



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

GENÓMICA COMPUTACIONAL

2026-1

Proyecto Final

**Análisis genómico y epidemiológico del virus del sarampión en México:
comparación entre cepas históricas y recientes**

ALUMNOS:

Baltodano Cuevas Erick Joel - 312245924
Martínez Hidalgo Paola Mildred - 319300217
Rueda Tokuhara Daiki Benedek - 4200003533

FECHA DE ENTREGA:
10 de diciembre del 2025

PROFESOR:
Sergio Hernández López

AYUDANTES:
**Rafael López Martínez
Jazmín de Jesús Santillán Manjarrez**



1. Introducción

Contexto biológico

”El virus del sarampión (Measles morbillivirus) pertenece a la familia Paramyxoviridae y al género Morbillivirus. Es un virus envuelto de ARN de cadena sencilla con polaridad negativa (complementaria a los RNA mensajeros) que forma una nucleocápside helicoidal recubierta de membrana de origen celular”. [3]

Es altamente contagioso, que se disemina a través de gotas en el aire y secreciones en forma de aerosol cuando una persona infectada tose o estornuda. Es un virus de transmisión respiratoria que infecta a humanos y primates.

Produce un cuadro exantemático, que va acompañado de una inmunosupresión transitoria que probablemente contribuye a una parte de las complicaciones por infecciones oportunistas. Su genoma codifica varias proteínas esenciales, incluyendo la proteína de superficie H (hemaglutinina), responsable de la unión a receptores celulares, y la proteína F (fusión), que permite la entrada del virus a la célula huésped. [4]

”Desde el año 2000 al 2023 se han prevenido un estimado de 60.3 millones de muertes a causa del sarampión y a pesar de que existe una vacuna segura y costoeficaz, se estima que en 2023 hubo 107 500 fallecimientos por sarampión en todo el mundo, mayoritariamente niños menores de cinco años no vacunados o que no habían recibido la pauta completa”. [5]

En México, se logró la eliminación de transmisión endémica del sarampión en 1997. En 2016 se declaró erradicada en todo México pero se registraron nuevos brotes en 2019-2020 vinculados a casos importados y disminución de la cobertura vacuna, pero brotes recientes sugieren reintroducciones del virus desde otros países. Analizar y comparar genéticamente las cepas históricas con las más recientes permite identificar patrones de transmisión, evaluar rutas de entrada al país y entender los mecanismos detrás de la reemergencia de esta enfermedad. [2]

Planteamiento del problema

Aunque México logró la eliminación del sarampión, en la última década se han registrado casos y brotes asociados a importaciones. Un análisis comparativo detallado entre las secuencias genómicas de las cepas históricas que circulaban en México antes de su eliminación y las secuencias de los virus responsables de los brotes recientes nos ayuda a poder reconstruir rutas epidemiológicas, confirmar la procedencia de las reintroducciones y evaluar si los patrones genéticos actuales implican cambios relevantes en la dinámica del virus.

Objetivos

Objetivo general: Analizar las diferencias genéticas y epidemiológicas entre las cepas del virus del sarampión que circularon en México antes de su erradicación y las que reaparecieron en brotes recientes.

Objetivos específicos:

- Recopilar secuencias genómicas del virus del sarampión asociadas a México antes de su eliminación y a brotes recientes.
- Identificar mutaciones, regiones variables y linajes genéticos presentes en los brotes recientes y compararlos con las cepas previas.
- Realizar alineamientos múltiples y análisis filogenéticos para determinar la relación evolutiva entre cepas históricas y recientes.

Pregunta de investigación

¿Qué diferencias genéticas existen entre las cepas del virus del sarampión que circularon en México antes de su erradicación y las que causaron los brotes recientes, y qué implicaciones tienen estas variaciones en su origen y reintroducción al país?

2. Materiales y Métodos

Fuente de datos

Las secuencias de nucleótidos se obtuvieron de la base de datos del National Center for Biotechnology Information Virus (NCBI Virus)[6].

Para encontrar los datos que necesitamos aplicamos en los filtros de "Identifiers and Classification", en la parte de "Virus/Taxonomy" escribimos "Measles morbillivirus, taxid:11234" y en "Location and Source" en la parte de "Geographic Region" añadimos "Mexico".

Los accession numbers que usaremos son:

Accession numbers	Año
AF411001.1	2001 (Publicación, no recolección)
KJ659872.1, KJ659873.1, KJ659874.1	2011
KP769784.1, KP769785.1	2015
MH042758.1	2018
MK628192.1, MN544303.1, MN544304.1, MN544305.1, MN544307.1, MN544306.1, MN544308.1	2019

Las secuencias nucleotídicas las descargamos en formato FASTA desde GenBank (NCBI), utilizando solo sus accession numbers para datos más limpios.

