

Aplicación de la filogenética al VIH: qué revela sobre su biología y epidemiología

Índice

2	Agradecimientos
3	Introducción
4	Genomas y genética
6	Secuenciación
8	Cómo se lee un árbol filogenético
10	Diversidad intrahuésped
13	Beneficios y riesgos de la filogenética del VIH
14	Limitaciones
15	Conclusión
16	Glosario

Agradecimientos

ONUSIDA traslada un agradecimiento especial a Lucie Abeler-Dörner, Ana Bulas Cruz, Josh Herbeck, Matthew Hall, Manon Ragonnet-Cronin, Chris Wymant, Katrina Lythgoe, Louise Karlsson y Christophe Fraser por sus aportaciones a esta publicación.

Introducción

La filogenética es el estudio de la *evolución* de los organismos en función de su similitud genética. Las técnicas filogenéticas son procedimientos de laboratorio que permiten leer el código genético presente en todos los organismos vivos y pueden utilizarse para comparar especies diferentes o distintos miembros de una misma especie; por ejemplo, para comparar virus como el VIH. En particular, sirven para identificar los pequeños cambios que se producen en el código genético de cada organismo de una generación a otra. Estos cambios son especialmente pronunciados en el VIH, lo que lo convierte en un excelente candidato para los estudios filogenéticos.

Esta breve introducción a la filogenética del VIH es el primero de tres informes sobre el tema. En él, se introduce el concepto de filogenética, se explica cómo se lee el código genético y se muestra cómo sus secuencias pueden revelar información sobre la biología y epidemiología del VIH. El segundo informe repasa qué se ha descubierto a partir de los estudios filogenéticos del VIH y cómo pueden y deben utilizarse para apoyar los programas de salud pública. El tercero analiza las consideraciones éticas de la filogenética del VIH, incluida la necesidad de proteger y salvaguardar la información sensible.

Genomas y genética

La filogenética utiliza la información genética de un organismo. Todos los animales, plantas, bacterias y virus almacenan su información genética en su *genoma*, que especifica cómo se desarrolla un organismo, cuál es su aspecto y cómo se reproduce. El genoma está compuesto bien por *ácido ribonucleico* (ARN), bien por *ácido desoxirribonucleico* (ADN).

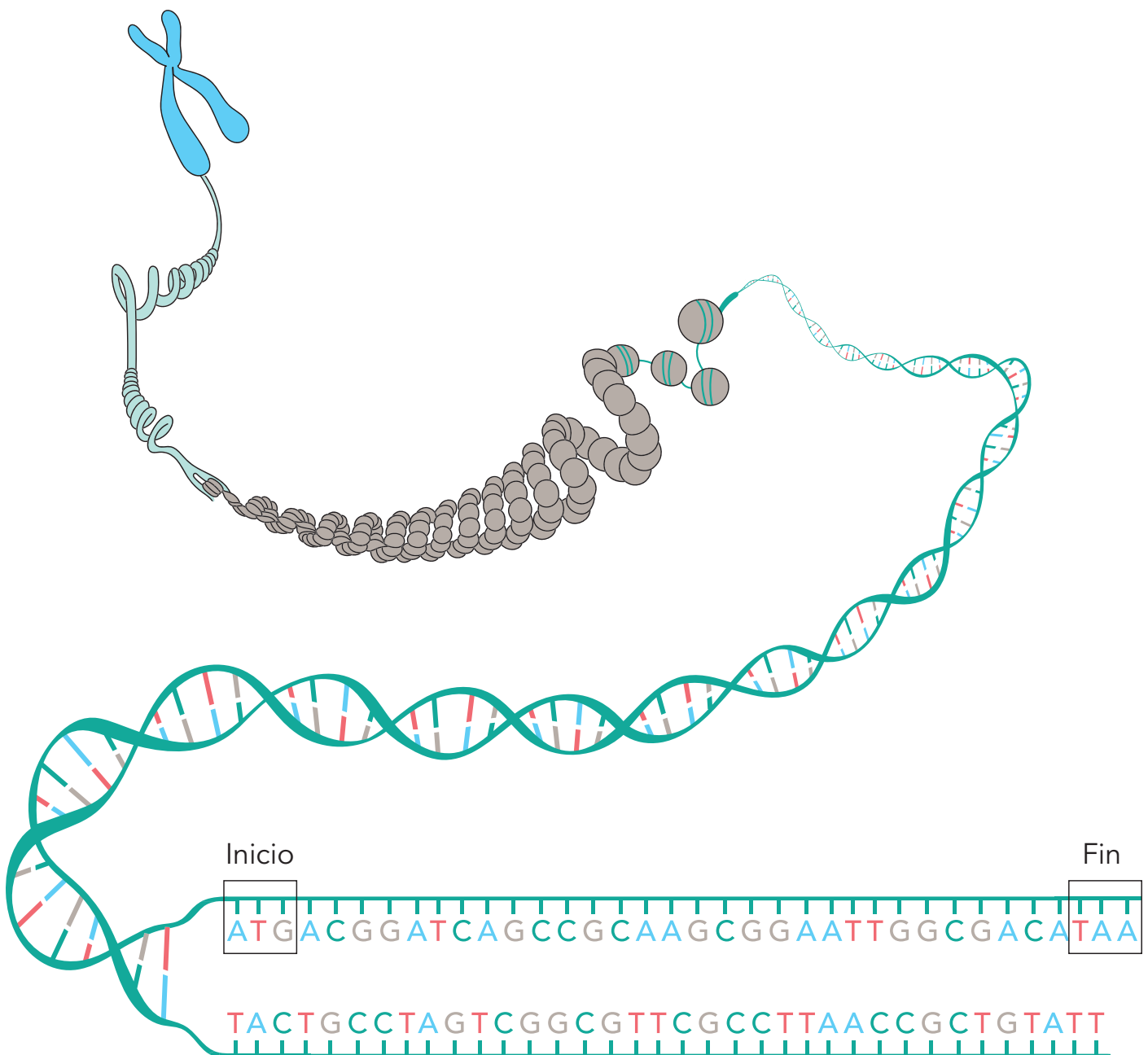
Tanto el ARN como el ADN son moléculas grandes formadas por otras de menor tamaño llamadas *nucleótidos* o *bases*. En el ADN, estos nucleótidos son adenina (A), citosina (C), guanina (G) y timina (T), mientras que, en el ARN, el lugar de la timina lo ocupa una molécula similar llamada uracilo (U). Estos nucleótidos funcionan como «letras» del alfabeto genético: al igual que en un libro, se unen en largas cadenas para codificar información genética. El libro, o conjunto de todo el material genético de un organismo, se denomina genoma. El ARN está formado por una única cadena lineal de A, C, G y U, mientras que el ADN se conforma mediante dos cadenas de A, C, G y T, dispuestas una junto a la otra y unidas por enlaces químicos. Una molécula de ADN parece una escalera enrollada o de caracol (la doble hélice) (Fig. 1). En esta doble hélice, la A de una hebra siempre está emparejada con una T de la otra hebra, mientras que una C siempre lo está con una G. Por lo tanto, la información de ambas hebras es complementaria y, de la misma manera, es posible reconstruir la segunda hebra de una molécula de ADN cuando se conoce la primera.

