alineamientos de secuencias de ADN utilizando ClustalW [<https://www.genome.jp/tools-bin/clustalw>] (Li, 2003), diseñar primer con Primer3web versión 4.1.0 [<http://primer3.ut.ee/>] (Untergasser *et al.*, 2012), hacer PCR virtuales con iPCR [[https://embnet.vital-it.ch/software/iPCR\\_form.html](https://embnet.vital-it.ch/software/iPCR_form.html)] (Lexa, Horak, & Brzobohaty, 2001) y, de manera destacada, a manejar la herramienta BLASTn de GenBank para la búsqueda de secuencias homólogas [<https://blast.ncbi.nlm.nih.gov/Blast.cgi>] (Altschul, Gish, Miller, Myers, & Lipman, 1990); este último programa además compara secuencias de nucleótidos o proteínas con bases de datos de secuencias y calcula la significación estadística (GenBank, 2020). Teniendo en cuenta lo anterior, se puede afirmar que los estudiantes se relacionaron con las más actuales y modernas herramientas para el aprendizaje de análisis de secuencias proteicas y de ácidos nucleicos.

Históricamente, la educación en ciencia ha requerido desarrollar un enfoque que propicie enseñanza contextualizada o en contexto del mundo real, con lo cual el aprendizaje se convierte en altamente significativo. Así, cuando la nueva información se relaciona con algún aspecto de lo ya existente en la estructura cognitiva del individuo, se produce un proceso que conduce al aprendizaje significativo (Novak, 1988). A partir de este argumento, se introdujo a los estudiantes en el conocimiento de la pandemia producida por el nuevo coronavirus (ncov 19) descubierto a finales de 2019 en Wuhan, provincia de Hubei, China (Wu *et al.*, 2020). Este coronavirus causa una infección severa

en las vías respiratorias. Desde su aparición a principios de diciembre de 2019 hasta el 26 de marzo de 2020, el número de casos ha aumentado exponencialmente hasta 858.785, con un total de 42.151 muertes y una recuperación de 178.119 pacientes (Dong, Du, & Gardner, 2020). La pandemia se ha generalizado a 180 países, donde las naciones con mayor número de infectados son China (82.290), Italia (105.792), Estados Unidos (189.035) y España (95.923). Colombia no ha sido ajeno a la pandemia, registrando su primer caso de contagio el 3 de marzo de 2020. Actualmente, el país cuenta con 906 casos reportados (Dong *et al.*, 2020).

Dentro de las políticas que han tomado la mayoría de los países que presentan individuos infectados por COVID-19 se encuentran: el aislamiento de los enfermos, la cuarentena de personas que posiblemente han estado expuestas al virus, el confinamiento de todos sus habitantes para prevenir nuevos contagios y la realización de un mayor número de pruebas diagnósticas. Una de estas pruebas se basa en la RT-qPCR, técnica que permite hacer una copia de DNA complementario a partir del RNA viral utilizando una enzima denominada transcriptasa inversa, la cual convierte el RNA de cadena sencilla en DNA de doble cadena. Con esta enzima es posible amplificar diferentes regiones del virus para posibilitar su caracterización. Para ello, se usan primers, o cadenas sencillas cortas de ADN (20 nucleótidos), y sondas de DNA marcadas con fluorocromos para identificar en tiempo real la amplificación de los fragmentos.

La PCR en tiempo real (o cuantitativa) permite medir la cantidad de copias del DNA viral en la muestra. Esta técnica es bastante confiable para identificar los virus. Sin embargo, cuenta con algunas limitaciones. Una de ellas es el tiempo necesario para su ejecución (4-6 h), lo que limita el número de pruebas que se pueden realizar al día. Otra limitante son los reactivos o kits para desarrollar las pruebas, puesto que en este momento se reporta escasez de estos elementos debido a la gran demanda a escala global. Además, la contaminación es un problema limitante, puesto que se requieren laboratorios con suficiente bioseguridad a fin de evitar que se presenten falsos positivos y falsos negativos, como ocurre en algunas ocasiones.

Por otra parte, estas pruebas permiten detectar pacientes asintomáticos, configurando un diagnóstico de gran importancia debido a que estos son transmisores

activos de la infección (Costa, 2004; Salazar, Sandoval, & Armendáriz, 2013).

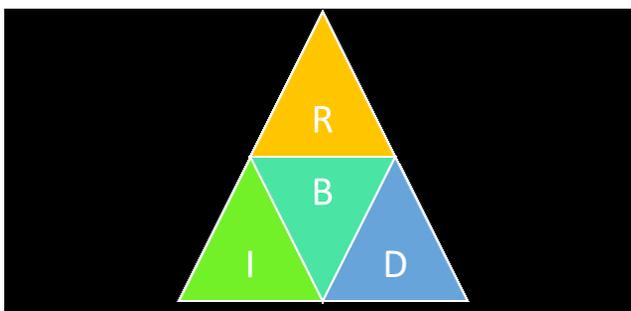
A partir de lo anterior, este trabajo tuvo como objetivo diseñar una prueba diagnóstica con los estudiantes del curso de biología molecular 2020-1 de la Universidad Jorge Tadeo Lozano, quienes trabajaron muy seriamente y propusieron las bases para desarrollar una prueba que podría convertirse en kit para identificar el COVID-19.

## MÉTODOS

### Revisión de literatura

Los estudiantes consultaron literatura relacionada con las técnicas moleculares actualmente utilizadas para la identificación de virus de RNA en las bases de datos ScienceDirect, Scopus, GenBank y Elsevier. Con base en la información recopilada, se planteó una metodología alternativa para la detección del SARS-CoV2. La figura 1 presenta la propuesta metodológica para el desarrollo de este estudio.

**Figura 1.** Esquema metodológico utilizado para el desarrollar un método alternativo para detectar el covid-19



Nota: R) reconocimiento de métodos moleculares y herramientas bioinformáticas para detección y análisis de virus de RNA; B) búsqueda de genomas de virus de interés; I) identificación de genes candidatos; D) diseño de primers y sondas.

Fuente: elaboración propia.

### Identificación de genes

Se recuperaron de la base de datos de GenBank (NCBI, 2020) los genomas completos de tres cepas

de coronavirus correspondientes a Wuhan nCoV-19 (NC\_045512.2), Human SARS (NC\_004718.3) y Bat SARS (DQ412043.1). Las secuencias de cada uno de los 11 genes que componen al virus Wuhan nCoV-19 fueron alineadas al genoma de las otras dos especies con ClustalW. Adicionalmente, se seleccionaron los genes que presentaron menor puntaje contra regiones del genoma del Human SARS y Bat SARS como posibles candidatos para la identificación molecular de COVID-19.