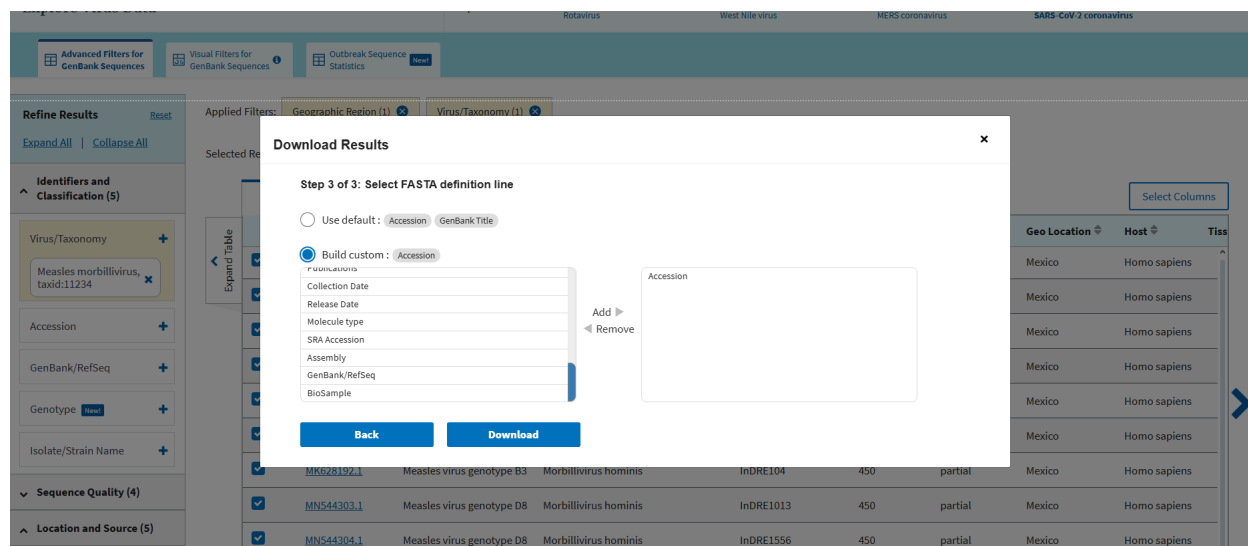


Figura 1: Descarga de datos en formato FASTA de las secuencias

Herramientas de software

El procesamiento y análisis de las secuencias genómicas del virus del sarampión se realizó empleando herramientas bioinformáticas de uso estándar en genómica comparativa.

- **NCBI Virus / GenBank:** Se utilizó para la obtención de las secuencias nucleotídicas del virus del sarampión asociadas a México, descargadas en formato FASTA a partir de sus respectivos accession numbers.

- **MUSCLE**: Se empleó para realizar el alineamiento múltiple de las secuencias nucleotídicas, utilizando los parámetros por defecto disponibles en la plataforma MEGA.
- **Google Colab**: Se empleó un entorno de trabajo en Google Colab, donde se ejecutó código en Python para procesar las secuencias ya alineadas y obtener las métricas utilizadas en el análisis.
- **MEGA12 (Molecular Evolutionary Genetics Analysis)**: Se utilizó para la construcción del árbol filogenético a partir del alineamiento múltiple, así como para el análisis evolutivo y la visualización de las relaciones filogenéticas entre las cepas históricas y recientes

3. Resultados

Métodos de análisis

El análisis genómico se llevó a cabo siguiendo un flujo de trabajo secuencial.

En primer lugar, las secuencias nucleotídicas del virus del sarampión asociadas a México, correspondientes tanto a cepas históricas como recientes, fueron descargadas desde la base de datos NCBI Virus en formato FASTA. Posteriormente, se realizó una revisión básica para asegurar la correcta correspondencia entre los accession numbers, los años de aislamiento y el origen geográfico de cada secuencia.

A continuación, las secuencias fueron alineadas mediante un alineamiento múltiple utilizando el programa MEGA utilizando MUSCLE 5, con parámetros por defecto, con el fin de identificar regiones conservadas y variables entre las cepas analizadas.

Después, se empleó un entorno de trabajo en Google Colab, donde se ejecutó código en Python para procesar las secuencias alineadas y obtener las métricas utilizadas en el análisis.

El alineamiento resultante se importó al software MEGA, donde se construyó un árbol filogenético de Maximum Likelihood utilizando el método de Tamura-Nei, con el objetivo de inferir las relaciones evolutivas entre las cepas históricas y las cepas recientes del virus del sarampión. Para evaluar la robustez de los nodos del árbol filogenético, se realizaron 1,000 réplicas de bootstrap.

Finalmente, el árbol filogenético fue visualizado y exportado desde MEGA para su análisis e inclusión en la sección de resultados.

Alineamiento

Para el alineamiento usaremos las 14 secuencias con los accession numbers ya vistos utilizando N-450, que es una técnica de vigilancia molecular del virus del sarampión que analiza una secuencia específica de 450 nucleótidos para rastrear brotes de otras regiones y países y la transmisión del virus.

