

Reporte nº 21– Provincia de Santa Fe – 13 de Mayo de 2021

Consortio interinstitucional para la Secuenciación del genoma y estudios genómicos de SARS-CoV-2
(Proyecto PAIS)

Estado de la situación de la PRIMERA OLA en la Provincia de Santa Fe: Caracterización genómica de SARS-CoV-2 en pacientes con diagnóstico de COVID-19 en la provincia de Santa Fe entre marzo y diciembre de 2020

RESUMEN

Se realizó la secuenciación del **genoma completo** de SARS-CoV-2 a partir de **212** muestras obtenidas de individuos con diagnóstico de la COVID-19 de la **provincia de Santa Fe** correspondientes al período comprendido entre marzo y diciembre del 2020. **Las secuencias obtenidas fueron asignadas a 6 linajes: B.1.499 (56,1%), N.5 (35,4%), N.3 (1,4%), B.1 (4,7%), B.1.1 (1,4%) y B.1.1.28 (0,9%).** Los linajes detectados coinciden con los reportados previamente en Argentina. Particularmente, en Santa Fe se observó un cambio en la proporción de los linajes en circulación a lo largo del período estudiado, donde el linaje B.1.499, identificado con mayor frecuencia entre junio y octubre, habría sido desplazado por el linaje N.5, detectado mayormente en noviembre y diciembre.

Como parte de la evolución natural del virus, se encontraron cambios aminoacídicos en distintas regiones del genoma viral de algunas secuencias, entre ellas en la región que codifica para la proteína Spike. Sin embargo, las implicancias de estos cambios se encuentran bajo análisis y no revisten por el momento consecuencias en la biología viral.

Estos resultados constituyen un punto de partida y muestran los linajes que circularon en la provincia de Santa Fe durante la primera ola de la pandemia de SARS-CoV-2 durante el año 2020. Futuros análisis son necesarios para caracterizar la segunda ola y así poder conocer cómo la introducción de las nuevas variantes fue alterando el patrón epidemiológico original de esta provincia.

En este informe se analiza la circulación de SARS-CoV-2 en la provincia de Santa Fe desde el 23/3/2020 al 22/12/2020 (SE 13/2020 a SE 52/2020).

El principal **objetivo** es describir la introducción, circulación y establecimiento de los linajes de SARS-CoV-2 en la provincia de Santa Fe y profundizar el estudio del origen y evolución de los brotes relevantes en la provincia.

Situación en el territorio nacional durante el período de análisis.

Respecto de la situación epidemiológica de la nación, al inicio del período de análisis de este informe, se encontraba en vigencia el Aislamiento Social, Preventivo y Obligatorio (ASPO) estricto en todo el territorio argentino, la circulación entre provincias se encontraba restringida y fronteras internacionales cerradas. En el mes de junio la mayor parte del territorio nacional pasó a fase de distanciamiento social (DISPO), incluida la provincia de Santa Fe (excepto la ciudad de Villa Ocampo). A partir del ingreso a esta fase, se comienzan a habilitar las reuniones familiares, las actividades culturales y religiosas, el traslado dentro de la provincia hasta 30 km salvo en los aglomerados Gran Rosario y Gran Santa Fe, los deportes al aire libre y actividades comerciales. A partir del 31 de julio se decretó que las personas residentes en la Provincia de Santa Fe que se trasladen hacia la ciudad de Rosario u otra ciudad con circulación local de la COVID-19, quedarían excluidas de participar a su regreso al lugar de residencia de actividades sociales por 10 días. En ese mismo mes se suspenden las reuniones familiares en todo el territorio provincial y en diciembre se establecen ciertas condiciones para las reuniones familiares de fin de año.

Situación epidemiológica de la provincia de Santa Fe durante el período de análisis

El establecimiento del ASPO a nivel Nacional, y luego su implementación por regiones, permitió que los casos reportados en la provincia de Santa Fe no tuvieran un aumento al inicio de la pandemia. Es decir, luego del primer caso reportado en la provincia el 14 de Marzo de 2020, la situación epidemiológica de la misma mostró un número de casos reducido hasta el mes de Julio, donde comenzó a registrarse un aumento de casos (Figura 1a). La mayor cantidad de casos se produjo en los departamentos más poblados, y su normalización por cantidad de habitantes (Figura 1b) resulta comparable, si bien se observa cierta variación en la incidencia.

El total de casos reportados por la provincia hasta la fecha (20/04/2021) es de 256.120 casos confirmados, con 234.934 recuperados y 4.367 fallecidos. Para la ventana temporal analizada en el presente informe, al 22/12/2020 el total de casos registrados fue de 168.607 casos, con 155.904 recuperados y 2.754 fallecimientos.

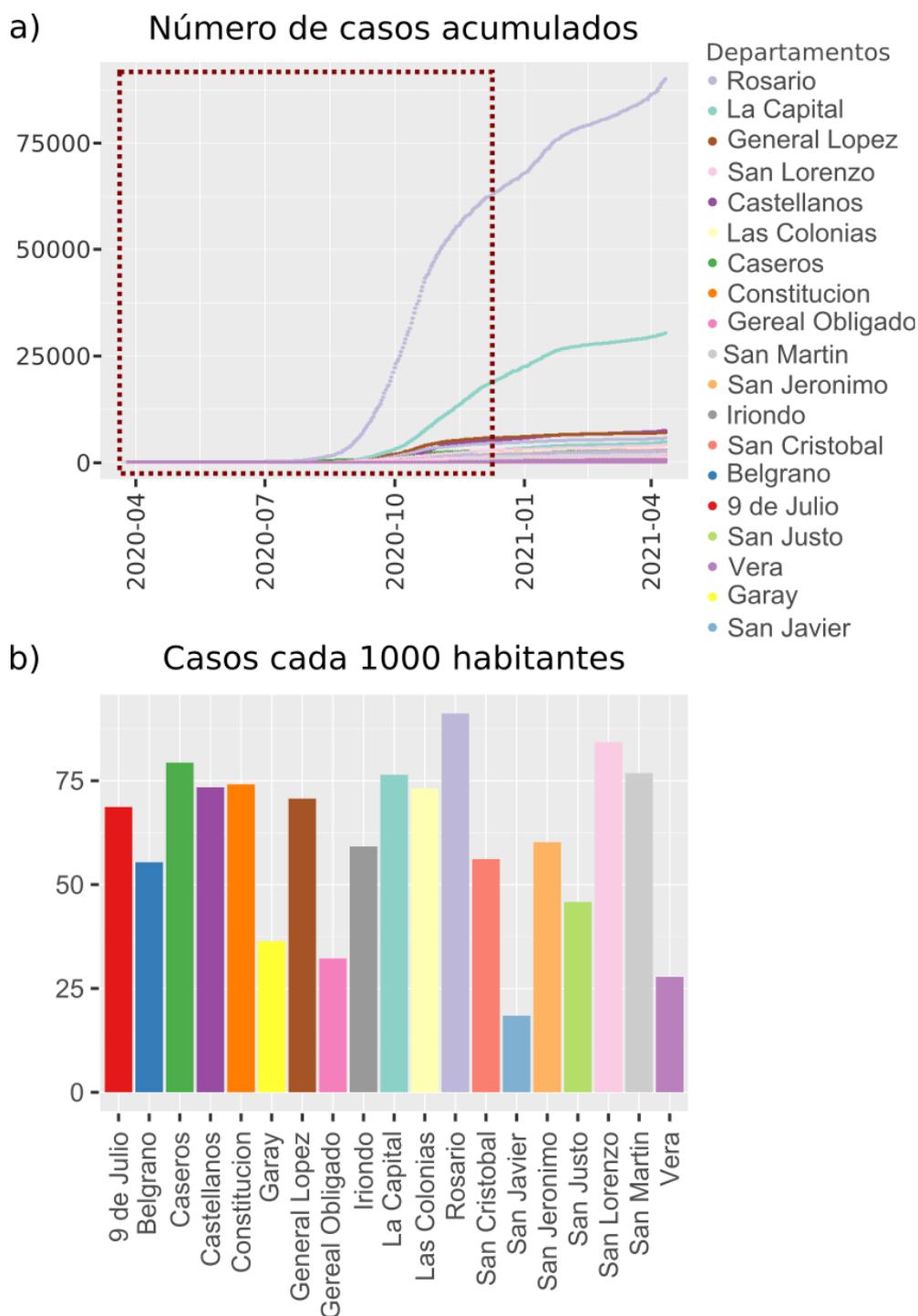


Figura 1: Casos positivos de la COVID-19 en la provincia de Santa Fe. a) Casos acumulados desde Abril de 2020 hasta Abril 2021, por Departamentos. La ventana temporal analizada se indica en el gráfico. b) Casos cada 1.000 habitantes, por Departamentos.

Caracterización genómica de SARS-CoV-2 de la provincia de Santa Fe

MUESTREO

Se incluyeron muestras de individuos residentes en 58 localidades con transmisión comunitaria en toda la provincia, entre los meses de marzo a diciembre de 2020, priorizando aquellas localidades con mayor cantidad de habitantes. Para el caso de la ciudad de Rosario, se seleccionaron muestras correspondientes a todos los barrios. La distribución de muestras seleccionadas según localidad fue la siguiente (Figura 2):

Santa Fe (51), Rosario (47), Rafaela (11), Santo Tomé (9), Esperanza (5), Coronda (4), Monte Vera (4), Villa Ocampo (4), Casilda (3), Helvecia (3), Hughes (3), Villa Constitución (3), Villa Gobernador Gálvez (3), Capitán Bermúdez (2), Ceres (2), Fighiera (2), Firmat (2), Funes (2), Laguna Paiva (2), Las Rosas (2), Maggiolo (2), Perez (2), Pueblo Esther (2), Puerto General San Martín (2), Ricardone (2), San Lorenzo (2), Soldini (2), Villa Amelia (2), Villa Cañas (2), Wheelwright (2), Zavalla (2), Acebal (1), Arroyo Seco (1), Avellaneda (1), Carcarañá (1), Carmen del Sauce (1), Cayastá (1), Colonia Margarita (1), Coronel Arnold (1), Fortin Olmos (1), Fray Luis Beltran (1), Fuentes (1), Los Molinos (1), Pavón Arriba (1), Reconquista (1), Rueda (1), San Carlos (1), San José de la Esquina (1), San Justo (1), Sancti Spiritu (1), Sanford (1), Santa Isabel (1), Sunchales (1), Teodelina (1), Totoras (1), Venado Tuerto (1) y Villada (1).