### Diseño de primers

Utilizando la secuencia de los genes candidatos identificados previamente mediante ClustalW que presentaron un score inferior a 70 %, se diseñaron primers con el programa Primer3 versión 4.1.0, empleando los parámetros %GC (40 < %GC < 60) y la probabilidad de *hair pins* (horquillas).

### Diseño de sondas

Se diseñaron sondas tipo TagMan a partir de unas regiones de 20 y 24 pb al interior de cada uno de los genes predichos como posibles marcadores moleculares. Se utilizó el programa MegaX para extraer la secuencia *reverse complement* de las regiones de cada uno de los genes, las cuales, por complementariedad, se hibridarán a los genes que se quiere identificar. A las sondas creadas se les incorporará dos fluorocromos (un aceptor y un donador).

### RT-qPCR

Para el diagnóstico del virus COVID-19 se realizará una RT-qPCR, la cual, inicialmente, sintetizará una cadena de cDNA a partir del RNA viral, para posteriormente amplificar y cuantificar en tiempo real la expresión de genes que permitirán su identificación de forma diferencial (Pinilla, 2019).

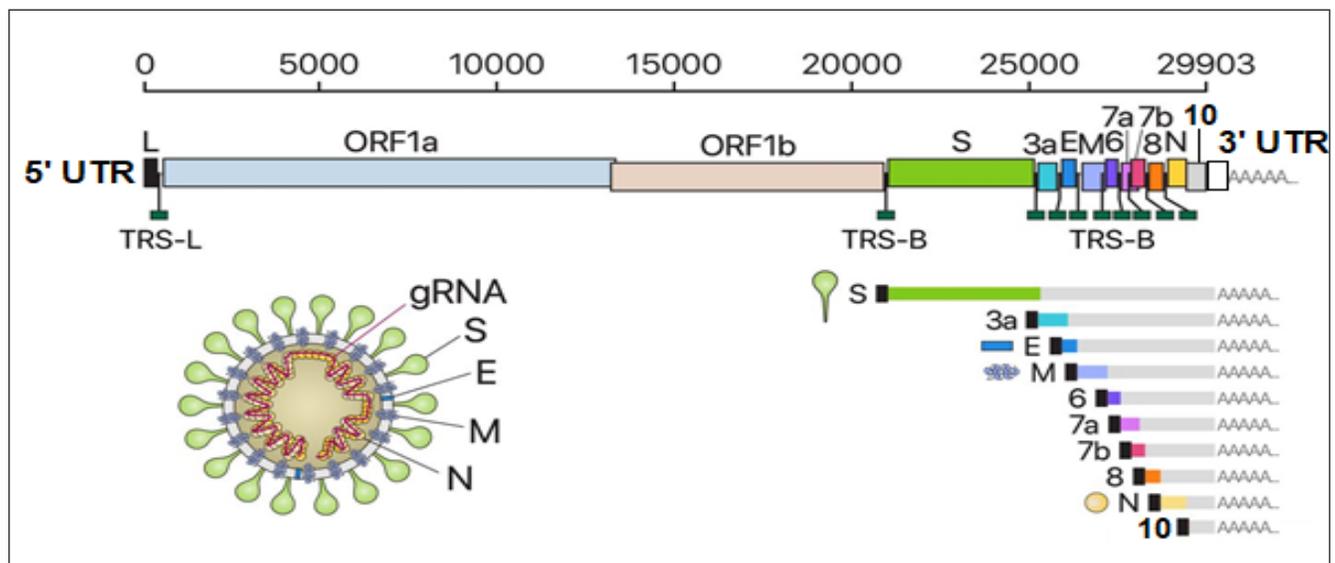
## RESULTADOS

El SARS-CoV-2 es un betacoronavirus responsable de la pandemia de COVID-19. Este posee un transcriptoma altamente complejo debido a numerosos eventos de recombinación, tanto canónicos como no canónicos. Su genoma está compuesto por 11 genes con

longitudes entre 114 y 3.822 pb y comprende cerca de 30.000 pb, distribuido de la siguiente manera: en el extremo 3' una cápside, en el extremo 5' una cola poliA, en su interior dos ORFs 1a y 1b, los cuales se traducen en 46 proteínas una vez colonizan la célula,

3 regiones de ARN que codifican para proteínas estructurales ("S" proteína espiga "spike protein", "M" proteína de membrana, "N" proteína de la cápside) y 6 proteínas accesorias (3a, 6, 7a, 7b, 8 y 10), como se muestra en la figura 2 (Kim *et al.*, 2020).

**Figura 2.** Presentación esquemática de la organización del genoma del SARS-CoV-2 los ARNm subgenómicos canónicos y la estructura del virión



Nota: las cajas grandes representan los ORFs, las cajas pequeñas las proteínas accesorias y el cuadro negro indica la secuencia líder.

Fuente: adaptado de Kim *et al.* (2020).

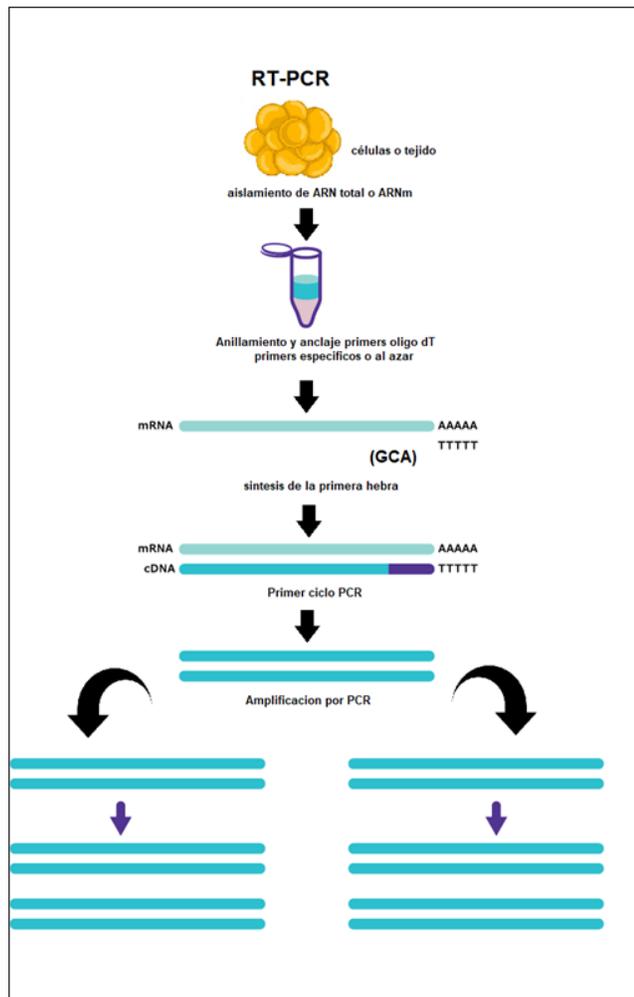
La reacción en cadena de la polimerasa en tiempo real (RT-qPCR) es una metodología sensible que se utiliza ampliamente en diagnóstico e investigación biológica. Su mayor virtud es poder amplificar fragmentos pequeños a partir de mínimas cantidades de DNA del cual se quiere identificar presencia en una muestra. En diagnóstico clínico, la RT-qPCR es utilizada para medir las cargas virales o bacterianas. Por esta razón, los autores consideran que este método es el mejor, más rápido y económico para hacer diagnóstico del coronavirus COVID-19. Sin embargo, se requiere tener alguna experiencia en el manejo de las herramientas bioinformáticas para reconocer los genes o regiones a amplificar, el diseño de primers y el diseño de sondas. La RT-qPCR es utilizada para mejorar la eficiencia diagnóstica como método rápido y temprano para identificar el agente etiológico (en este caso, el coronavirus COVID-19) y

de esta manera tomar las medidas de prevención y control del contagio y la enfermedad en fases tempranas con el objetivo de evitar su dispersión (Corman *et al.*, 2020).