El alineamiento múltiple se realizó con MUSCLE 5 sobre las secuencias procesadas.

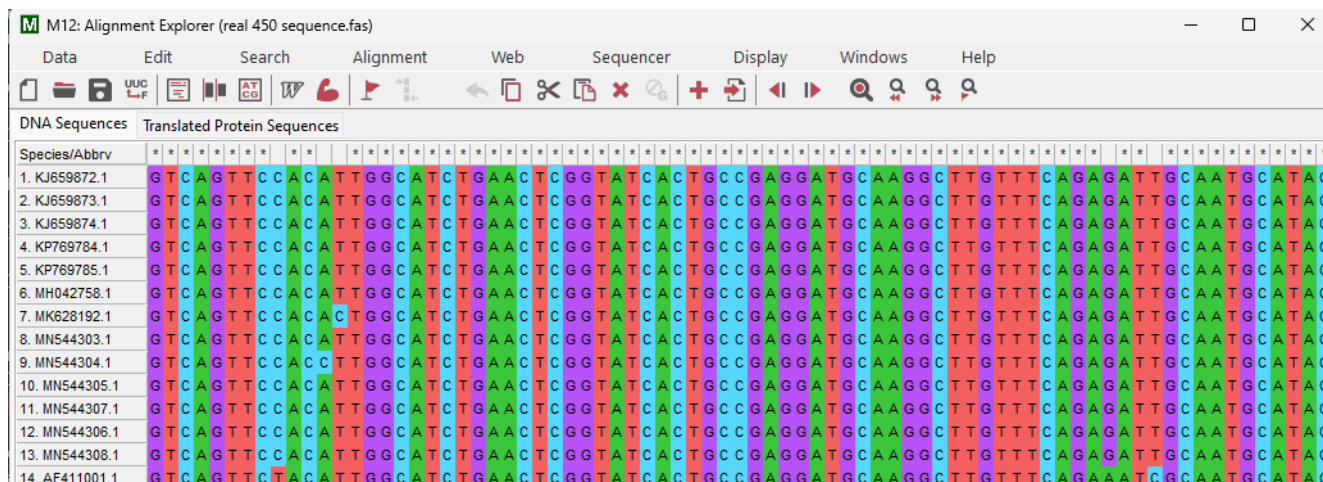


Figura 2: Alineamiento entre las secuencias analizadas.

Árbol filogenético

Se construyó un árbol filogenético de Máxima Verosimilitud (Maximum likelihood) mediante el método Tamura-Nei de con 1000 réplicas de bootstrap. El análisis incluyó un total de 14 secuencias.

M12: Analysis Preferences	
Likelihood Phylogeny Reconstruction	
Option	Setting
ANALYSIS	
Statistical Method →	Maximum Likelihood
PHYLOGENY TEST	
Test of Phylogeny →	None
SUBSTITUTION MODEL	
Substitutions Type →	Nucleotide
Model/Method →	Tamura-Nei model
RATES AND PATTERNS	
Rates among Sites →	Uniform Rates
DATA SUBSET TO USE	
Gaps/Missing Data →	Use all sites
Select Codon Positions →	<input checked="" type="checkbox"/> 1st <input checked="" type="checkbox"/> 2nd <input checked="" type="checkbox"/> 3rd <input checked="" type="checkbox"/> Noncoding Sites
TREE INFERENCE OPTIONS	
ML Heuristic Method →	Nearest-Neighbor-Interchange (NNI)
Initial Tree for ML →	Make initial tree automatically (Default - NJ/MP)
Branch Swap Filter →	None
SYSTEM RESOURCE USAGE	
Number of Threads →	4

Figura 3: Configuración del árbol filogenético

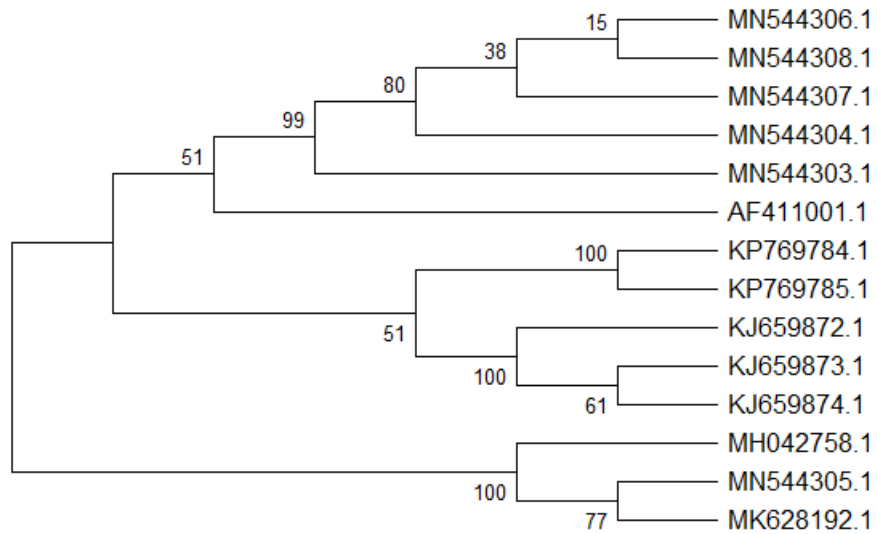


Figura 4: Árbol filogenético de las secuencias de nucleótidos

Resultados finales

Analizamos un total de 14 secuencias con N-450 todas recolectadas en México:

- 1 del año 2001.
- 3 del año 2011.
- 2 del año 2015.
- 1 del año 2018
- 6 del año 2019

=== Composición nucleotídica ===									
	ID	A	T	G	C	%A	%T	%G	%C
0	KJ659872.1	132	85	131	102	29.333333	18.888889	29.111111	22.666667
1	KJ659873.1	132	84	131	103	29.333333	18.666667	29.111111	22.888889
2	KJ659874.1	132	84	131	103	29.333333	18.666667	29.111111	22.888889
3	KP769784.1	140	73	127	110	31.111111	16.222222	28.222222	24.444444
4	KP769785.1	140	73	127	110	31.111111	16.222222	28.222222	24.444444
5	MH042758.1	133	72	130	115	29.555556	16.000000	28.888889	25.555556
6	MK628192.1	132	73	131	114	29.333333	16.222222	29.111111	25.333333
7	MN544303.1	135	77	129	109	30.000000	17.111111	28.666667	24.222222
8	MN544304.1	132	75	131	112	29.333333	16.666667	29.111111	24.888889
9	MN544305.1	135	74	128	113	30.000000	16.444444	28.444444	25.111111
10	MN544307.1	133	76	131	110	29.555556	16.888889	29.111111	24.444444
11	MN544306.1	133	76	131	110	29.555556	16.888889	29.111111	24.444444
12	MN544308.1	133	76	131	110	29.555556	16.888889	29.111111	24.444444
13	AF411001.1	110	67	132	102	24.444444	14.888889	29.333333	22.666667

Figura 5: Composición nucleotídica de las secuencias

=== MATRIZ DE IDENTIDAD (%) ===														
	KJ659872.1	KJ659873.1	KJ659874.1	KP769784.1	KP769785.1	MH042758.1	MK628192.1	MN544303.1	MN544304.1	MN544305.1	MN544307.1	MN544306.1	MN544308.1	AF411001.1
KJ659872.1	0.000000	99.777778	99.777778	92.888889	92.888889	89.555556	89.111111	93.555556	92.888889	89.555556	93.333333	93.333333	93.333333	91.970803
KJ659873.1	99.777778	0.000000	100.000000	93.111111	93.111111	89.333333	88.888889	93.777778	93.111111	89.333333	93.555556	93.555556	93.555556	92.214112
KJ659874.1	99.777778	100.000000	0.000000	93.111111	93.111111	89.333333	88.888889	93.777778	93.111111	89.333333	93.555556	93.555556	93.555556	92.214112
KP769784.1	92.888889	93.111111	93.111111	0.000000	100.000000	91.555556	91.111111	94.000000	94.222222	91.111111	94.222222	94.222222	94.222222	92.944039
KP769785.1	92.888889	93.111111	93.111111	100.000000	0.000000	91.555556	91.111111	94.000000	94.222222	91.111111	94.222222	94.222222	94.222222	92.944039
MH042758.1	89.555556	89.333333	89.333333	91.555556	91.555556	0.000000	99.111111	91.333333	91.555556	99.111111	91.555556	91.555556	91.555556	90.267640
MK628192.1	89.111111	88.888889	88.888889	91.111111	91.111111	99.111111	0.000000	90.888889	91.111111	99.111111	91.111111	91.111111	91.111111	89.781022
MN544303.1	93.555556	93.777778	93.777778	94.000000	94.000000	91.333333	90.888889	0.000000	98.888889	91.333333	99.333333	99.333333	99.333333	93.917275
MN544304.1	92.888889	93.111111	93.111111	94.222222	94.222222	91.555556	91.111111	98.888889	0.000000	91.111111	99.555556	99.555556	99.555556	93.673966
MN544305.1	89.555556	89.333333	89.333333	91.111111	91.111111	99.111111	99.111111	91.333333	91.111111	0.000000	91.111111	91.111111	91.111111	89.781022
MN544307.1	93.333333	93.555556	93.555556	94.222222	94.222222	91.555556	91.111111	99.333333	99.555556	91.111111	0.000000	100.000000	100.000000	94.160584
MN544306.1	93.333333	93.555556	93.555556	94.222222	94.222222	91.555556	91.111111	99.333333	99.555556	91.111111	100.000000	0.000000	100.000000	94.160584
MN544308.1	93.333333	93.555556	93.555556	94.222222	94.222222	91.555556	91.111111	99.333333	99.555556	91.111111	100.000000	100.000000	0.000000	94.160584
AF411001.1	91.970803	92.214112	92.214112	92.944039	92.944039	90.267640	89.781022	93.917275	93.673966	89.781022	94.160584	94.160584	94.160584	0.000000