Actualmente, todos los animales, plantas y bacterias disponen de ADN en sus genomas, pero los virus pueden contar tanto con ADN como con ARN. El VIH contiene un genoma compuesto de ARN y es un ejemplo de retrovirus, lo que significa que este virus puede replicarse en forma de ADN y luego insertarse en el genoma de las células infectadas. Por ello, los medicamentos usados para tratarlo se denominan *antirretrovirales*. Esta capacidad del VIH de esconderse en el genoma humano es una de las razones por las que una infección por el VIH es extremadamente difícil de curar. Su genoma está formado por aproximadamente 10 000 nucleótidos de ARN. En comparación, el genoma humano es mucho mayor, con más de 3200 millones de nucleótidos de ADN.

Nuestra composición genética, junto con el entorno natural y social en el que se crece, influye en diferentes aspectos de nuestro cuerpo:

nuestra apariencia, nuestra altura o peso y qué tan vulnerables somos a distintas enfermedades. En una persona, la mayoría de estos rasgos están determinados por muchas, a veces cientos, de secciones diferentes del genoma. En el caso de los virus, la relación entre el genoma y las características de un determinado virus suele ser más directa. Esto se debe a que el genoma es mucho más pequeño y cada segmento de ADN o ARN suele ser responsable de una propiedad de dicho organismo.

Figura 1: el ADN suele estar dispuesto en forma helicoidal, como una escalera de caracol. Esta «doble hélice» suele estar enrollada alrededor de proteínas circulares y comprimida progresivamente hasta formar una gran estructura en «x» llamada cromosoma.



Secuenciación

La secuenciación genética se refiere al proceso de determinar la secuencia de A, C, G y T (o U) en una muestra. Las técnicas actuales de secuenciación solo pueden realizarse con ADN. Esto significa que el genoma de los virus de ARN, como el VIH, debe convertirse en ADN antes de poder secuenciarlo. También es importante disponer de suficiente ADN para que las máquinas lectoras puedan analizarlo. Por ello, se emplea una técnica conocida como *reacción en cadena de la polimerasa* (PCR), para generar el suficiente ADN para la secuenciación. La PCR duplica las cadenas de ADN hasta obtener una cantidad medible de ADN.