El COVID-19 es un virus cuyo material genético es RNA de cadena sencilla. Por ello, se presenta un inconveniente, pues la PCR amplifica únicamente fragmentos de cadena doble, es decir, DNA. Sin embargo, métodos como la RT-PCR (reacción en cadena de la polimerasa con transcriptasa inversa o reversa) representan una solución para estos casos debido a que permiten convertir una cadena sencilla de RNA en una doble cadena de DNA complementaria (cdNA) con una alta fidelidad (Costa, 2004). En la figura 3 se presenta el proceso mediante el cual, empleando un kit de aislamiento para RNA total, se obtiene el material genético del virus. Posteriormente, utilizando

un primer de TTTTs, se origina la primera cadena de cDNA y luego, por PCR, se obtienen cadenas dobles del cDNA del virus (figura 3).

**Figura 3.** Proceso de RT-PCR por el cual se obtiene una cadena doble de DNA complementario

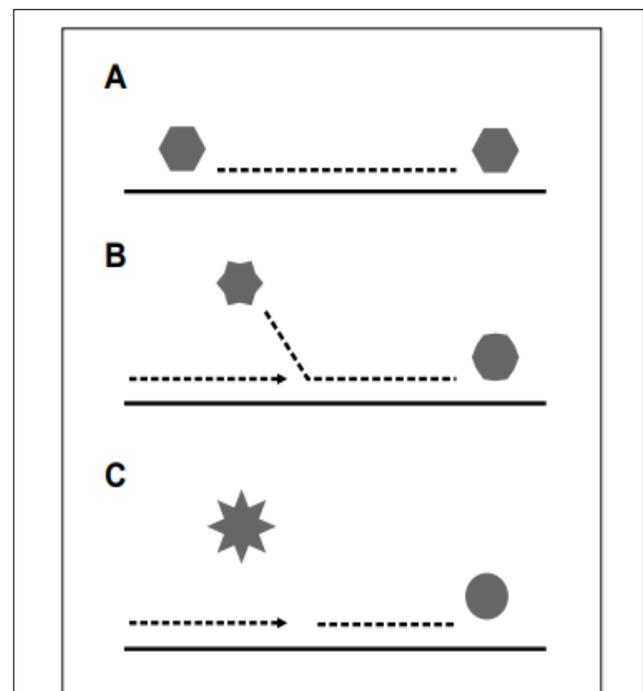


Fuente: adaptada de <https://www.sigmaaldrich.com/life-science/molecular-biology/pcr/rt-pcr.html>.

Al tener la cadena de cDNA se prosigue con la amplificación de uno o varios fragmentos de esta molécula. El COVID-19 tiene un RNA de 29,903 pb (figura 1). La qPCR es un método que amplifica regiones o fragmentos de ADN y los detecta de manera paralela dentro del mismo microtubo utilizado. Por ello, se utilizan sondas, que son fragmentos marcados de ADN de cadena sencilla complementarios a las zonas que se

van a amplificar, de manera que al unirse a los fragmentos amplificados permite identificarlos y cuantificarlos. Los autores del presente trabajo seleccionaron sondas de hidrólisis (Taqman) que corresponden a fragmentos de ADN de cadena sencilla marcados en el extremo 3' con un fluorocromo donador que reporta la fluorescencia, emitiendo luz al ser chocado por la ADN polimerasa, y un aceptor en el extremo 5', llamado quencher, el cual atrapa la fluorescencia que libera el donador. Las moléculas donadora y aceptora deben estar cercanas, de manera que cuando la sonda está intacta, la fluorescencia emitida por el donador no es observable, al ser absorbida por el aceptor (figura 4A). Durante la amplificación, la Taq polimerasa de ADN rompe o libera el extremo libre 5' de la sonda que contiene al donador al iniciar la síntesis, permitiéndole emitir la fluorescencia (figura 4B), que es captada por el lector (figura 4C) (Costa, 2004).

**Figura 4.** Proceso e activación de la fluorescencia en la qPCR



Nota: A) la sonda se hibrida a la zona complementaria del fragmento que se va a amplificar, el extremo donador está cerca al quencher, no hay fluorescencia; B) la polimerasa que ha iniciado la síntesis de DNA choca contra el donador y lo libera, produciéndose fluorescencia; C) la fluorescencia es captada y transmitida por el lector.

Fuente: tomada de Costa (2004).

### Identificación de genes candidatos para detección del virus COVID-19

Las secuencias de los genomas de Wuhan nCOV-19, Human SARS y Bat SARS descargados de GenBank presentaron longitudes de 29.903, 29.751 y 29.749 pb, respectivamente. El virus Wuhan- nCOVID-19,

estructuralmente está compuesto por 11 genes distribuidos como se presenta en la tabla 1, con longitudes entre 114 y 3.822 pb. A diferencia del SARS humano, el SARS aislado en murciélagos es 0,5 % más largo y su composición está bien definida a nivel de UTRs, CDs y proteínas.

**Tabla 1.** Composición estructural de los genes del virus Wuhan n-COVID-19, Human SARS y Bat SARS utilizados en este estudio

21289 pb	3822	828	228	669	186	366	132	366	1280	114		
ORF 1	Gen S	ORF 3a	Gen E	Gen M	ORF 6	ORF 7a	ORF 7b	ORF 8	Gen N	ORF 10		
1	2	3	4	5	6	7	8	9	10	11		
Human SARS												
21221	3768	825	465	231	666	353	369	135	120	255	1269	297
ORF1ab	Gen S	SARS3a	Sars3b	Gen E	Gen M	Sars6	Sars7a	Sars7b	Sars8a	Sars8b	Gen N	Sars9b
1	2	3	4	5	6	7	8	9	10	11	12	13
Bat SARS												
13167	7887	3726	825	231	666	192	369	135	366	1263	294	213
P1a	P1b	Spike	ORF3a	Gen E	Gen M	ORF6	Sars7a	Sars7b	ORF8	N	ORF 9a	ORF 9b
1	2	3	4	5	6	7	8	9	10	11	12	13

Las secuencias de los once genes presentados por el Wuhan COVID-19 coronavirus (NC\_045512.2) fueron alineados con los genomas de las cepas de Human-SARS y Bat-SARS mediante ClustalW. Los puntajes obtenidos por el alineamiento de cada una de las secuencias con cada uno de los 11 genes coincidió con que los genes 2 y 9 presentes en COVID-19 podrían ser utilizados para su identificación por RT-qPCR, considerando que los dos alineamientos presentaron puntajes menores a 70 % contra regiones de los dos coronavirus (figuras 5 y 6). Los alineamientos de estos dos genes al genoma del virus Human-SARS se pueden apreciar en las figuras 7 y 8.