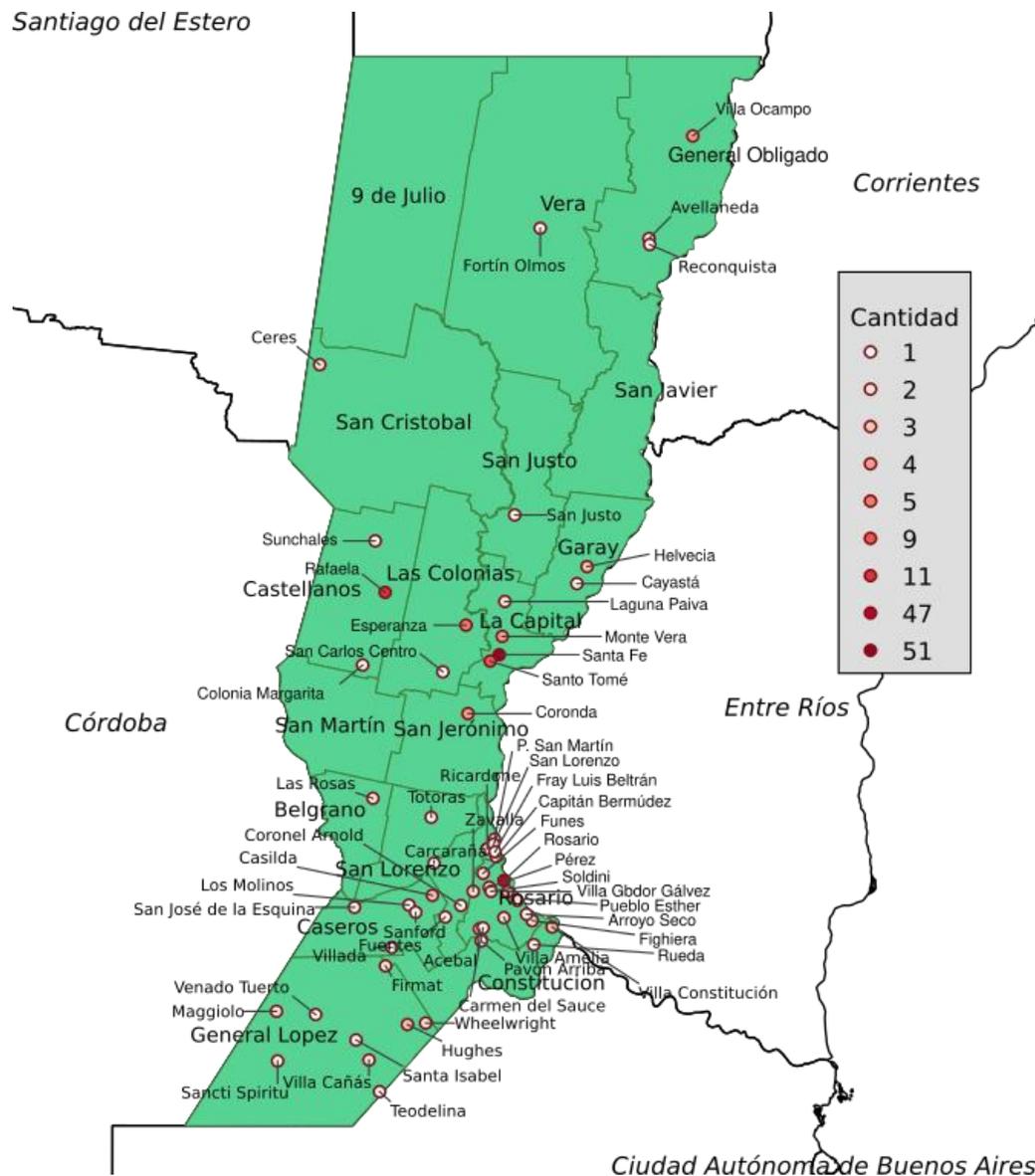


Figura 2. Distribución de localidades seleccionadas en la provincia de Santa Fe. Todas las muestras que se analizaron corresponden a extractos de ARN obtenidos de hisopados nasofaríngeos confirmados para la COVID-19 por la técnica de referencia de RT-PCR en tiempo real y que cumplían con el criterio de una carga viral medida en valores de Ct por debajo de 30. Estos criterios de selección pueden implicar una subestimación de la diversidad de linajes, sobre todo de los casos que son independientes entre sí y no tienen ningún nexo con *clusters* epidemiológicos identificados.

Estrategia de secuenciación empleada:

La secuenciación fue realizada según dos estrategias distintas:

-Nodo Rosario: Las secuencias se obtuvieron con el protocolo de ArgenTAG: secuenciación de genomas completos por RT-PCR de amplicones de ~1500 pb (2 reacciones 12-plex por muestra). Los productos de RT-PCR multiplex se combinaron a su vez en grupos de 12, 48 o 96 muestras con códigos de barras ArgenTAG para su secuenciación conjunta en un secuenciador MinION de Oxford Nanopore (profundidad promedio 2839X).

-Nodo Rafaela: las secuencias fueron obtenidas utilizando el protocolo ARTIC (<https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye>), como se detalla en reportes anteriores.

RESULTADOS OBTENIDOS

Secuencias obtenidas

En total, se obtuvieron un total **212 secuencias genómicas de SARS-CoV-2** provenientes de muestras clínicas de pacientes con la COVID-19.

Distribución de linajes encontrados en el periodo analizado

Las secuencias del genoma de SARS-CoV-2 obtenidas a partir de individuos infectados de distintos países han sido clasificadas según una taxonomía dinámica de linajes propuesta por el grupo del Dr. Rambaut, cuya última actualización corresponde al 01/04/2021 (Rambaut y col., 2020). Así, de acuerdo con cambios nucleotídicos y su agrupamiento filogenético, el SARS-CoV-2 se clasifica en varios linajes principales y numerosos linajes derivados, denominados con letras y números.

Para el análisis de las secuencias de Santa Fe, se realizó la asignación rápida de linajes mediante el programa Pangolin COVID-19 Lineage Assigner (<https://pangolin.cog-uk.io/>). Esta herramienta permite determinar el linaje a través de la identificación de patrones complejos mediante *machine learning*.

Se realizó la confirmación de linajes asignados mediante análisis filogenético, como se describe en el Anexo 1. El análisis e interpretación de los resultados fue realizado por los integrantes del Nodo de Evolución y los grupos de trabajo de la provincia de Santa Fe.

Mediante análisis filogenético se determinó que todas las secuencias de la provincia de Santa Fe pertenecieron al linaje B.1 o sus (sub)linajes derivados, que posee amplia distribución tanto a nivel mundial como en Argentina.

En mayor detalle, se encontraron representantes de los linajes B.1.499 (119 secuencias: 56,1 %), N.5 (75 secuencias: 35,4 %), N.3 (3 secuencias: 1,4 %), B.1 (10 secuencias: 4,7 %), B.1.1 (3 secuencias: 1,4 %) y B.1.1.28 (2 secuencias: 0,9 %).

En el gráfico de la Figura 3 se muestra la distribución de linajes hallados en la provincia de Santa Fe hasta el momento. La Figura 4 muestra la distribución de linajes identificados por quincena en el período analizado. Se observa una mayor frecuencia del linaje B.1.499 en los primeros meses del periodo muestreado (Junio a Octubre), el cual podría haberse desplazado por el linaje N.5, detectado mayormente en los meses de Noviembre y Diciembre 2020.

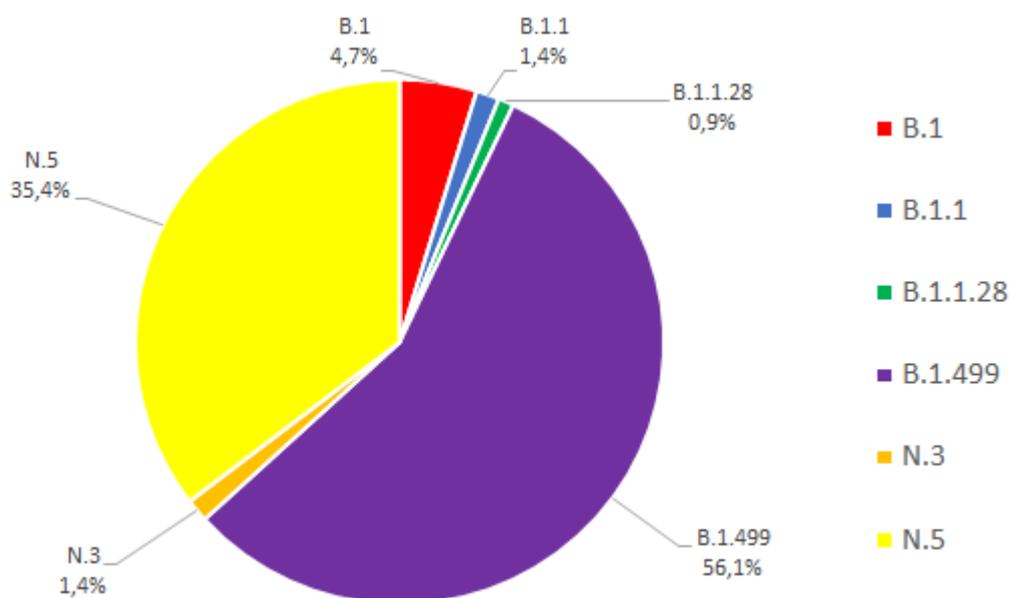


Figura 3 . Distribución de Linajes de SARS-CoV-2 hallados en la Provincia de Santa Fe (período marzo a diciembre 2020).