Figura 6: Matriz de comparación de identidades entre secuencias

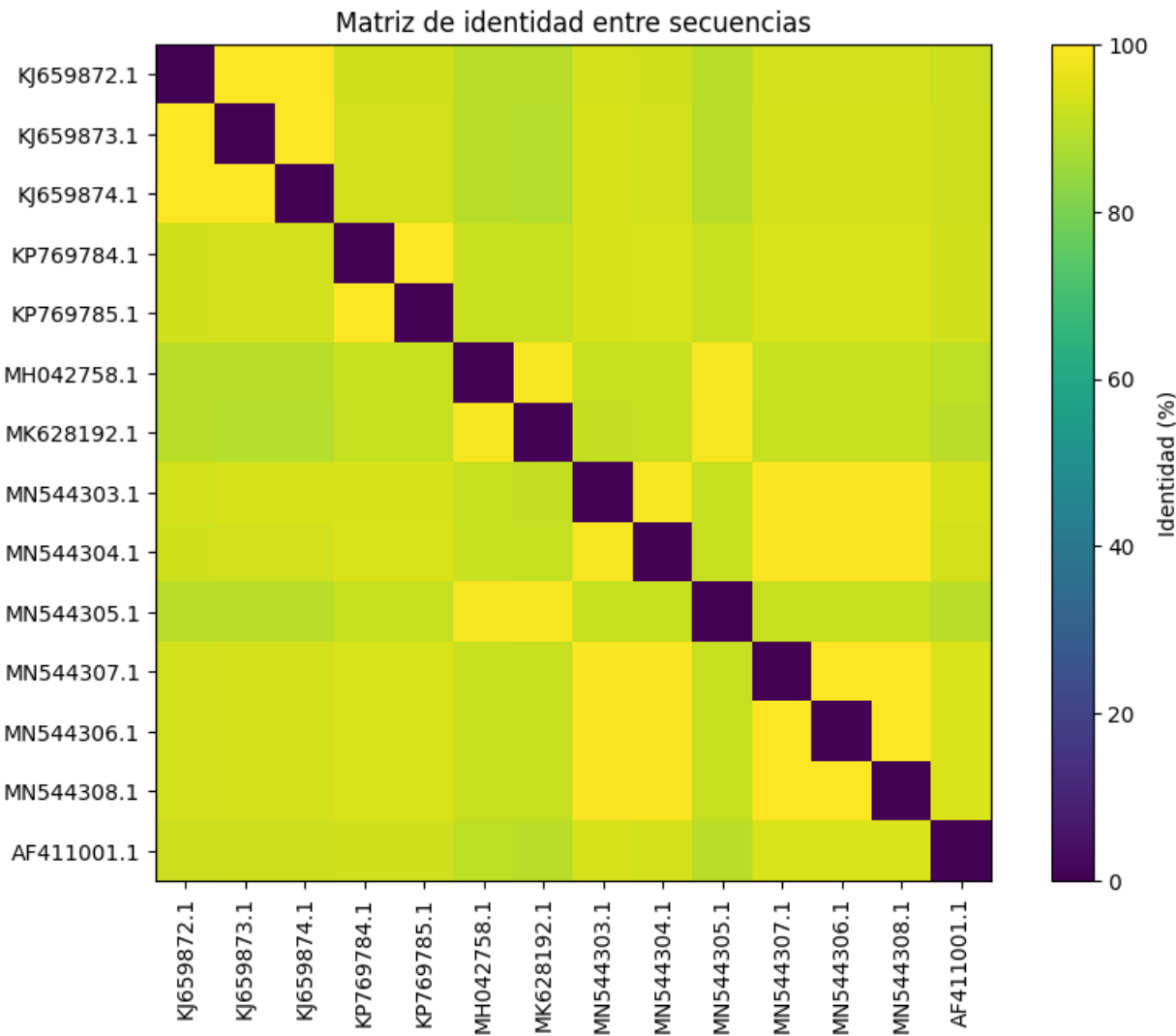


Figura 7: Mapa de calor de comparación de identidades entre secuencias


```
=== DIFERENCIAS SIGNIFICATIVAS ENTRE CEPAS (>= 3 bases) ===
```

Total de diferencias significativas encontradas: 86

	Seq1	Año1	Seq2	Año2	Inicio	Fin	Longitud	Fragmento_Seq1	Fragmento_Seq2
0	KJ659872.1	2011	KP769784.1	2015	415	417	3	TCT	CTC
1	KJ659872.1	2011	KP769785.1	2015	415	417	3	TCT	CTC
2	KJ659872.1	2011	MH042758.1	2018	201	203	3	GGT	AAC
3	KJ659872.1	2011	MH042758.1	2018	238	240	3	GAT	AGC
4	KJ659872.1	2011	MH042758.1	2018	316	318	3	CAG	TCA
81	MN544305.1	2019	MN544306.1	2019	399	401	3	GAC	AGT
82	MN544305.1	2019	MN544308.1	2019	316	318	3	TCA	CAG
83	MN544305.1	2019	MN544308.1	2019	399	401	3	GAC	AGT
84	MN544305.1	2019	AF411001.1	2001	195	197	3	AAA	GGG
85	MN544305.1	2019	AF411001.1	2001	316	318	3	TCA	CAG