**Figura 5.** Puntaje de alineamiento (%) de los 11 genes presentes en el Wuhan nCOV-19 vs. el genoma de Human SARS

Gen 1 nCov-19 Wuhan vs. Human SARS	= 76.7861
Gen 2 nCov-19 Wuhan vs. Human SARS	= 68.9499
Gen 3 nCov-19 Wuhan vs. Human SARS	= 75.7256
Gen 4 nCov-19 Wuhan vs. Human SARS	= 91.1392
Gen 5 nCov-19 Wuhan vs. Human SARS	= 84.6411
Gen 6 nCov-19 Wuhan vs. Human SARS	= 64.6552
Gen 7 nCov-19 Wuhan vs. Human SARS	= 81.4189
Gen 8 nCov-19 Wuhan vs. Human SARS	= 64.5161
Gen 9 nCov-19 Wuhan vs. Human SARS	= 29.7297
Gen 10 nCov-19 Wuhan vs. Human SARS	= 87.6471
Gen 11 nCov-19 Wuhan vs. Human SARS	= 87.234

Nota: los puntajes en el recuadro rojo corresponden a los genes candidatos para el diagnóstico.

Fuente: elaboración propia.

**Figura 6.** Puntaje de alineamiento (%) de los 11 genes presentes en el Wuhan nCoV-19 vs. el genoma de Bat SARS

Gen 1 nCov-19 Wuhan vs. Bat SARS	= 75.0189
Gen 2 nCov-19 Wuhan vs. Bat SARS	= <b>67.9904</b>
Gen 3 nCov-19 Wuhan vs. Bat SARS	= 70.5805
Gen 4 nCov-19 Wuhan vs. Bat SARS	= 91.7722
Gen 5 nCov-19 Wuhan vs. Bat SARS	= 82.9716
Gen 6 nCov-19 Wuhan vs. Bat SARS	= 71.5517
Gen 7 nCov-19 Wuhan vs. Bat SARS	= 83.1081
Gen 8 nCov-19 Wuhan vs. Bat SARS	= 74.1935
Gen 9 nCov-19 Wuhan vs. Bat SARS	= <b>67.2297</b>
Gen 10 nCov-19 Wuhan vs. Bat SARS	= 87.6471
Gen 11 nCov-19 Wuhan vs. Bat SARS	= 91.4894

Nota: los puntajes en el recuadro rojo corresponden a los genes candidatos para el diagnóstico.

Fuente: elaboración propia.

**Figura 7.** Human SARS vs. gen 9 (ORF 8) del nCoV-19 (Wuhan)

HUMANSARS_ATATTAGGTTTTACCTACC 9genCovid-19_ATGAAATTTCTTGTTTT	CAAACCATTCTTAGGACAAGCAGCAATTACAACATCAAATTGCGCTAAGA -----
HUMANSARS_ATATTAGGTTTTACCTACC 9genCovid-19_ATGAAATTTCTTGTTTT	GATTAGCACAAACGTGTGTTTAAACAATTATATGCCTTATGTGTTTACATTA -----CATGTACTCAACATCAA----CCATATGTAGTTGATGAC * ** * ** * ** * ** * ** * **
HUMANSARS_ATATTAGGTTTTACCTACC 9genCovid-19_ATGAAATTTCTTGTTTT	TTGTTCCAATTGTGTACTTTTACTAAAAGTACCAATTCTAGAATTAGAGC CCGTGTC--CTATTCACTTCTATTCTAAATGGTA-TATTAGAGTAGGAGC ** * * * ** * ** * ** * * ** * ** * **
HUMANSARS_ATATTAGGTTTTACCTACC 9genCovid-19_ATGAAATTTCTTGTTTT	TTCACTACCTACAACCTATTGCTAAAAATAGTGTTAAGAGTGTGCTAAAT TAGAAAATCAGCACCT-----TTAA-----TTGAAT * * * * * ** * ** * ** * ** * **
HUMANSARS_ATATTAGGTTTTACCTACC 9genCovid-19_ATGAAATTTCTTGTTTT	TATGTTTGGATGCCGGCATTAAATTATGTGAAGTCACCCAAATTTCTAAA TGTGCGTGGATG--AGGCTGGTTCT--AAATCACCCA--TTCAGTACA * ** * ** * ** * ** * ** * ** * ** * ** * **
HUMANSARS_ATATTAGGTTTTACCTACC 9genCovid-19_ATGAAATTTCTTGTTTT	TTGTTCCAAATCGCTATGTGGCTATTGTTGTTAAGTATTTGCT-TAGGTT TCGATATCGGTAATTATACAGTT--TCCTGTTTACCTTTTACAATTAATT * * * * * * ** * * * ** * ** * ** * ** * **
HUMANSARS_ATATTAGGTTTTACCTACC 9genCovid-19_ATGAAATTTCTTGTTTT	CTCTAATCTGTGTAACCTGCTGCTTTTGGTGTACTTATCTAATTTTGGT GCCAGGAACCTAAATGGGTAGTCTTGTAGTGCCTTGTTCGTTCTATGAA * * * * * ** * ** * ** * ** * ** * **
HUMANSARS_ATATTAGGTTTTACCTACC 9genCovid-19_ATGAAATTTCTTGTTTT	GCTCCTCTTATTGTAATGGCGTTAGAGAATTGTATCTTAATTCGCTCAA GACTTTTTAGAGTATCATGACGTTTCGTG--TTGTTTATGATTTTCATCAA * ** * * * ** * ** * ** * ** * ** * ** * **

Fuente: elaboración propia mediante ClustalW (Aligned score =29,72).