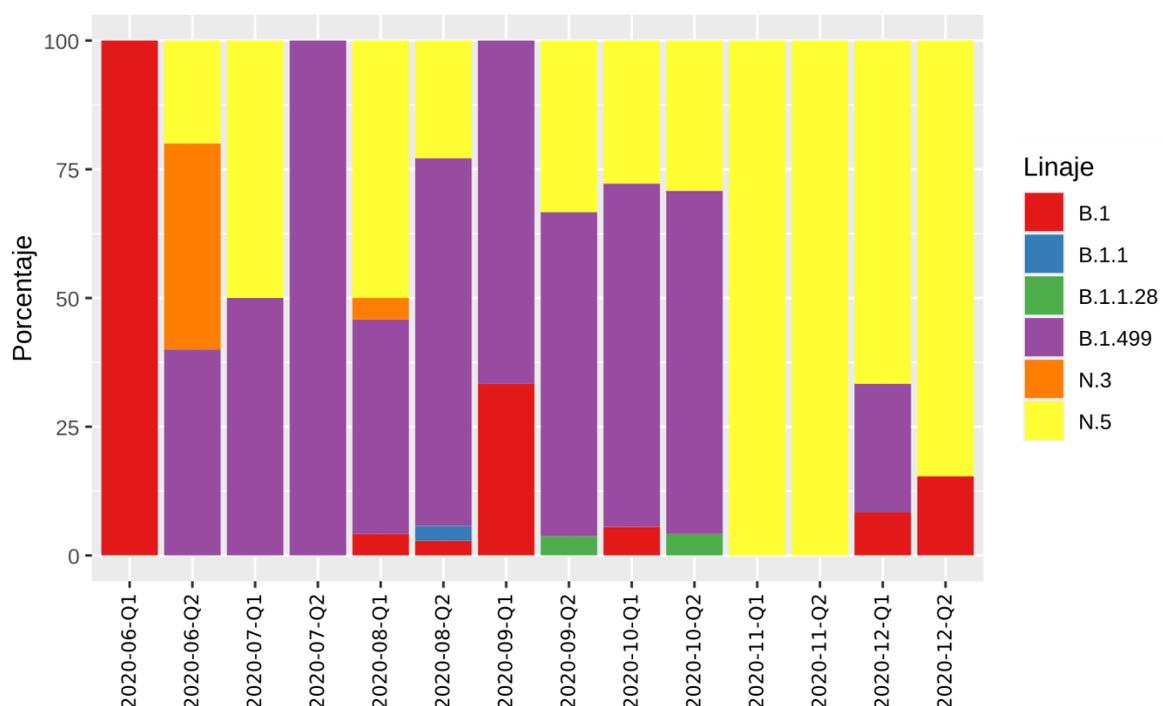


Figura 4: Distribución de los linajes observados en la provincia de Santa Fe, por quincena. Se muestran los datos obtenidos a partir de Junio 2020 donde se produce el aumento de casos en la provincia.

Análisis filogenético de linajes individuales

Linaje B.1

El linaje B.1 se encontró en el 4,7% de las secuencias incluidas en el análisis, para el que se observaron al menos cinco introducciones independientes (Figura 5). Este linaje fue uno de los principales causantes de los brotes en Europa y América del Norte. Si bien las diez secuencias B.1 de Santa Fe se relacionan filogenéticamente con secuencias de distintos países de los cinco continentes, en otros casos algunas formaron grupos que podrían representar clusters de circulación local dentro de la provincia. Al respecto, un grupo de secuencias con alto soporte estadístico incluyen a las localidades Reconquista y Villa Ocampo, en el norte de Santa Fe (clado D). Estas secuencias se agrupan con alto soporte estadístico con secuencias de Resistencia (provincia de Chaco), muy cercanas geográficamente.

Linaje B.1.1

El linaje B.1.1 se encontró en 3 de las secuencias analizadas, las cuales corresponden a la ciudad de Santa Fe (Figura 6). Una de ellas es una secuencia del 23/03/2020, que constituye uno de los primeros casos de la COVID-19 diagnosticados desde el inicio de la pandemia, en Argentina y en la provincia de Santa Fe. Corresponde a un caso importado por viaje a Europa. Por su parte, las otras dos secuencias encontradas corresponden al mes de agosto y conforman un clado monofilético con alto soporte estadístico junto con secuencias de la provincia de Buenos Aires, evidenciando una posible circulación de estos virus entre ambas provincias.

Linaje N.5

El linaje N.5 (alias B.1.1.33.5) surge del B.1.1.33, de origen brasilero. Es reconocido como un linaje argentino (75% de las secuencias reportadas a nivel global para este linaje, <https://cov-lineages.org/lineages>), y además se halló en Estados Unidos, Reino Unido, España y Chile. De las secuencias analizadas, 75 se clasificaron en este sublinaje. La mayoría de estas secuencias se distribuyeron en dos grandes clados (Figura 7). Uno de ellos es un grupo monofilético que incluye secuencias del norte de Santa Fe y secuencias de la ciudad de San Francisco, Córdoba, de cercanía geográfica (clado C). El otro clado, agrupa la mayoría de las secuencias de Santa Fe de este linaje con otras secuencias de Argentina (clado F). A su vez dentro del mismo, un subclado de alto soporte estadístico incluye enteramente secuencias del sur de la provincia, consistente con transmisión viral en regiones circunscritas. Las demás secuencias se agrupan por fuera de estos dos grandes clados, con otras localidades de Buenos Aires (clados D, E, G).

Linaje N.3

El linaje N.3 (alias B.1.1.33.3), detectado por primera vez en abril de 2020, surge del B.1.1.33 al establecerse este último en la Argentina. Es reconocido como un linaje argentino, ya que casi la totalidad de las secuencias reportadas en base de datos provienen de nuestro país (n=84), con excepción de un reporte de Chile y uno de Hong-Kong. De las secuencias analizadas, se hallaron tres correspondientes a este sublinaje, pertenecientes a la ciudad de Ceres y Santa Fe (Figura 7). Se observa que conforman un clado monofilético de alto soporte que incluye secuencias de Córdoba y Buenos Aires desde abril 2020 a enero 2021 (clado A). Por su parte las dos muestras de Ceres se agrupan juntas con alto soporte, indicativo de que estas muestras pertenecen a la misma cadena de transmisión (clado B).

Linaje B.1.499

Este linaje, previamente B.1.3.2, de origen argentino y detectado inicialmente en marzo del 2020, ha sido reportado en diferentes provincias del país, principalmente en la región del AMBA. El mismo fue detectado en la mayoría de las secuencias analizadas, representando el 56% de las mismas. Se observan 6 grupos monofiléticos de alto soporte estadístico, cada uno conformado mayoritariamente por secuencias del Norte o del Sur de la provincia. Las demás secuencias se encuentran asociadas de forma independiente con otras secuencias de Argentina, lo cual sugiere múltiples introducciones de este linaje a la provincia, posiblemente desde distintas regiones del país (Figura 8).

Linaje B.1.1.28

Este linaje es mayoritariamente brasilero y se encuentra en circulación desde marzo 2020. A la fecha sólo se han reportado 3 secuencias pertenecientes a este linaje en la Argentina (Figura 9). De las secuencias analizadas, se encontraron dos secuencias correspondientes a este linaje correspondientes a las ciudades de Rueda y Pueblo Esther, en los meses de septiembre y octubre respectivamente. Las mismas se agrupan con alto soporte estadístico junto con otras secuencias brasileras. Sin embargo, ninguno de los casos reportó antecedentes de viaje. Cabe destacar que este linaje es el que da origen a los linajes P.1 (variante 501Y.V3) en Manaus y P.2 en Río de Janeiro, que constituyen variantes de preocupación e interés, respectivamente.