Figura 8: Diferencias de bases entre secuencias

```
>>> Total de mutaciones reales entre años:
```

Par_Años	Mutaciones
0 2011 → 2001	3
1 2011 → 2015	6
2 2011 → 2018	9
3 2011 → 2019	18
4 2015 → 2001	2
5 2015 → 2018	4
6 2015 → 2019	8
7 2018 → 2001	2
8 2018 → 2019	10
9 2019 → 2001	4
10 2019 → 2019	20

```
>>> Matriz Año1 vs Año2 (conteo de mutaciones):
```

Año2	2001	2015	2018	2019
Año1				
2011	3	6	9	18
2015	2	0	4	8
2018	2	0	0	10
2019	4	0	0	20

Figura 9: Diferencia del total de mutaciones entre años

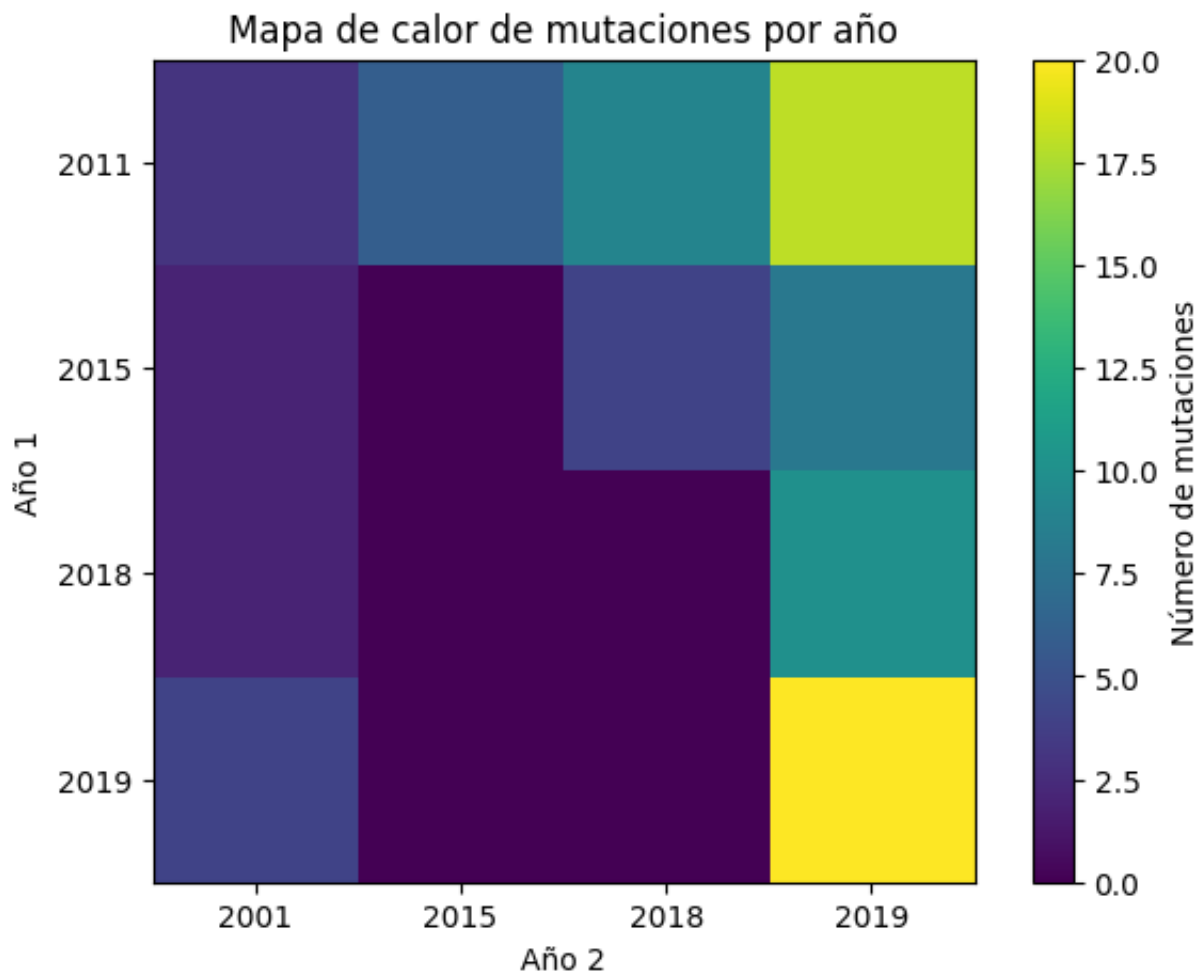


Figura 10: Mapa de calor de diferencia del total de mutaciones entre años

4. Discusión

Interpretaciones y discusiones

El análisis de mutaciones reales entre las secuencias demuestra un patrón temporal claro:

- Las comparaciones entre cepas cercanas en años muestran muy pocas mutaciones (2–6 cambios).
- Las comparaciones entre 2001 y 2019 muestran incrementos mayores (hasta 20 mutaciones).
- Entre cepas del mismo año 2019, se observan diferencias de hasta 20 mutaciones, lo cual indica que coexistieron múltiples linajes circulantes ese mismo año.