**Figura 8.** Alineamiento genoma Human SARS (NC\_004718.3), vs. gen 2 (ORF 8) del virus nCoV-19 (Wuhan, NC\_045512.2).

```

NC_045512.2_21563-25384  ATGTTTGTT
NC_004718.3             ATGTTTATT
                        *****
NC_045512.2_21563-25384  TTTCT---TGTTTTATTGCCACTAGT-----CTCTAGTCAGTGTGT
NC_004718.3             TTCTTATTATTTCTTACTCTCACTAGTGGTAGTGACCTTGACCGGTGCAC
                        ** * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  TAATCTT-ACAA---CCAGAACTCAATTACCCCTGCATACACTAATTC
NC_004718.3             CACTTTTGATGATGTTCAAGCTCCTAATTACACTCAACATAC-----TTC
                        * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  TTTCACACGTGGTGTATTACCCTGACAAAGTTTTTCAGATCCTCAGTTT
NC_004718.3             ATCTATGAGGGGGGTTTACTATCCTGATGAAATTTTTAGATCAGACACTC
                        * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  TACATTCAACTCAGGACTTGTCTTACCTTTCTTTCCAATGTTACTTGG
NC_004718.3             TTTATTTAACTCAGGATTTATTTCTCCATTTTATTCTAATGTTACAGGG
                        * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  TTCCATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAA
NC_004718.3             TTTCACTAT-----TAATCATAC-----GTTTGCAA
                        ** * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  CCCTGTCCTACCATTTAATGATGGTGTATTGCTTCCACTGAGAAGT
NC_004718.3             CCCTGTCATACCTTTAAGGATGGTATTTATTTGCTGCCACAGAGAAAT
                        ***** * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  CTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTCTGAAGACC
NC_004718.3             CAAATGTTGTCCGTGGTGGGTTTTGGTTCTACCATGAACAACAAGTCA
                        * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  CAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTG
NC_004718.3             CAGTCGGTGATTATTATTAACAATTCTACTAATGTTGTTATACGAGCATG
                        ***** * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  TGAATTTCAATTTTGAATGATCCATTTTGGGTGTTTATTACCACAAAA
NC_004718.3             TAACTTTGAATTGTGTGACAACCCTTCTTTGCTGTTTCTAAACCCATGG
                        * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  ACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGTGCGAAT

```

```

NC_004718.3      GTA-CACAGAC---ACATACTATGATATTCGA-----TAATGCATTT
                * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  AATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGG
NC_004718.3      AATTGCACTTTTCGAGTACATATCTGATGCCTTTTCGCTTGATGTTTCAGA
                ***** * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  AAAACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTG
NC_004718.3      AAAGTCAGGTAATTTAAACACTTACGAGAGTTTGTGTTAAAAATAAAG
                *** ***** * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  ATGGTTATTTTAAAATATATTCTAAGCACACGCCTATTAATTTAGTGCCT
NC_004718.3      ATGGGTTTCTCTATGTTTATAAGGGCTATCAACCTATAGATGTAGTTCGT
                ***** * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  GATCTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGATTTGCCAAT
NC_004718.3      GATCTACCTTCTGGTTTTAACACTTTGAAACCTATTTTAAAGTTGCCTCT
                ***** * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  AGGTATTAACATCACTAGGTTTCAAACCTTACTTGCTTTACATAGAAGTT
NC_004718.3      TGGTATTAACATTACAAATTTAGAGCCATTCTTAC-----AGCC
                ***** * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  ATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTGCAGCT
NC_004718.3      TTTTCA--CCTGCTCAAGACATTTGGGGCACGTCAGCTG-----CAGCC
                * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  TATTATGTGGTTATCTTCAACCTAGGACTTTTCTATTAATATAATGA
NC_004718.3      TATTTTGTGGCTATTTAAAGCCAACACTACATTTATGCTCAAGTATGATGA
                ***** * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  AAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCAG
NC_004718.3      AAATGGTACAATCACAGATGCTGTTGATTGTTCTCAAAATCCACTTGCTG
                ***** * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  AAACAAAGTGTACGTTGAAATCCTTCACTGTAGAAAAAGGAATCTATCAA
NC_004718.3      AACTCAAATGCTCTGTTAAGAGCTTTGAGATTGACAAAGGAATTTACCAG
                ** * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  ACTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCCCTAA
NC_004718.3      ACCTCTAATTTCAGGGTTGTTCCCTCAGGAGATGTTGTGAGATTTCCCTAA
                * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384  TATTACAAACTTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTTG
    
```

NC\_004718.3 TATTACAACTTGTGTCCTTTTGGAGAGGTTTTAATGCTACTAAATTCC  
\*\*\*\*\*

NC\_045512.2\_21563-25384 CATCTGTTTATGCTTGG AACAGGAAGAGAATCAGCAACTGTGTTGCTGAT  
NC\_004718.3 CTTCTGTCTATGCATGGGAGAGAAAAAATTTCTAATTGTGTTGCTGAT  
\*\*\*\*\*

NC\_045512.2\_21563-25384 TATTCTGTCCTATATAATTCCGCATCATTTTCCACTTTTAAGTGTTATGG  
NC\_004718.3 TACTCTGTGCTCTACAACCAACATTTTTTCAACCTTAAGTGCTATGG  
\*\*\*\*\*

NC\_045512.2\_21563-25384 AGTGTCTCCTACTAAATTAATGATCTGCTTTACTAATGTCTATGCAG  
NC\_004718.3 CGTTTCTGCCACTAAGTTGAATGATCTTTGCTTCTCCAATGTCTATGCAG  
\*\*\*\*\*

NC\_045512.2\_21563-25384 ATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAATCGCTCCAGGGCAA  
NC\_004718.3 ATTCTTTGTAGTCAAGGAGATGATGTAAGACAAATAGCGCCAGGACAA  
\*\*\*\*\*

NC\_045512.2\_21563-25384 ACTGGAAAGATTGCTGATTATAATTATAAATTACCAGATGATTTTACAGG  
NC\_004718.3 ACTGGTGTTATTGCTGATTATAATTATAAATTGCCAGATGATTTTATGGG  
\*\*\*\*\*

NC\_045512.2\_21563-25384 CTGCGTTATAGCTTGG AATTCTAACAATCTTGATTCTAAGGTTGGTGGTA  
NC\_004718.3 TTGTGTCCTTGCTTGG AATACTAGGAACATTGATGCTACTTCAACTGGTA  
\*\*\*\*\*

NC\_045512.2\_21563-25384 ATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTT  
NC\_004718.3 ATTATAATTATAAATATAGGTATCTTAGACATGGCAAGCTTAGGCCCTTT  
\*\*\*\*\*

NC\_045512.2\_21563-25384 GAGAGAGATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTA  
NC\_004718.3 GAGAGAGACATATCTAATGTGCCTTTCTCCCCTGATGGCAAACCTTGCAC  
\*\*\*\*\*

NC\_045512.2\_21563-25384 TGGTGTTGAAGTTTTAATTGTTACTTTCTTTACAATCATATGGTTTCC  
NC\_004718.3 CCCACCTGC---TCTTAATTGTTATTGGCCATTAAATGATTATGGTTTTT  
\*\*\*\*\*