Linaje B.1

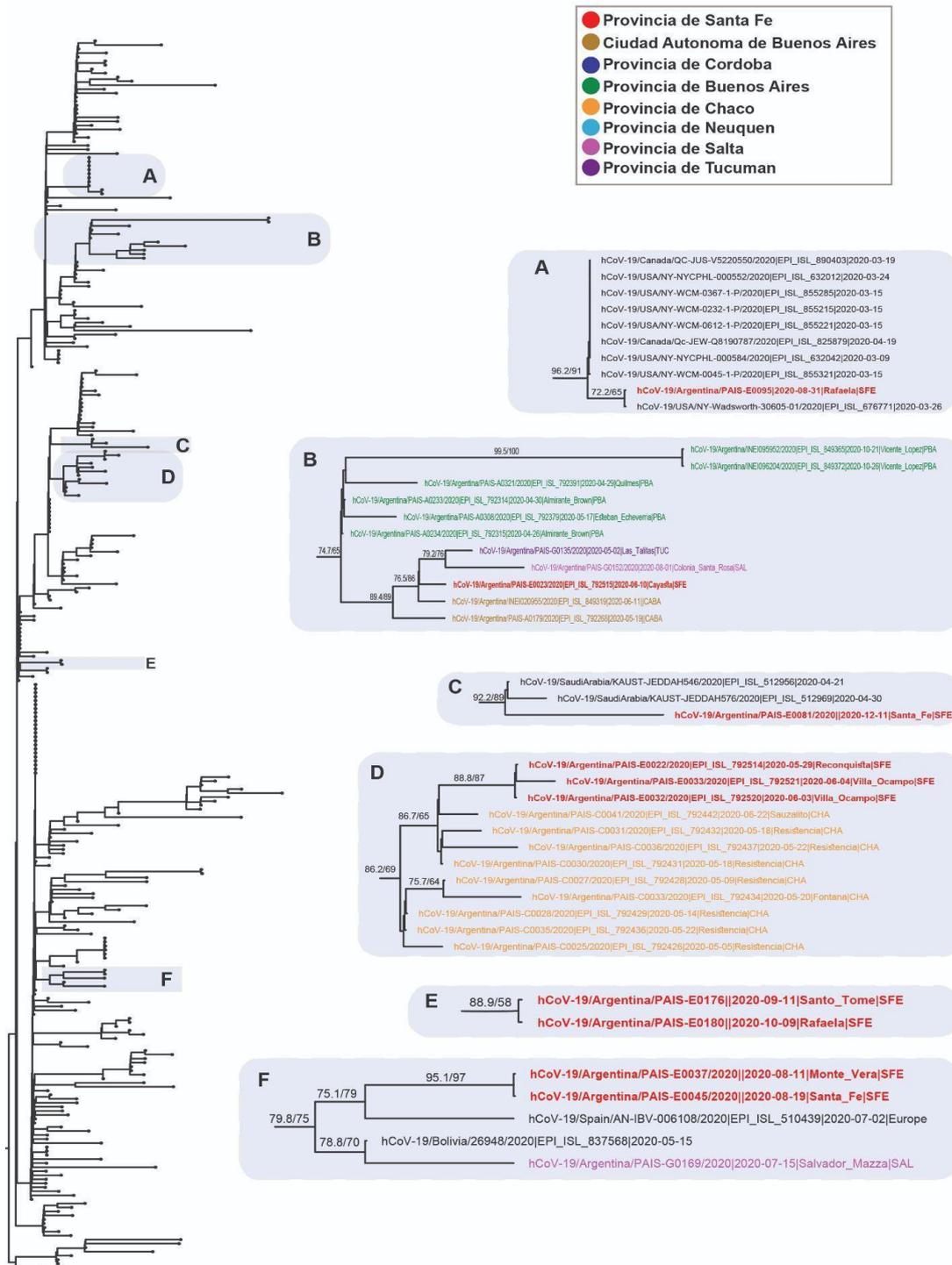


Figura 5. Árbol filogenético del linaje B.1 de SARS-CoV-2.

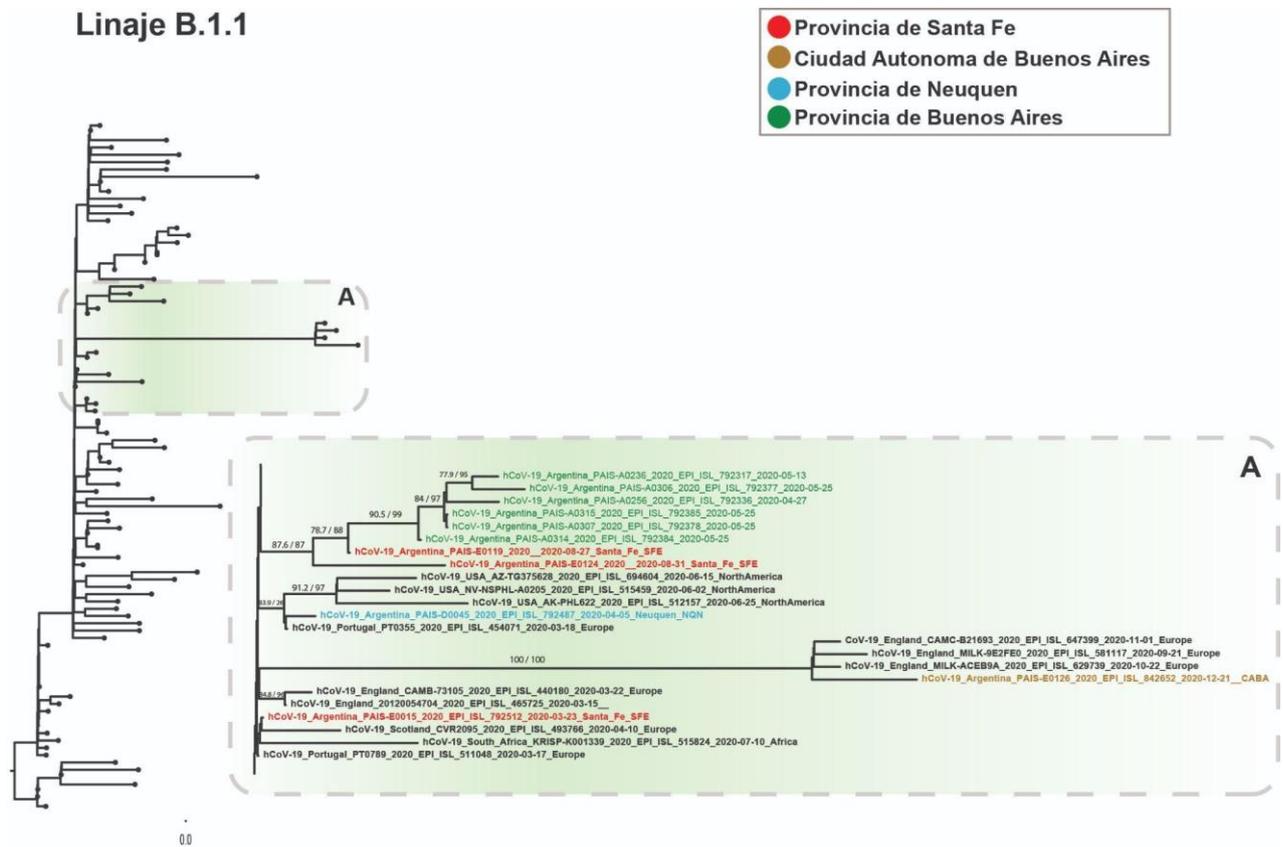
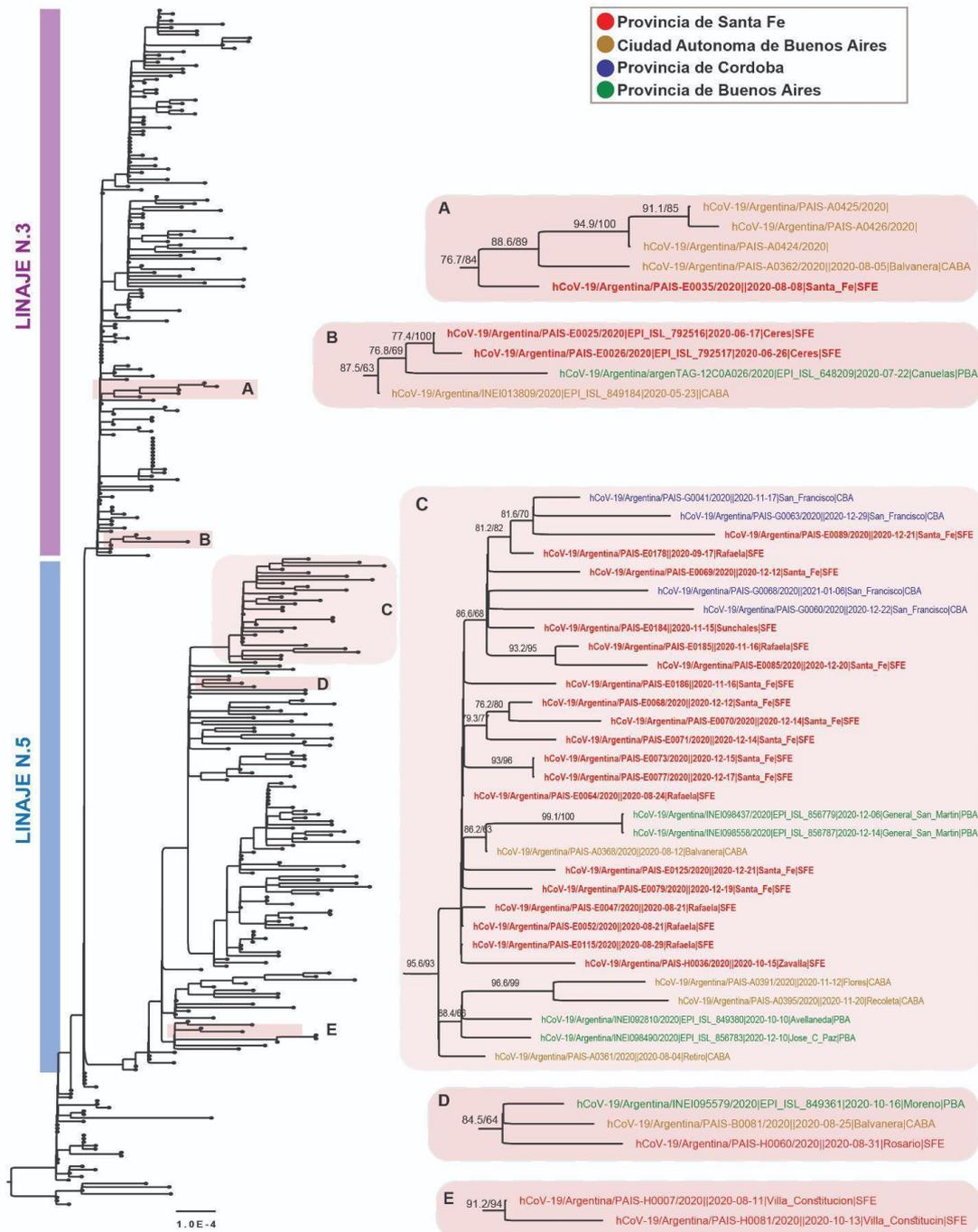


Figura 6. Árbol filogenético del linaje B.1.1 de SARS-CoV-2.