El número de mutaciones aumenta conforme crece la distancia temporal entre las muestras. La presencia de hasta 20 mutaciones entre aislamientos obtenidos dentro del mismo año (2019), como se observa en la Figura 9, sugiere la circulación simultánea de variantes genéticamente diferenciadas. Esto podría estar asociado a múltiples eventos de introducción, viajes internacionales o cadenas de transmisión independientes. Considerando además el contexto epidemiológico cercano al inicio de la pandemia de COVID-19 en 2020, es posible interpretar estos patrones como un reflejo del incremento en la movilidad y en el flujo de personas, factores que favorecen la dispersión viral a nivel global.

Además de la acumulación temporal de mutaciones, resulta relevante analizar la localización y el impacto potencial de algunos cambios puntuales detectados:

- En la comparación entre la cepa histórica AF411001.1 (2001) y la cepa reciente MN544305.1 (2019), se observa un cambio de AAA (Lys, lisina) [7] a GGG (Gly, glicina) [7] en la posición 195–197 del fragmento N-450. Esta mutación ocurre en la proteína de la nucleocápside (N) y podría afectar la interacción con el ARN viral o la estabilidad de la nucleocápside. [1]
- Entre las cepas de 2011 (KJ659872.1) y 2018 (MH042758.1), se identificó el cambio GAT (Asp, ácido aspártico) [7] a AGC (Ser, serina) [7] en las posiciones 238–240, también dentro de la proteína N. Este cambio de un aminoácido ácido a uno polar podría alterar regiones expuestas de la proteína, aunque se requiere análisis estructural para confirmar su impacto funcional. [1]
- En el mismo par de años, en las posiciones 316–318 se observó el cambio CAG (Gln, glutamina) [7] a TCA (Ser, serina) [7]. Este cambio de una glutamina a una serina podría modular la interacción de la proteína N con la polimerasa viral (proteína L) o con la fosfoproteína (P). [1]
- Entre las cepas de 2011 (KJ659872.1) y 2015 (KP769784.1), en las posiciones 415–417 se detectó el cambio TCT (Ser, serina) [7] a CTC (Leu, leucina) [7]. Este cambio de un aminoácido polar a uno hidrofóbico podría afectar la estructura terciaria de la proteína N, aunque es necesario validar si se encuentra en una región funcionalmente relevante. [1]

La composición porcentual de A, T, G y C es altamente conservada entre todas las secuencias, con variación mínima:

- A: 24–31 %
- T: 14–19 %
- G: 28–29 %
- C: 22–26 %

La composición nucleotídica de las secuencias mostró una alta estabilidad entre muestras, sin cambios que sugieran presiones evolutivas o mutaciones fuertes. Las variaciones observadas (< 5 %) están dentro del rango aceptado para el gen del virus del sarampión y coinciden con lo reportado previamente en estudios de vigilancia molecular.

Nuestra matriz de identidad nos dió una altísima identidad entre la mayoría de las secuencias (89–100 %) con 3 grupos muy evidentes:

- Grupo 1 (México 2011): Secuencias KJ659872–874. Identidad entre ellas: 99–100
- Grupo 2 (México 2015): KP769784.1 y KP769785.1. Alta similitud entre sí y con el grupo de 2011.
- Grupo 3 (México 2019): MK628192.1 y las secuencias MN544303–MN544308. Presentan mayor divergencia respecto a los grupos anteriores, reflejando la acumulación de mutaciones y posiblemente múltiples introducciones.

La proteína N es vital para que el virus se reproduzca. Los pequeños cambios que vimos en ella probablemente no alteren mucho su función, lo que explica por qué el virus se mantiene tan estable. Lo interesante es que en 2019 vimos varios tipos de virus circulando a la vez, lo que sugiere múltiples entradas al país. Para saber si alguno de estos tipos nuevos era más contagioso o preocupante, habría que estudiar no solo esta pieza (N), sino también las piezas que usa el virus para entrar en nuestras células (H y F).

Limitaciones del trabajo

A pesar de que el análisis filogenético permitió identificar relaciones evolutivas entre las secuencias de sarampión reportadas en México, este trabajo presenta varias limitaciones importantes.