NC\_045512.2\_21563-25384 AACCCTAATGGTGTGGTTACCAACCATACAGAGTAGTAGTACTTTCT  
NC\_004718.3 ACACCCTACTGGCATTGGCTACCAACCTTACAGAGTTGTAGTACTTTCT  
\*\*\*\*\*

NC\_045512.2\_21563-25384 TTTGAACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTAC

```
NC_004718.3      TTTGAACTTTTAAATGCACCGCCACGTTTGTGGACCAAATTATCCAC
***** ** ***** ** ** ***** ** ** ***** ** **
NC_045512.2_21563-25384  TAATTTGGTTAAAAACAAATGTGTCAATTTCAACTTCAATGGTTAACAG
NC_004718.3      TGACCTTATTAAGAACCAGTGTGTCAATTTAATTTAATGGACTCACTG
* * * ***** ** ***** ** ** ***** ** **
NC_045512.2_21563-25384  GCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAACAA
NC_004718.3      GTACTGGTGTGTTAACTCCTTCTTCAAAGAGATTCAACCATTTCACAA
* * * ***** ** ** ** * * * ***** ** ** *****
NC_045512.2_21563-25384  TTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAGAC
NC_004718.3      TTTGGCCGTGATGTTTCTGATTTCACTGATTCCGTTTCGAGATCCTAAAAC
***** * * * ***** ***** * * * * ***** * * *
NC_045512.2_21563-25384  ACTTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTGAGTGTTA
NC_004718.3      ATCTGAAATATTAGACATTTACCTTGCCTTTTGGGGGTGAAGTGTA
* * * * * ***** ***** ** ***** ***** ***** *
NC_045512.2_21563-25384  TAACACCAGGAACAAATACTTCTAACCAGTTGCTGTTCTTTATCAGGAT
NC_004718.3      TTACACCTGGAACAAATGCTTCATCTGAAGTTGCTGTTCTATATCAAGAT
***** ***** ***** * ***** ***** ***** **
NC_045512.2_21563-25384  GTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCC
NC_004718.3      GTTAACTGCACTGATGTTTCTACAGCAATTCATGCAGATCAACTCACACC
***** ***** ** * * * ***** ***** ***** ** **
NC_045512.2_21563-25384  TACTTGGCGTGTATTATTCTACAGGTTCTAATGTTTTTCAAACACGTGCAG
NC_004718.3      AGCTTGGCGCATATATTCTACTGGAAACAATGATTCCAGACTCAAGCAG
***** * ***** ** ***** ** * * * * *****
NC_045512.2_21563-25384  GCTGTTTAATAGGGGCTGAACATGTCAACAACCTCATATGAGTGTGACATA
NC_004718.3      GCTGTCTTATAGGAGCTGAGCATGTCGACACTTCTATGAGTGCACATT
***** * ***** ***** ***** ** * ***** *****
NC_045512.2_21563-25384  CCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCC
NC_004718.3      CCTATTGGAGCTGGCATTGTGCTAGTTACCATAC-----AGTTTCT
** ***** ** * * * * ***** ** * * * *
NC_045512.2_21563-25384  TCGGCGGGCACGTAGTG-TAGCTAGTCAATCCATCATTGCCTACACTATG
NC_004718.3      TTATT----ACGTAGTACTAGCCAAA-AATCTATTGTGGCTTATACTATG
* ***** ***** * ***** ** * * * * *****
```

```
NC_045512.2_21563-25384   TCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTATTGCCAT
NC_004718.3              TCTTTAGGTGCTGATAGTTCAATTGCTTACTCTAATAACACCATTGCTAT
** * ***** ** * ***** ***** ***** ***** ***** **
NC_045512.2_21563-25384   ACCCACAAATTTTACTATTAGTGTACCACAGAAATTCTACCAGTGTCTA
NC_004718.3              ACCTACTAACTTTTCAATTAGCATTACTACAGAAGTAATGCCTGTTTCTA
*** * * * * * * * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384   TGACCAAGACATCAGTAGATTGTACAATGTACATTTGTGGTGATTCAACT
NC_004718.3              TGGCTAAAACCTCCGTAGATTGTAATATGTACATCTGCGGAGATTCTACT
** * * * * * * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384   GAATGCAGCAATCTTTTGTGCAATATGGCAGTTTTTGTACACAATTA
NC_004718.3              GAATGTGCTAATTTGCTTCTCCAATATGGTAGCTTTTGCACACAATA
***** ** * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384   CCGTGCTTAACTGGAATAGCTGTTGAACAAGACAAAACACCCAAGAAG
NC_004718.3              TCGTGCACTCTCAGGTATTGCTGCTGAACAGGATCGCAACACACGTGAAG
***** * * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384   TTTTGCACAAGTCAAACAAATTTACAAAACACCACCAATTAAGATTTT
NC_004718.3              TGTTGCTCAAGTCAAACAAATGTACAAAACCCCAACTTTGAAATATTTT
* * * * * * * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384   GGTGGTTTTAATTTTTCACAAATATTACCAGATCCATCAAACCAAGCAA
NC_004718.3              GGTGGTTTTAATTTTTCACAAATATTACCTGACCCTCTAAAGCCAATA
***** ***** ***** ***** ***** ***** ***** *****
NC_045512.2_21563-25384   GAGGTCATTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATG
NC_004718.3              GAGGTCTTTTATTGAGGACTTGCTCTTAATAAGGTGACACTCGCTGATG
***** ***** * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384   CTGGCTTCATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGA
NC_004718.3              CTGGCTTCATGAAGCAATATGCGAATGCCTAGGTGATATTAATGCTAGA
***** * * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384   GACCTCATTGTGCGCAGAAAGTTAACGGCCTTACTGTTTTGCCACCTTT
NC_004718.3              GATCTCATTGTGCGCAGAAAGTTCAATGGACTTACAGTGTTGCCACCTCT
** ***** * * * * * * * * * * * * * * * * * * * * * *
NC_045512.2_21563-25384   GCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTA
NC_004718.3              GCTCACTGATGATGATTGCTGCCTACACTGCTGCTCTAGTTAGTGGTA
***** ***** ***** ***** ***** ***** ***** *****
```