Proyecto Argentino Interinstitucional de genómica de SARS-CoV-2

Linajes N.3 y N.5 - parte 1



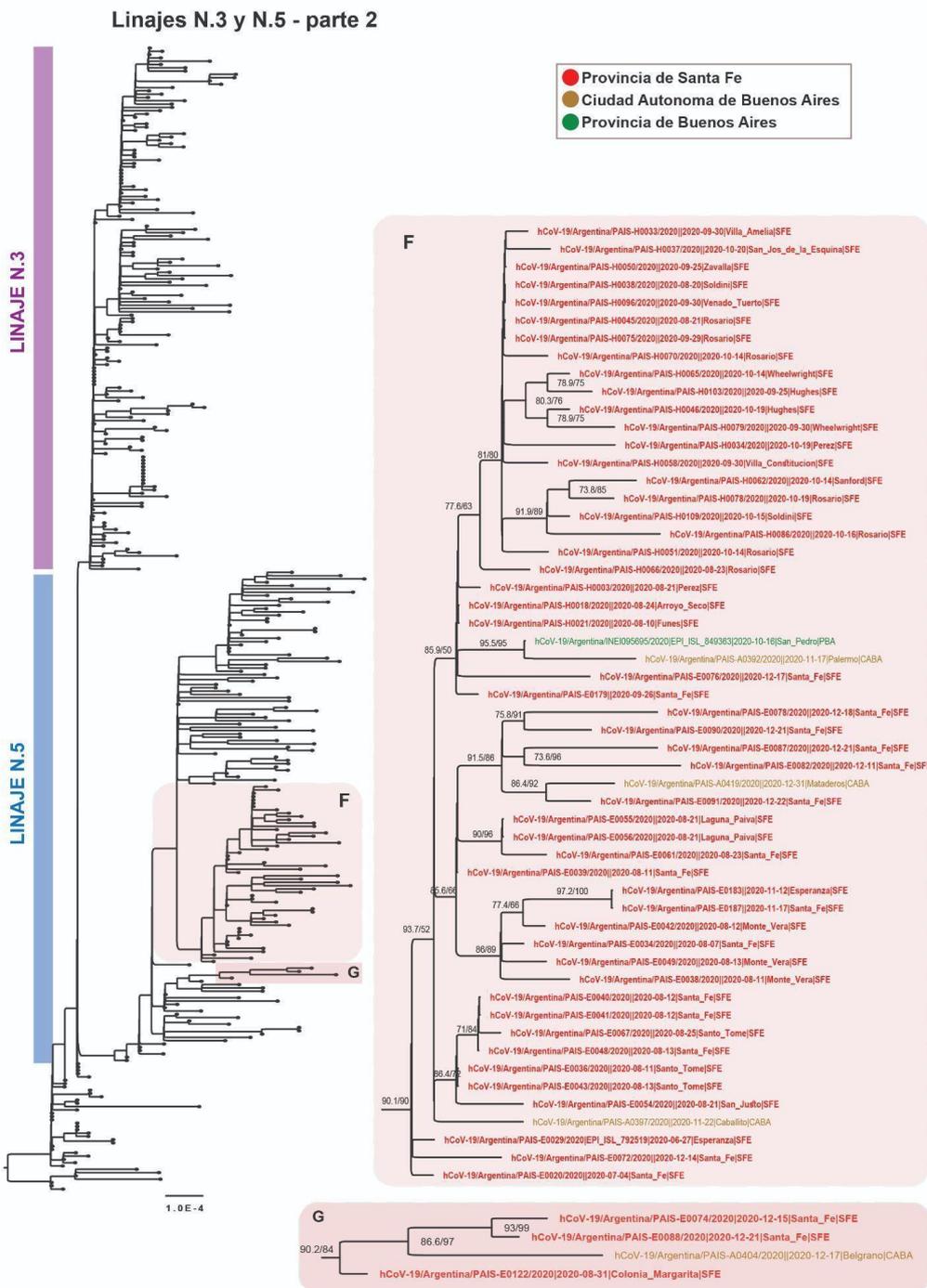
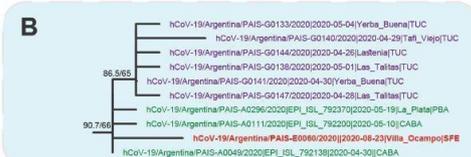
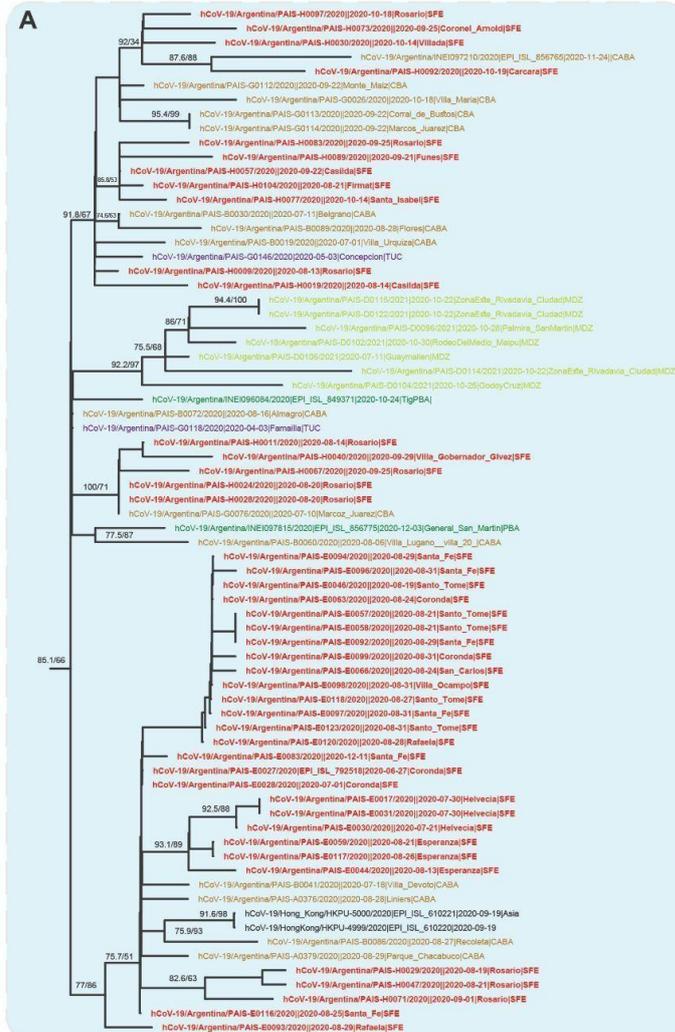
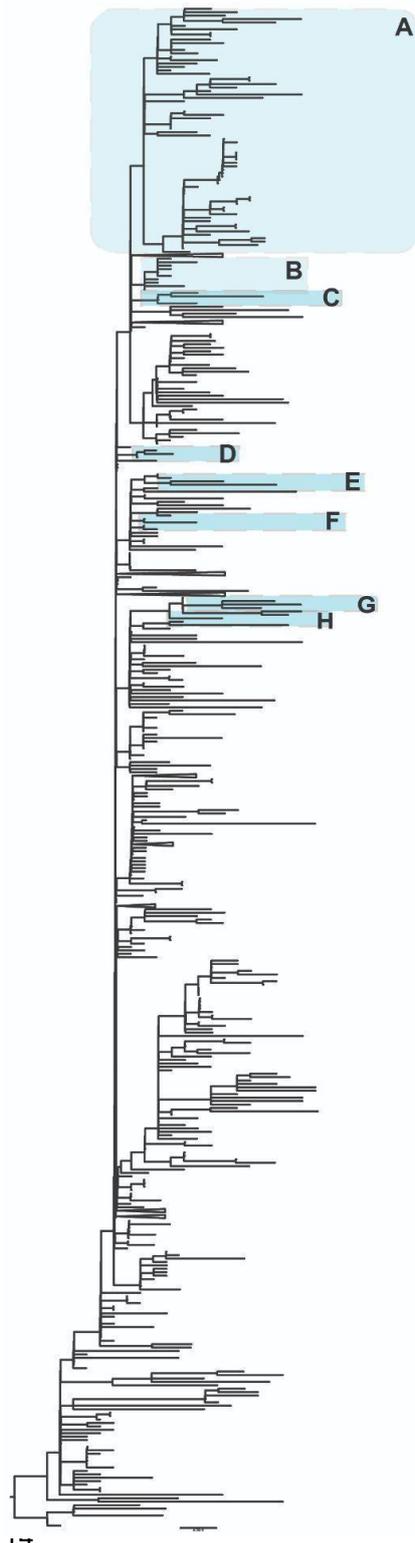


Figura 7. Árbol filogenético de los linajes N.5 y N.3 de SARS-CoV-2.