1. El estudio se basó únicamente en la región N-450 del gen N, debido a que no todas las secuencias disponibles en la base de datos NCBI Virus contenían la región completa o genes adicionales.
2. Aunque N-450 es el marcador estándar recomendado por la OMS para la genotipificación del virus del sarampión, su capacidad de discriminación es limitada y no siempre permite distinguir de manera precisa entre variantes circulantes dentro de un mismo genotipo.
3. El tamaño de la muestra fue reducido. Solo se identificaron 14 secuencias provenientes de México en la base de datos, distribuidas en años específicos (2001, 2011, 2015, 2018 y 2019). Esta disponibilidad impide una representación completa del virus en el país.
4. Trabajar únicamente con secuencias mexicanas limita la capacidad de inferir patrones de importación, exportación o rutas de transmisión internacional.

Futuras líneas de investigación

1. Ampliar el muestreo geográfico. Incorporar secuencias de más años y de regiones diversas del país permitiría reconstruir de manera más robusta la historia evolutiva del virus en México y detectar posibles reintroducciones desde otros países.
2. Al ampliar el muestreo y obteniendo muestras de secuencias más completas, podemos incluir regiones genómicas adicionales como el gen H, ofreciendo mayor variabilidad e identificación de las cadenas de transmisión.
3. Si podemos utilizar la secuenciación del genoma completo, podemos obtener mucha más información y nos permitiría identificar mutaciones relevantes para vigilancia epidemiológica, diagnóstico molecular y potenciales fallas vacunales.

5. Conclusión

En este trabajo se realizó un análisis genómico comparativo de 14 secuencias del virus del sarampión asociadas a México, correspondientes a los años 2001, 2011, 2015, 2018 y 2019, todas pertenecientes a la región N-450 del gen N. El análisis mostró que la composición nucleotídica del virus se mantiene altamente conservada a lo largo del tiempo; sin embargo, también se identificaron diferencias genéticas relevantes entre las secuencias. En total, se detectaron 86 mutaciones significativas, lo que evidencia la acumulación progresiva de cambios genéticos en el virus. A partir del análisis de identidad y de la reconstrucción filogenética, fue posible agrupar las cepas principalmente de acuerdo con su año de aislamiento, destacando la presencia de múltiples linajes genéticamente diferenciados, especialmente durante el año 2019.

El árbol filogenético permitió observar una clara relación temporal entre las secuencias, donde las cepas correspondientes a años cercanos, como 2011 y 2015, presentaron una mayor similitud genética. La secuencia histórica de 2001 se ubicó en una posición basal dentro del árbol, mientras que las cepas de 2019 formaron un conjunto más diverso y claramente diferenciado. Este comportamiento indica que los casos observados en años recientes no provienen de un único linaje del virus que haya permanecido circulando, sino que probablemente corresponden a varias introducciones independientes del virus a lo largo del tiempo.

El principal hallazgo de este estudio es la evidencia genómica de la circulación simultánea de múltiples variantes del virus del sarampión en México durante 2019, reflejada en la alta diversidad genética observada entre cepas de ese mismo año. Además, el incremento en el número de mutaciones conforme aumenta la distancia temporal entre los aislamientos respalda este comportamiento evolutivo. En conjunto, estos resultados muestran que el país sigue siendo vulnerable a la introducción del virus del sarampión desde el exterior, lo que subraya la importancia de conservar

niveles elevados de vacunación y fortalecer la vigilancia genómica como herramientas clave para la detección temprana de variantes y la prevención de brotes.

Referencias

- [1] J. Benítez, A. Gonzalez, N. Malats, A. Osorio, M. Robledo, S. Rodríguez, and M. Urioste. Glosario de genética. Fundación Instituto Roche, n.d. Consulta: 8 de diciembre de 2025.
- [2] Secretaría de Salud (Dirección General de Epidemiología). Informe diario del brote de sarampión en México, 2025. Gobierno de México — gob.mx, 2025. Corte al 11 de junio de 2025. Última consulta: (fecha de acceso).
- [3] Rafael Fernández-Muñoz, Juan Carabaña, Montserrat Caballero, Paloma B Liton, Beatriz M Duque, Dolores García-Villalón, and María Luisa Celma. Epidemiología molecular del virus del sarampión. *Revista Española de Salud Pública*, 73(5):605–608, 1999.
- [4] Àngela Domínguez García and Eva Borràs López. El sarampión. *7a Monografía de la Sociedad Española de Epidemiología, Capítol*, 7, 2008.
- [5] A. A. Minta, M. Ferrari, S. Antoni, et al. Progress toward measles elimination — worldwide, 2000–2023. *MMWR Morbidity and Mortality Weekly Report*, 73:1036–1042, 2024.
- [6] National Center for Biotechnology Information. NCBI Virus database. <https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>, 2025. [Internet; access: 2025-12-06].
- [7] Wikipedia contributors. Dna and rna codon tables. Wikipedia, 2025. Última modificación: 16 de noviembre de 2025. Consulta: 8 de diciembre de 2025.