NC\_045512.2\_21563-25384 TGTGACTTATGTCCCTGCACAAGAAAAGAACTTCACAACCTGCTCCTGCCA  
NC\_004718.3 TGTCACGTATGTGCCATCCCAGGAGAGGAACTTCACCACAGCGCCAGCAA  
\*\*\* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 TTTGTCATGATGGAAAAGCACACTTCCTCGTGAAGGTGTCTTTGTTCA  
NC\_004718.3 TTTGTCATGAAGGCAAAGCATACTCCCTCGTGAAGGTGTTTTGTGTTT  
\*\*\*\*\* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 AATGGCACACACTGGTTTGTAAACACAAAGGAATTTTATGAACCACAAAT  
NC\_004718.3 AATGGCACTTCTTGTTTATTACACAGAGGAACTTCTTTTCTCCACAAAT  
\*\*\*\*\* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 CATTACTACAGACAACACATTTGTGTCTGGTAACTGTGATGTTGTAATAG  
NC\_004718.3 AATTACTACAGACAATACATTTGTCTCAGGAAATTGTGATGTCGTTATTG  
\*\*\*\*\* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 GAATTGTCAACAACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCA  
NC\_004718.3 GCATCATTAAACAACACAGTTTATGATCCTCTGCAACCTGAGCTTGACTCA  
\* \* \* \* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 TTCAAGGAGGAGTTAGATAAATATTTAAGAATCATAACATCACCAGATGT  
NC\_004718.3 TTCAAAGAAGAGCTGGACAAGTACTTCAAAAATCATAACATCACCAGATGT  
\*\*\*\*\* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 TGATTTAGGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAACATTCAAA  
NC\_004718.3 TGATCTTGGCGACATTTAGGCATTAACGCTTCTGTGTCGCAACATTCAAA  
\*\*\*\*\* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 AAGAAATTGACCGCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTC  
NC\_004718.3 AAGAAATTGACCGCTCAATGAGGTCGCTAAAAATTTAAATGAATCACTC  
\*\*\*\*\* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 ATCGATCTCCAAGAACTTGAAAGTATGAGCAGTATATAAAATGGCCATG  
NC\_004718.3 ATTGACCTTCAAGAATTGGGAAAATATGAGCAATATATTAATGGCCTTG  
\* \* \* \* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 GTACATTTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAGTAATGGTGA  
NC\_004718.3 GTATGTTTGGCTCGGCTTCATTGCTGGACTAATTGCCATCGTCATGGTTA  
\*\*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*  
NC\_045512.2\_21563-25384 CAATTATGCTTTGCTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGT  
NC\_004718.3 CAATCTTGCTTTGTTGCATGACTAGTTGTTGCAGTTGCCTCAAGGGTGCA  
\*\*\*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*\*\*\*\* \*\* \* \*\* \* \*\* \*

```

NC_045512.2_21563-25384   TGTTCTTGTTGATCCTGCTGCAAATTTGATGAAGACGACTCTGAGCCAGT
NC_004718.3               TGCTCTTGTTGTTCTTGCTGCAAGTTTGATGAGGATGACTCTGAGCCAGT
** ***** ** ***** ***** ** *****
NC_045512.2_21563-25384   GCTCAAAGGAGTCAAATTACATTACACATAA-----
NC_004718.3               TCTCAAGGGTGTCAAATTACATTACATAAACGAACTTATGGATTGTT
***** ** *****
    
```

Fuente: elaboración propia mediante ClustalW (Aligned score = 67,99).

### Diseño de primers-oligonucleótidos

Utilizando el programa Primer3 versión 4.1.0 y la secuencia completa de los dos genes candidatos (genes 2 y 9) como posibles marcadores moleculares para la identificación de COVID-19, se diseñaron dos pares de primers con una longitud de 20 pb cada uno, los cuales amplifican en su totalidad cada uno de los genes. Sus características se presentan en la tabla 2.

A pesar de que el gen número 8 también produjo una puntuación baja, se descartó su uso debido a que este correspondía a una secuencia muy corta dentro del genoma del Wuhan nCoV-19 (figura 1) y por tanto no cumple con las características para el diseño de primers.

**Tabla 2.** Características de los primers diseñados para la identificación de COVID-19

Primer/Gen	Secuencia	TM (°C)	%G-C	Producto
CVD-19UJTL-02 Forward CVD-19UJTL-02 Reverse	5" ACACTACTGATGCTGTCCGT 3" 5" GCCCTATTAAACAGCCTGC 3"	58,74 58,69	50 55	3822 (pb)
CVD-19UJTL-09 Forward CVD-19UJTL-09 Reverse	5" ACTGTAGCTGCATTTACCA 3" 5" AGAACCGCCTCATCCACG 3"	57,43 59,18	45 55	366 (pb)

### Diseño de sondas

Para el diseño de las sondas se seleccionaron 2 regiones dentro de los genes 2 y 9, de 20 y 24 pb, respectivamente. Con ayuda del programa MegaX, se obtuvieron las cadenas complementarias a esas regiones

utilizando la opción *reverse complement*. Las regiones seleccionadas se hibridarán por complementariedad a los genes que se requiere identificar. Además, a estas sondas se les incorporará dos tipos de fluorocromos (un aceptor y un donador). Las sondas diseñadas se presentan a continuación:

**Sonda COVID-19- Gen2 (Size = 20pb)**

5'GATCCACAGACTTGAGATTCTTGACATTACACCATGTTCTTTTGGTGGTGTTCAGTGTATAACACCAGGAACAAATAC-TTCTAACCAGGTTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTATTATTCTACAGGTTCTAATGTTTTTCAAACACGT-3' → **CADENA A AMPLIFICAR DELIMITADA POR LOS PRIMERS**

5'-ATGTTCTTTTGGTGGTGTCA-3' → **REGIÓN A LA QUE SE UNIRÁ LA SONDA**

3'-TACAAGAAAACCACCACAGT-5' → **SONDA**

**Sonda COVID-19- Gen9 (Size = 24pb)**

5'TAGTTTACAGTCATGTACTCAACATCAACCATATGTAGTTGATGACCCGTGCCTATTCACTTCTATTCTAAATGGTATATTAGAGTAGGAGCTAGAAAATCAGCACCTTTAA-3' → **CADENA A AMPLIFICAR DELIMITADA POR LOS PRIMERS**

5'-ATGTACTCAACATCAACCATATGTA-3' → **REGIÓN A LA QUE SE UNIRÁ LA SONDA**

3'-TACATGAGTTGTAGTTGGTATACAT -5' → **SONDA**

Una vez tenemos los primers y las sondas diseñadas para amplificar un fragmento de los genes S y ORF8, es posible afirmar que contamos con los componentes necesarios para realizar la identificación del COVID-19; los demás reactivos, como enzimas, tampones y cofactores, se consiguen fácilmente en el mercado.

**DISCUSIÓN**

Se propuso un diseño para el desarrollo de la reacción de RT-qPCR y de esta forma identificar el virus COVID-19 responsable de la pandemia que actualmente se desarrolla en todo el planeta. Para ello, se utilizaron métodos moleculares (RT-qPCR) y programas bioinformáticos que permitieron analizar las secuencias del virus y a partir de ellas diseñar los primers y sondas para su amplificación y cuantificación. Por medio de este ejercicio, los estudiantes lograron comprender las variables de las dos reacciones: RT-PCR y la PCR Tiempo Real.