Proyecto Argentino Interinstitucional de genómica
de SARS-CoV-2

Linaje B.1.499 - parte 1



Proyecto Argentino Interinstitucional de genómica
de SARS-CoV-2

Linaje B.1.499 - parte 2

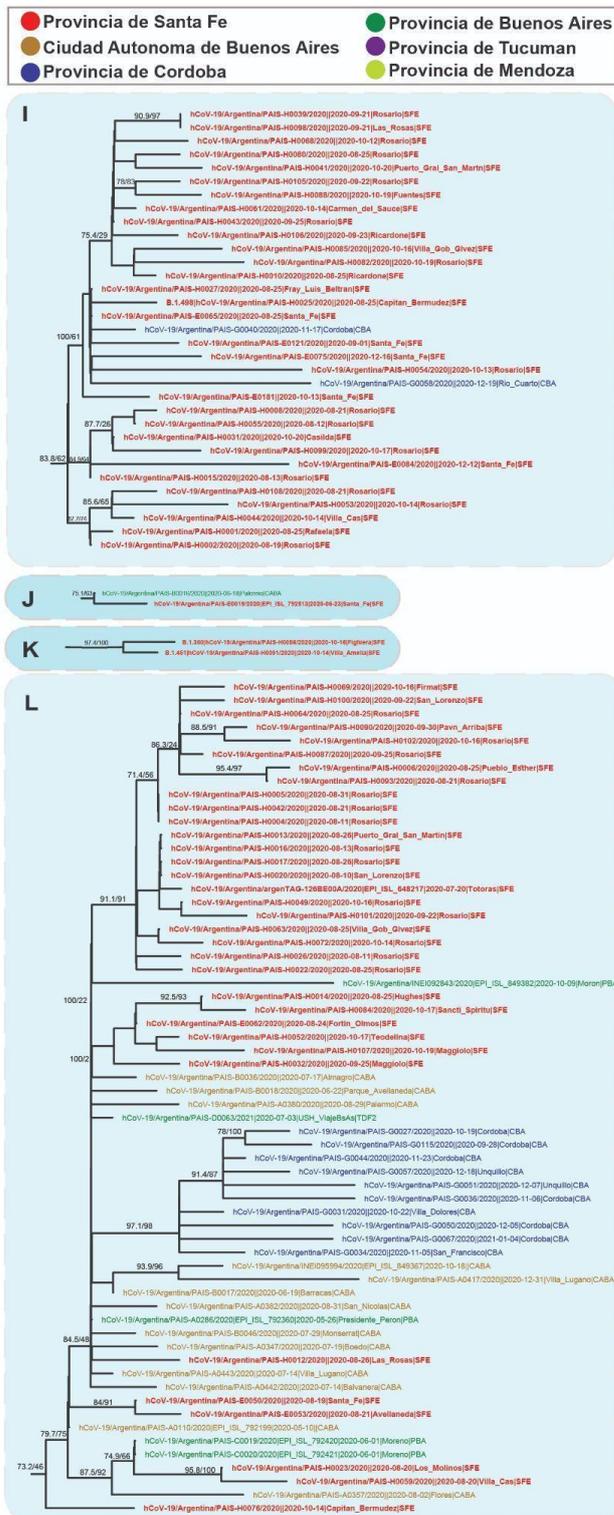


Figura 8. Árbol filogenético del linaje B.1.499 de SARS-CoV-2.

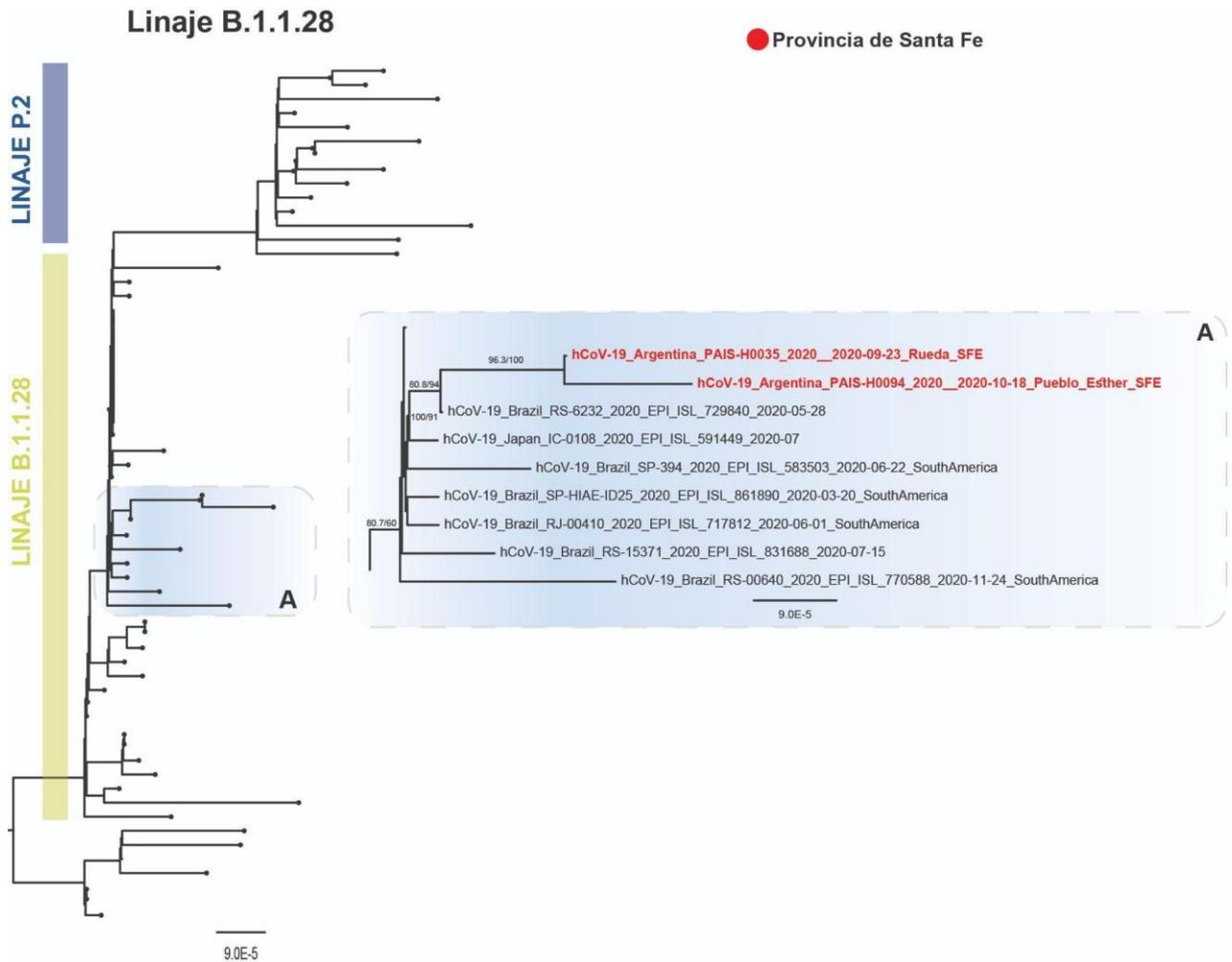


Figura 9. Árbol filogenético del linaje B.1.1.28 de SARS-CoV-2 y su linaje derivado P.2.

Análisis de mutaciones que podrían afectar la detección de los métodos diagnósticos moleculares disponibles en el mercado.

Dado que existe gran diversidad de métodos moleculares disponibles en el país para el diagnóstico de la infección por SARS-CoV-2, cada uno con sondas y cebadores específicos localizados en diferentes regiones del genoma, se analizó la presencia de mutaciones en dichas regiones en los genomas de SARS-CoV-2 obtenidos en la provincia de Santa Fe y reportados en este informe. Este análisis se pudo realizar sólo para aquellos ensayos que tienen la información disponible. Cabe aclarar que la presencia de estas mutaciones no implica necesariamente una disminución en la sensibilidad de dichos ensayos.

Se analizaron los cebadores y sondas de los siguientes ensayos (ver referencias en Anexo 1):

- Protocolo de Charité, Berlín. Genes reconocidos: E, N y RdRp.
- Protocolo de The University of Hong Kong. Genes reconocidos: N y RdRp.
- Protocolo de CDC Estados Unidos. Gen reconocido: N (regiones N1 y N2).
- Protocolo de CDC China. Genes reconocidos: ORF1ab y N
- Protocolo de Ministry of Public Health, Tailandia. Gen reconocido: N
- Protocolo de National Institute of Infectious Diseases, Tokio, Japón. Gen reconocido: N
- Protocolo de Institut Pasteur, Paris. Genes reconocidos: E y RdRp.
- Protocolo de KiCqStart de Sigma-Aldrich. Gen reconocido: S

Se reconocieron las secuencias nucleotídicas de los cebadores y sondas utilizando el algoritmo BLASTn a través de la aplicación de un *script* desarrollado en el nodo bioinformático del Consorcio. Se registraron las mutaciones encontradas en los genomas de Santa Fe, cuya ubicación es reportada en este informe con respecto a la secuencia de referencia hCoV-19/Wuhan/WIV04/2019:

- Gen RdRp:
 - Se detectó un genoma con la mutación a12719g, donde se ubica la región de detección de la sonda del gen RdRp IP2 del protocolo del Instituto Pasteur.
 - Se detectaron 6 genomas con la mutación c12723t, donde se ubica la región de detección de la sonda del gen RdRp IP2 del protocolo del Instituto Pasteur.
 - Se detectó un genoma con la mutación t12727c, donde se ubica la región de detección de la sonda del gen RdRp del protocolo del Instituto Pasteur.
 - Se detectó un genoma con la mutación c12729t, donde se ubica la región de detección de la sonda del gen RdRp del protocolo del Instituto Pasteur.
 - Se detectaron 4 genomas con la mutación g13445t, donde se ubica la región de detección del cebador antisentido del gen ORF1ab del protocolo del CDC de China.
 - Se detectaron 34 genomas con la mutación c14120t, donde se ubica la región de detección de la sonda del gen RdRp IP4 del protocolo del Instituto Pasteur.

Proyecto Argentino Interinstitucional de genómica
de SARS-CoV-2

- Se detectó un genoma con la mutación c14183t, donde se ubica la región de detección del cebador antisentido del gen RdRp IP4 del protocolo del Instituto Pasteur.
- Se detectó un genoma con la mutación t15507c, donde se ubica la región de detección del cebador antisentido del gen RdRp del protocolo de Charité, Berlín.

- Gen S:

- Se detectó un genoma con la mutación g23587t, donde se ubica la región de detección del cebador sentido del gen S del protocolo Sigma.
- Se detectó un genoma con la mutación c23613t, donde se ubica la región de detección de la sonda del gen S del protocolo Sigma.

- Gen E:

- Se detectó un genoma con la mutación c26370t, donde se ubica la región de detección del cebador antisentido del gen E del protocolo del Instituto Pasteur y de Charité, Berlín.

- Gen N:

- Se detectó un genoma con la mutación g28321t, donde se ubica la región de detección del cebador sentido del gen N del protocolo de Tailandia y del CDC de China.
- Se detectó un genoma con la mutación c28335t, donde se ubica la región de detección del cebador sentido del gen N del protocolo de Tailandia.
- Se detectó un genoma con la mutación g28362t, donde se ubica la región de detección del cebador antisentido del gen N del protocolo de Tailandia.
- Se detectó un genoma con la mutación a28366g, donde se ubica la región de detección del cebador antisentido del gen N del protocolo de Tailandia.
- Se detectó un genoma con la mutación c28775t, donde se ubica la región de detección de la sonda del gen N del protocolo Charité, Berlín.
- Se detectó un genoma con la mutación c28720t, donde se ubica la región de detección del cebador sentido del gen N del protocolo Charité, Berlín.
- Se detectó genoma con la mutación t28825g, donde se ubica la región de detección del cebador antisentido del gen N del protocolo Charité, Berlín.
- Se detectaron 79 genomas con la mutación ggg28881aac, donde se ubica la región de detección del cebador sentido del gen N del protocolo del CDC de China.
- Se detectaron 6 genomas con la mutación c28887t, donde se ubica la región de detección del cebador sentido del protocolo del CDC de China.
- Se detectó un genoma con la mutación g28960t, donde se ubica la región de detección del cebador antisentido del protocolo del CDC de China.

