

Investigación desarrollada en la Universidad de O'Higgins revela la diversidad oculta del SARS-CoV-2

- *El estudio, publicado en Scientific Reports, revista de Nature Portfolio, analizó 96 muestras chilenas con una profundidad promedio cercana a 60.000x. Es la primera investigación en Chile que utiliza secuenciación ultra profunda del genoma completo para caracterizar la evolución del SARS-CoV-2 dentro de las personas infectadas.*

Una investigación concebida y desarrollada íntegramente en la Universidad de O'Higgins (UOH) permitió observar una dimensión del SARS-CoV-2 que la vigilancia genómica convencional generalmente no detecta: las variantes minoritarias que aparecen dentro de cada persona durante la infección.

El trabajo, titulado "Ultra-deep sequencing reveals intra-host diversity and selective constraint in SARS-CoV-2 samples from Chile", fue publicado en la revista científica *Scientific Reports*. El estudio fue coordinado por el Dr. Alex Di Genova, académico del Instituto de Ciencias de la Ingeniería y director del Centro UOH de Bioingeniería (CUBI). En la investigación participaron también la académica Carol Moraga y el académico Mauricio Latorre.

La investigación analizó 96 muestras de SARS-CoV-2 obtenidas en Chile entre 2020 y 2022, periodo que abarcó la circulación y el reemplazo de distintos linajes, desde las variantes iniciales hasta Lambda, Delta y Ómicron. La preparación de las bibliotecas, la secuenciación, el procesamiento bioinformático

y la interpretación de los resultados se realizaron en la UOH.

Un error afortunado que abrió una nueva pregunta

El proyecto surgió en 2023, durante una prueba del secuenciador DNBSeg-G400. El equipo utilizó por error una celda de secuenciación con una capacidad mucho mayor que la necesaria para las muestras disponibles. Como resultado, los genomas virales fueron secuenciados a una profundidad promedio cercana a 60.000×, muy por encima de lo requerido para identificar la variante predominante en cada paciente.

En lugar de descartar ese exceso de información, los investigadores se preguntaron qué procesos biológicos podían observarse con ese nivel de resolución. La respuesta estaba en las variantes intra-hospedero o iSNVs: pequeños cambios genéticos presentes en una fracción de los virus que coexisten dentro de una misma persona infectada.

“Todo comenzó con un error durante una prueba del secuenciador: utilizamos una celda mucho más grande de la que necesitábamos y terminamos secuenciando las muestras a una profundidad enorme. En vez de quedarnos solamente con el genoma principal del virus, aprovechamos esos datos para mirar qué estaba ocurriendo dentro de cada paciente. Ese error terminó abriendo una pregunta científica mucho más interesante”, explica Alex Di Genova.

Más allá de la variante dominante

La vigilancia genómica tradicional reconstruye un genoma consenso, es decir, la versión predominante del virus presente en una muestra. Esta información fue fundamental durante la pandemia para seguir la aparición y expansión de variantes como Gamma, Lambda, Delta y Ómicron. Sin embargo, el consenso no muestra toda la diversidad viral que existe dentro de un

individuo.

La secuenciación ultra profunda permitió examinar esas variantes minoritarias con criterios estrictos para separar señales biológicas reales de posibles errores técnicos. En total, el equipo generó más de 2.000 millones de lecturas y, después de un proceso conservador de filtrado, identificó 2.890 iSNVs respaldadas por una alta profundidad de secuenciación.

La mayoría se encontraba en frecuencias bajas o intermedias y el cambio más común fue de citosina a timina, un patrón consistente con procesos de mutación y edición del ARN descritos previamente para el SARS-CoV-2.

Señales tempranas de evolución viral y coinfecciones

El análisis mostró que la evolución del SARS-CoV-2 dentro de los pacientes estuvo dominada por selección purificadora. Esto significa que la mayoría de las mutaciones que aparecen durante la infección son eliminadas porque pueden afectar negativamente el funcionamiento del virus.

Esta restricción no fue igual en todo el genoma. La proteína Spike y algunos genes accesorios, como *ORF3a*, *ORF7a* y *ORF8*, concentraron una mayor proporción de cambios que modifican proteínas. Estas señales permiten identificar regiones donde la diversidad viral puede tener mayores consecuencias funcionales, aunque no implican por sí solas la aparición de nuevas variantes.

El estudio detectó además 13 coinfecciones candidatas. En estas muestras se observó un linaje viral predominante junto con mutaciones compatibles con un segundo linaje presente en menor proporción. Los casos coincidieron con periodos en que distintas variantes circulaban simultáneamente o se

reemplazaban en Chile.

“Los iSNVs son una ventana a las primeras etapas de la evolución viral dentro del hospedero. La mayoría de estos cambios no llegará a establecerse, pero nos permite observar dónde se genera la diversidad, qué regiones del virus están más restringidas y cuáles acumulan señales de cambios funcionales. Es una capa de información que normalmente se pierde en la vigilancia basada únicamente en genomas de consenso”, señala el académico.

Ciencia realizada de principio a fin en la UOH

Una característica central del estudio es que sus etapas experimentales y computacionales fueron desarrolladas en la Universidad de O'Higgins. Las muestras se procesaron y secuenciaron localmente en el equipo DNBSeg-G400, mientras que el análisis bioinformático se ejecutó mediante flujos de trabajo reproducibles en KÜTRAL, la infraestructura de cómputo de alto rendimiento de la UOH. Solo la ejecución del flujo principal de análisis requirió cerca de 1500 horas de CPU.

El trabajo integró las capacidades del Laboratorio de Biología Computacional, el Laboratorio de Bioingeniería, SeqUOH, CUBI y KÜTRAL. Además, los datos, el código empleado en los análisis y las figuras se encuentran disponibles públicamente en el repositorio del CBLab, mientras que las lecturas de secuenciación fueron depositadas en repositorios científicos internacionales.

“Este trabajo demuestra que en la UOH tenemos la capacidad de desarrollar un estudio genómico completo: desde generar los datos hasta analizarlos e interpretarlos. Esa capacidad no sirve solamente para estudiar el COVID-19; queda instalada para responder a futuras enfermedades infecciosas y para abordar otros problemas biomédicos relevantes para Chile”,

destaca Di Genova.

Los resultados muestran que la secuenciación ultra profunda puede complementar la vigilancia genómica tradicional al revelar procesos tempranos y sutiles de evolución viral. Al mismo tiempo, el estudio evidencia que desde la Región de O'Higgins es posible generar investigación genómica de frontera, con infraestructura y capacidades científicas desarrolladas en la propia Universidad.