Aunque, en general, se utilizaron de forma adecuada las herramientas bioinformáticas para encontrar los genes, diseñar los primers y las sondas, se presentó un error conceptual, puesto que al inicio se realizaron las comparaciones entre genes y genomas. No obstante, se hizo necesario comparar gen con gen para detectar la identidad de cada uno de ellos y así detectar zonas de homología y zonas variables, con el fin de diseñar los primers de tal manera que se aseguraran de amplificar fragmentos únicos del COVID-19. Finalmente, esto se consiguió gracias a que los estudiantes aprendieron del error, lo corrigieron y lograron identificar los genes más variables.

ClustalW es una herramienta sencilla que permite explicar conceptos básicos de biología molecular y hacer comparaciones entre secuencias equidistantes y de una longitud similar (Thompson, Plewniak, & Poch, 1999). Al realizar comparaciones entre secuencias de

diferente longitud, y con base en la plataforma que se utilice, el resultado de los alineamientos difiere, ya que cada plataforma maneja sus propios parámetros por defecto. Por ende, cuando se realizan alineamientos múltiples se asignan diferentes puntuaciones a los gaps e inserciones para optimizarlos (Gaskell, 2000).

La realización de este tipo de ejercicios académicos por parte de los estudiantes de biología molecular permite afianzar los conocimientos adquiridos, incentivar la creatividad y gusto por las ciencias ómicas y despertar su curiosidad por comprender nuevas herramientas como la bioinformática, lo cual nos permite adentrarnos y comprender nuevos conceptos y procesos relacionados con la vida, contribuyendo a descubrir los misterios ocultos de la naturaleza.

## CONCLUSIONES

Se logró el objetivo planteado para esta investigación al analizar las secuencias de los virus y diseñar primers y sondas, los cuales son los insumos más importantes para realizar el diagnóstico de COVID-19.

El trabajo presentado consistió en la aplicación de la metodología de aprendizaje basado en problemas (ABP) a manera de un ejercicio de aprendizaje en contexto, el cual fue llevado al plano actual de la pandemia de COVID-19 con un enfoque interdisciplinar en el que se debían aplicar conocimientos teóricos de métodos en biología molecular y emplear diferentes herramientas bioinformáticas.

ClustalW y primers 3 son dos herramientas bioinformáticas fáciles de usar que resultan muy valiosas, ya que permiten a los estudiantes afianzar los conocimientos adquiridos en clase y desarrollar su capacidad de análisis.

Ejercicios académicos como estos ayudan a los estudiantes a comprender la importancia de la biología molecular, la bioinformática y las ciencias ómicas para el desarrollo de estrategias que permitan detectar virus emergentes y contribuir al desarrollo de la ciencia.

## REFERENCIAS

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403-410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Biomedical Information Science and Technology Initiative Consortium [BISTIC]. (2000). *Definition Committee NIH working definition of bioinformatics and computational biology*. <http://www.bisti.nih.gov/CompuBioDef.pdf>
- Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., ... & Mulders, D. G. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, 25(3), 2000045. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>
- Costa, J. (2004). Reacción en cadena de la polimerasa (PCR) a tiempo real. *Enfermedades infecciosas y microbiología clínica*, 22(5), 299-305. [https://doi.org/10.1016/S0213-005X\(04\)73092-X](https://doi.org/10.1016/S0213-005X(04)73092-X)
- Dong, E., Du, H., & Gardner, L. (2020) An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis*, published online February 19. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)
- Furge, L. L., Stevens-Truss, R., Moore, D. B., Langeland, J. A. (2009). Vertical and horizontal integration of bioinformatics education: A modular, interdisciplinary approach. *Biochem Mol Biol Educ.*, 37(1): 26-36. <https://doi.org/10.1002/bmb.20249doi: 10.1002/bmb.20249>.
- Gaskell, G. J. (2000). Multiple sequence alignment tools on the Web. *BioTechniques*, 29(1), 60-62. <https://doi.org/10.2144/00291io01>
- Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., & Chang, H. (2020). The architecture of SARS-CoV-2 transcriptome. *bioRxiv*, (preprint). <https://doi.org/10.1101/2020.03.12.988865>
- Lexa, M., Horak, J., & Brzobohaty, B. (2001). Virtual PCR. *Bioinformatics*, 17(1), 192-193. <https://doi.org/10.1093/bioinformatics/17.2.192>

- Li, K. B. (2003). ClustalW-MPI: ClustalW analysis using distributed and parallel computing. *Bioinformatics*, 19(12), 1585-1586. <https://doi.org/10.1093/bioinformatics/btg192>
- National Center for Biotechnology Information [NCBI]. (2020). *GenBank. Bat SARS coronavirus Rm1, complete genome. >DQ412043.1, FASTA.* <https://www.ncbi.nlm.nih.gov/nuccore/DQ412043.1?report=fasta>.
- National Center for Biotechnology Information [NCBI]. (2020). *GenBank. SARS coronavirus, complete genome >NC\_004718.3, FASTA.* [https://www.ncbi.nlm.nih.gov/nuccore/NC\\_004718.3?report=fasta](https://www.ncbi.nlm.nih.gov/nuccore/NC_004718.3?report=fasta).
- National Center for Biotechnology Information [NCBI]. (2020). *GenBank. Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome.>NC\_045512.2, FASTA.* [https://www.ncbi.nlm.nih.gov/nuccore/NC\\_045512.2?report=fasta](https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.2?report=fasta).
- Novak, J. D. (1988). *Teoría y Práctica de la educación*. Madrid: Alianza Editorial.
- Pinilla, G. (2019). *Biología molecular: ADN recombinante y sus aplicaciones*. Bogotá: Manual Moderno.
- Prieto, L. (2006). Aprendizaje activo en el aula universitaria: el caso del aprendizaje basado en problemas, en *Miscelánea Comillas. Revista de Ciencias Humanas y Sociales*, 64(24), 173-196.
- Salazar, A., Sandoval, A., & Armendáriz, J. (2013). *Biología molecular fundamentos y aplicaciones en las ciencias de la salud*. Ciudad de México: McGraw Hill Education.
- Thompson, J. D., Plewniak, F., & Poch, O. (1999). A comprehensive comparison of multiple sequence alignment programs. *Nucleic acids research*, 27(13), 2682-2690. <https://doi.org/10.1093/nar/27.13.2682>
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic acids research*, 40(15), e115-e115. <https://doi.org/10.1093/nar/gks596>
- Vincze, T., Posfai, J., & Roberts, R. J. (2003). NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Res.*, 31, 3688-3691. <https://doi.org/10.1093/nar/gkg526>
- Wu, F., Zhao, S., Yu, B., Chen, Y. M., Wang, W., Song, Z. G., ... & Zhang, Y. Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, 579, 265-269. <https://doi.org/10.1038/s41586-020-2008-3>