Proyecto Argentino Interinstitucional de genómica
de SARS-CoV-2

- Se detectó un genoma con la mutación a28967g, donde se ubica la región de detección del cebador antisentido del protocolo del CDC de China.
- Se detectaron dos genomas con la mutación g28968t, donde se ubica la región de detección del cebador antisentido del protocolo del CDC de China.
- Se detectaron 72 genomas con la mutación g29134a, donde se ubica la región de detección del cebador sentido del gen N del protocolo de Japón.
- Se detectaron 68 genomas con la mutación c29272t, donde se ubica la región de detección del cebador antisentido del gen N del protocolo de Japón.
- Se detectó un genoma con la mutación g29179t, donde se ubica la región de detección del cebador sentido del protocolo del CDC de Estados Unidos.
- Se detectaron dos genomas con la mutación a28182g, donde se ubica la región de detección de la sonda del protocolo del CDC de Estados Unidos y de la Universidad de Hong Kong.
- Se detectaron 50 genomas con la mutación t29148c, donde se ubica la región de detección del cebador sentido del gen N del protocolo de la Universidad de Hong Kong.

Detalle de las sustituciones no sinónimas (cambios aminoacídicos) en las proteínas codificadas en los genomas obtenidos a partir de las muestras de Santa Fe.

Para determinar las sustituciones no sinónimas en comparación con la secuencia de referencia hCoV-19/Wuhan/WIV04/2019, se utilizó la herramienta CoVsurver y se corroboraron mediante inspección visual. Los cambios fueron comparados con los reportados en secuencias de SARS-CoV-2 a nivel mundial utilizando la herramienta COV-GLUE de GISAID (Singer et al. 2020) y la herramienta outbreak.info.

En concordancia con la presencia exclusiva del linaje B.1 y sus derivados en las secuencias virales estudiadas, **todos** los genomas de SARS-CoV-2 poseen la mutación **Spike: D614G**, asociada a mayor transmisibilidad tanto en modelos *in vitro* como *in vivo* (Korber et al, 2020; Plante, et al. 2020). Asimismo, la mayoría de las secuencias (201/212) poseen la mutación **NSP12 (RdRp)/P323L** -previamente asociada con un potencial poder mutagénico mayor del SARS-CoV-2 por estudios bioinformáticos (Eskier, D. et al, 2020.; Chand, G.B., et. al. 2020).

Los análisis se focalizaron en el gen de la proteína de superficie Spike (S) del SARS-CoV-2 por su importancia biológica en la infección, transmisión, tratamiento e impacto en la vacunación.

Linaje B.1

Las 10 secuencias de Santo Tomé (1), Rafaela (2), Villa Ocampo (2), Santa Fe Capital (2), Cayastá (1), Monte Vera (1) y Reconquista (1) identificadas como B.1 poseen cambios descritos previamente en secuencias de SARS-CoV-2 a nivel mundial, incluyendo el cambio en la región Spike_D614G. En solo una de ellas (E0081) se identificó la mutación en Spike T778I. Además la secuencia E0023 de Cayastá presentó una mutación en Spike P812L.

Linaje B.1.1

Las 3 secuencias identificadas como B.1.1 son de Santa Fe Capital en las cuales se observaron cambios reportados previamente en secuencias de SARS-CoV-2, incluyendo el cambio en Spike D614G. A su vez, una de ellas (E118) posee la mutación en Spike Q414R, frecuente en secuencias de Argentina reportadas anteriormente.

Linaje B.1.1.28

En las dos secuencias de Rueda y Pueblo Esther identificadas como B.1.1.28 se observaron cambios descritos previamente en secuencias de SARS-CoV-2. Además, ambas comparten un cambio en el gen N S202I, reportado por primera vez a nivel mundial. En la secuencia de Rueda (H0035) se identificó el cambio en Spike V1176F, característico de las secuencias agrupadas en este linaje.

Linaje N.5

En las 75 secuencias de la provincia de Santa Fe identificadas como N.5 se observaron cambios descritos previamente en secuencias de SARS-CoV-2 a nivel mundial; 15 de ellos presentes en la región Spike, entre

ellas 3 identificadas en todas o la mayoría de las secuencias de Santa Fe: D614G (75); Q414R (75); T778I (70) y 12 presentes sólo en algunas de ellas (Tabla 1).

Tabla 1. Mutaciones poco frecuentes en secuencias de Santa Fe del linaje N.5 observadas en Spike

Mutación en Spike	Secuencias*	Localidades	Reportada en Argentina
G1219C	H0036, E0047, E0052, E0064, E0068 , E0069 , E0070 , E0071, E0073, E0077, E0079, E0085 , E0089 , E0115, E0125, E0178, E0184, E0185, E0186	Zavalla, Rafaela, Santa Fe Capital, Sunchales	si
I402V ¹	E0089	Santa Fe Capital	si
A829S	H0062, H0079	Rosario, Sandford	no
N1192T	E0068 , E0070	Santa Fe Capital	no
A647S	E0067	Santo Tomé	no
M153T	E0069	Santa Fe Capital	no
Q675H	E0087	Santa Fe Capital	no
M177I	E0089	Santa Fe Capital	no
A1020S	E0085	Santa Fe Capital	no
A1222S ²	H0046	Hughes	no
V11L ³	E0082	Santa Fe Capital	no
L5F	E0072	Santa Fe Capital	si

*En negrita se muestran las secuencias donde se identificó más de una mutación poco frecuente en Spike.

¹La mutación Spike_I402V es un cambio poco frecuente, y según un estudio se propone que este cambio podría estar asociado a un leve incremento de la unión al receptor ACE2 (Starr et al. 2021).

²La mutación Spike_A1222S se identificó sólo en dos secuencias de EE.UU y Suiza al 14/04/2021.

³El cambio Spike_V11L se observó en sólo 8 secuencias de SARS-CoV-2 a nivel mundial al 14/04/2021.

En una de las secuencias agrupada en este linaje (H007), proveniente de Villa Constitución, se identificó la mutación E59stop en el ORF8 resultando en una transcripción truncada de la proteína. Se ha observado que variantes virales con este tipo de mutaciones en ORF8 y con cambios en Spike serían más transmisibles y requerirían un monitoreo epidemiológico para el control de la pandemia (Pereira 2021).

Linaje N.3

En las 3 secuencias de Ceres (2) y Santa Fe Capital (1) pertenecientes al linaje N.3 se observaron cambios reportados previamente en secuencias de SARS-CoV-2 a nivel mundial, incluyendo en la región Spike el cambio D614G. Además, las 3 secuencias comparten un cambio en la región E_V58F que fue identificado previamente en secuencias de Argentina agrupadas en el linaje N.3.

Linaje B.1.499

En las secuencias identificadas como B.1.499 provenientes de varias localidades de Santa Fe se observan cambios reportados previamente en secuencias de SARS-CoV-2 a nivel mundial, incluyendo en la región Spike el cambio D614G, presente en todas ellas y el cambio en Spike A626S identificado en 32 de estas secuencias. Se identificaron además 16 mutaciones en Spike menos frecuentes y que se detallan en la Tabla 2.

Tabla 2: Mutaciones poco frecuentes en secuencias de Santa Fe del linaje B.1.499 observadas en Spike

Mutación en Spike	Secuencias*	Localidades	Reportado en Argentina
Spike_S161F ¹	H0082	Rosario	no
Spike_G181A ²	H0030 , H0077, H0089, H0057, H0104, H0083, H0073 , H0097 , H0092 , H0019	Villada, Casilda, Coronel Arnold, Santa Isabel, Rosario, Funes, Carcarañá, Firmat	no
Spike_F833Y ³	E0084	Santa Fe Capital	no

Proyecto Argentino Interinstitucional de genómica
de SARS-CoV-2

Spike_F338L	H0012	Las Rosas	no
Spike_F32L	E0019	Santa Fe Capital	no
Spike_V1228L	H0088	Fuentes	si
Spike_T95I	H0039, H0098	Las Rosas, Rosario	si
Spike_S813I	H0090	Pavon Arriba	no
Spike_P1162L ⁴	H0042, H0049, H0063, H0064, H0069, H0072, H0087, H0090 , H0093, H0100, H0101, H0102	Rosario, Villa G. Galvez, Firmat, Pavon Arriba, San Lorenzo	si
Spike_E654Q	E0056, H0048	Acebal, Santa Fe Capital	no
Spike_D253G	H0030, H0097, H0092, H0073 , E0086	Villada, Coronel Arnold, Carcarañá, Rosario, Santa Fe Capital	si
Spike_A706V	H001	Rafaela	no
Spike_A684V	H0085	Villa Gobernador Galvez	si
Spike_A67V	H0076	Capitán Bermúdez	si
Spike_A27V	H0052	Teodolina	no

Spike_A1078V	H0095	Figuera	no
--------------	-------	---------	----

*En negrita se muestran las secuencias donde se identificó más de una mutación poco frecuente en Spike.

¹La mutación Spike_S161F no se encuentra reportada en otras secuencias de SARS-CoV-2 a nivel global al 15/04/21, siendo la primera vez que se reporta este cambio.

²Las secuencias con la mutación en Spike G181A forman un clado junto con otras secuencias de Argentina con alto soporte estadístico en la filogenia del linaje B.1.499 (Figura 8).

³La mutación Spike F833Y ha sido reportada en una sólo secuencia de EE.UU al 15/04/2021.

⁴En las secuencias con la mutación en Spike P1162L también se identificó el cambio Spike A626S y se agrupan en un clado con alto soporte estadístico en la filogenia del linaje B.1.499 (Figura 8).

CONCLUSIONES GENERALES

En este informe se reporta el análisis detallado de 212 secuencias de SARS-CoV-2 provenientes de muestras clínicas obtenidas de pacientes de la provincia de Santa Fe diagnosticados con la COVID-19 durante el período comprendido entre el 23/3/2020 al 22/12/2020 (SE 13/2020 a SE 52/2020).

Todas las secuencias corresponden al linaje B.1 o sus derivados, lo que resulta compatible con la amplia distribución que presenta este linaje a nivel mundial y es congruente con lo observado para otras secuencias de Argentina y para la mayoría de los países de América del Sur. Cabe destacar que el linaje B.1 fue detectado mayormente en muestras de Junio 2020, indicando que podría ser el inicio de la cadena de transmisión en la provincia (con ciertas introducciones posteriores) para luego ser desplazado por los demás linajes.

Se observó una alta frecuencia de secuencias agrupadas en los linajes N.5 y B.1.499. En particular, el linaje B.1.499 es de amplia distribución en el país y resultó mayoritario en los primeros meses del periodo muestreado, luego su circulación podría haberse desplazado por el linaje N.5, detectado en alta frecuencia en los meses de Noviembre y Diciembre 2020. Las secuencias de Santa Fe corresponden al 60% de las secuencias de este linaje en el país. Asimismo, dentro de estos linajes se identificaron clusters de secuencias con estrecha cercanía geográfica, manifestando las cadenas de transmisión virales dentro de la provincia de Santa Fe y/o con las provincias vecinas (Chaco, Córdoba y Buenos Aires).

Se identificaron 2 secuencias del linaje B.1.1.28, relacionadas entre sí y con secuencias de Brasil en pacientes sin antecedentes de viaje a ese país, y al momento (04/2021) sólo 5 secuencias de este linaje fueron reportadas en Argentina, evidenciando su baja circulación en la población.

Se detectaron algunos cambios nucleotídicos en zonas de reconocimiento utilizadas por métodos de diagnóstico molecular de los que se dispone información, sin embargo, la presencia de estas mutaciones no implica necesariamente una disminución en la sensibilidad de dichos métodos.

Por último, se encontraron algunos cambios aminoacídicos (sustituciones no sinónimas) en distintas regiones del genoma viral, algunos de ellos en la región que codifica para la proteína Spike en las secuencias analizadas. Se observaron diferentes combinaciones de mutaciones en esta región, la mayoría de ellas reportadas anteriormente a nivel global y que se corresponden con la clasificación de los diferentes linajes detectados. Se destaca la mutación Spike_S161F en una secuencia de Rosario, reportada por primera vez a nivel global.

Participantes en este reporte:

Nodo de toma, procesamiento de muestras clínicas y análisis epidemiológicos:

Centro de Tecnología en Salud Pública. Facultad de Ciencias Bioquímicas y Farmacéuticas y Hospital Provincial del Centenario, Rosario: Ana Laura Cavatorta, Julián Acosta, Eduardo Codino, Agustina Cerri.

IDICER-CONICET/UNR y Facultad de Ciencias Médicas (UNR): Silvana Spinelli.

Laboratorios privados de la ciudad de Rosario: SILAB (Juan Pablo Scapini), Hospital Español (Larisa Altube, Daniel Ezpeleta).

Laboratorio Central de Santa Fe: Carlos Pastor, Guillermo Ojeda, Gabriela Rompato, Viviana Mugna.

Nodo de secuenciación de Rosario:

Laboratorio Mixto de Biotecnología Acuática, Facultad de Ciencias Bioquímicas y Farmacéuticas (UNR) y Secretaría de Ciencia, Tecnología e Innovación de Santa Fe. Silvia Arranz, Vanina Villanova y Victoria Posner.

Proyecto Argentino Interinstitucional de genómica
de SARS-CoV-2

CIFASIS-CONICET/UNR y Facultad de Ciencias Exactas, Ingeniería y Agrimensura (UNR). Elizabeth Tapia, Pilar Bulacio, Joaquín Ezpeleta, Ignacio Garcia Labari, Flavio Spetale, Javier Murillo, Laura Angelone y Sofía Lavista Llanos.

Nodo de secuenciación de Rafaela: Instituto de Investigación de la Cadena Láctea (IDICAL, INTA-CONICET): María Florencia Eberhardt, Matías Irazoqui, Ariel Amadio.

Nodo evolución: Carolina Torres, Débora Marcone, Laura Mojsiejczuk, María Dolores Blanco Fernández, Andrés Culasso, Humberto Debat, Mariana Viegas.

Grupo “Virología Humana” del IBR-CONICET/UNR y Facultad de Ciencias Bioquímicas y Farmacéuticas (UNR). Adriana Giri, Pablo Casal, Elisa Bolatti, Agustina Cerri, Diego Chouhy, María Florencia Re y Gastón Viarengo.

Financiamiento:

Proyecto IP COVID-19 N°08, Focem COF 03/11 COVID-19, Covid Federal (SF-10), Covid Federal (SF-11), Covid Federal (SF-12) y ArgenTag S.A.S.

ANEXO 1

Análisis filogenéticos

Las secuencias de Argentina fueron analizadas junto con secuencias de referencia de distintos linajes (<https://github.com/hCoV-2019/pangolin>) y secuencias con mejor score de alineamiento por BLAST (entre 10 hits por secuencia incógnita, contra la base de datos de GISAID al 31/01/2021). Los alineamientos se construyeron con el programa MAFFT (<https://mafft.cbrc.jp/alignment/server/>) con parámetros por *default*. El modelo evolutivo apropiado se seleccionó con ModelFinder (Kalyaanamoorthy et al., 2017, <https://doi.org/10.1038/nmeth.4285>) de acuerdo con criterio de Información Bayesiano y los análisis filogenéticos se realizaron por Máxima Verosimilitud con el programa IQ-TREE v. 2.1.2 COVID-edition (Minh y col., 2020, <https://doi.org/10.1093/molbev/msaa015>). Se utilizaron los métodos de *SH-like approximate likelihood ratio test* (1000 réplicas) (Guindon et al., 2010, <https://doi.org/10.1093/sysbio/syq010>) y *Ultrafast bootstrap Approximation* (10000 réplicas) (Hoang y col., 2018, <https://doi.org/10.1093/molbev/msx281>) como métodos para evaluar la confiabilidad de los grupos obtenidos. En las ramas de los árboles (Figuras 5-9) se indican los valores de SH-like/UFB para los grupos relevantes.

Análisis de mutaciones que podrían afectar la detección de los ensayos diagnósticos moleculares disponibles en el mercado.

Se analizaron los cebadores y sondas de los siguientes test:

- Protocolo de Charité, Berlín. Genes reconocidos: E, N y RdRp (https://www.who.int/docs/default-source/coronaviruse/protocol-v2-1.pdf?sfvrsn=a9ef618c_2)
- Protocolo de The University of Hong Kong. Genes reconocidos: N y RdRp. (https://www.who.int/docs/default-source/coronaviruse/peiris-protocol-16-1-20.pdf?sfvrsn=af1aac73_4)
- Protocolo de CDC Estados Unidos. Gen reconocido: N (regiones N1 y N2) (<https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html>)
- Protocolo de CDC China. Genes reconocidos: ORF1ab y N (http://ivdc.chinacdc.cn/kyjz/202001/t20200121_211337.html)
- Protocolo de Ministry of Public Health, Tailandia. Gen reconocido: N (https://www.who.int/docs/default-source/coronaviruse/conventional-rt-pcr-followed-by-sequencing-for-detection-of-ncov-rirl-nat-inst-health-t.pdf?sfvrsn=42271c6d_4)
- Protocolo de National Institute of Infectious Diseases, Tokio, Japón. Gen reconocido: N (https://www.who.int/docs/default-source/coronaviruse/method-niid-20200123-2.pdf?sfvrsn=fbf75320_7)
- Protocolo de Institut Pasteur, Paris. Genes reconocidos: E y RdRp. (https://www.who.int/docs/default-source/coronaviruse/real-time-rt-pcr-assays-for-the-detection-of-sars-cov-2-institut-pasteur-paris.pdf?sfvrsn=3662fcb6_2)
- Protocolo de KiCqStart de Sigma-Aldrich. Gen reconocido: S (<https://www.sigmaaldrich.com/covid-19/main-es.html>)

REFERENCIAS

1. Eskier D, Suner A, Oktay Y, Karakülah G. 2020. Mutations of SARS-CoV-2 nsp14 exhibit strong association with increased genome-wide mutation load. PeerJ 8:e10181 <https://doi.org/10.7717/peerj.10181>
2. Korber B y col. 2020. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell. 182(4):812-827.e19. doi: 10.1016/j.cell.2020.06.043. Epub 2020 Jul 3. PMID: 32697968; PMCID: PMC7332439.
3. Plante JA y col (2020). Spike mutation D614G alters SARS-CoV-2 fitness. Nature 592:116–121. doi: 10.1038/s41586-020-2895-3.
4. Rambaut y col. 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 5:1403–1407. <https://doi.org/10.1038/s41564-020-0770-5>.
5. Singer y col. 2020. CoV-GLUE: A Web Application for Tracking SARS-CoV-2 Genomic Variation Preprints 2020, 2020060225 <https://doi.org/10.20944/preprints202006.0225.v1>.
6. Starr y col. 2021. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. Cell 182(5):1295–1310.e20. DOI: 10.1016/j.cell.2020.08.012.
7. Pereira 2021. SARS-CoV-2 variants combining spike mutations and the absence of ORF8 may be more transmissible and require close monitoring. Biochemical and Biophysical Research Communications 550:8-14. doi.org/10.1016/j.bbrc.2021.02.080.