La primera técnica de secuenciación desarrollada, conocida como secuenciación de Sanger, analiza simultáneamente todas las moléculas de ADN amplificadas. Solo se registra el nucleótido más común en cada posición, sin tener en cuenta las variaciones menos frecuentes. Así, si en la posición 1000 del genoma, dos partículas virales presentan el nucleótido A y una el T, se registraría una A en esta posición. El resultado es una secuencia compuesta por el nucleótido más común para cada posición, lo que se conoce como *secuencia de consenso*. Debido a limitaciones técnicas, el procedimiento de Sanger genera secuencias de aproximadamente 500–800 nucleótidos. Para obtener un genoma completo del VIH de 10 000 nucleótidos, es necesario alinear secuenciaciones parcialmente solapadas de los diferentes fragmentos.

Desde 2005, se han desarrollado métodos de secuenciación «profunda» o de «nueva generación», que han reemplazado al de Sanger en muchas aplicaciones, como en los estudios filogenéticos. Dado que estos procedimientos requieren dividir el ADN en fragmentos, ensamblar un genoma completo después de la secuenciación es más complejo que con la secuenciación de Sanger y requiere un mayor poder de cómputo. Sin embargo, esta limitación puede salvarse parcialmente mediante protocolos más nuevos y avanzados. Otra opción es utilizar tecnologías de secuenciación completamente diferentes, como la de Oxford Nanopore. Esta tecnología puede leer fragmentos sin ninguna restricción en cuanto a su longitud. En teoría, puede leer un genoma completo de VIH, o *haplotipo*, en un solo paso. Por ahora, estos sistemas son poco eficientes para analizar grandes cantidades de muestras y son mucho más propensos a errores que la secuenciación profunda. No obstante,

este campo está avanzando rápidamente y es probable que los datos generados con estos métodos estén disponibles próximamente.

La secuenciación de nueva generación tiene una gran ventaja: debido a que los fragmentos de ADN se organizan en un patrón espacial sobre una superficie similar a un microchip durante el proceso de secuenciación, es posible registrar la secuencia de nucleótidos de fragmentos diferentes por separado. Por lo tanto, los métodos de secuenciación profunda nos permiten estudiar tanto la diversidad de virus presentes en diferentes personas como la que existe dentro de un mismo individuo. Esto posibilita el estudio de *variantes minoritarias*, es decir, desviaciones de nucleótidos respecto de la secuencia de consenso.

Tradicionalmente, el VIH se secuencia a partir del suero de una persona, que es la parte acelular de una muestra de sangre. Si las personas que viven con el VIH reciben un tratamiento antirretroviral efectivo, la cantidad de partículas virales en el suero es tan baja que no es suficiente para realizar la secuenciación. Sin embargo, existe una alternativa: puesto que el VIH puede insertar su material genético en el genoma de una persona y estas copias del virus no desaparecen al recibir el tratamiento, cuando este se interrumpe, la infección resurge. Existen métodos para preparar células humanas de cara a la secuenciación de estos *provirus*, pero son más complejos de implementar y se usan menos que aquellos que emplean el suero como punto de partida.

.....

Cómo se lee un árbol filogenético

La filogenética es el estudio de la *evolución* de los organismos en función de su similitud genética. Al igual que los árboles genealógicos, los árboles filogenéticos muestran la relación genética entre diferentes secuencias. La filogenética utiliza modelos matemáticos para deducir la forma en la que determinadas secuencias genéticas del pasado habrían evolucionado para dar lugar a las secuencias genéticas observables en un conjunto de datos en un momento determinado.

Los árboles filogenéticos retroceden hasta un punto de origen: el que se deduce que es el ancestro común más reciente de todas las secuencias del conjunto de datos. Este concepto es análogo al de un progenitor, que es el ancestro común más reciente de un conjunto de hermanos; o al de un abuelo o una abuela, que es el antepasado directo más cercano de todos los primos de una familia. Un ejemplo lo proporcionaría un árbol evolutivo que muestre cuándo se separaron las aves y los mamíferos de los reptiles. Para los virus como el VIH, los árboles filogenéticos pueden reconstruir la forma de propagación de un virus en una población.

Las técnicas filogenéticas parten del hecho de que las secuencias genéticas de todos los seres vivos cambian con el tiempo. Algunos nucleótidos se sustituyen por otros debido a duplicaciones imprecisas. Estos cambios se llaman *mutaciones*. Las mutaciones se acumulan a lo largo del tiempo y se pueden transmitir a la siguiente generación. Esto significa que las secuencias procedentes de organismos con una relación de parentesco cercana se parecen más (es decir, tienen en común una mayor cantidad de nucleótidos) que las de aquellos con una relación más lejana. Los virus acumulan mutaciones más rápidamente que las bacterias, las plantas o los animales, y el VIH suma mutaciones más rápidamente que la mayoría de los virus. Esto hace que el análisis filogenético del VIH sea mucho más esclarecedor que el de otros virus.

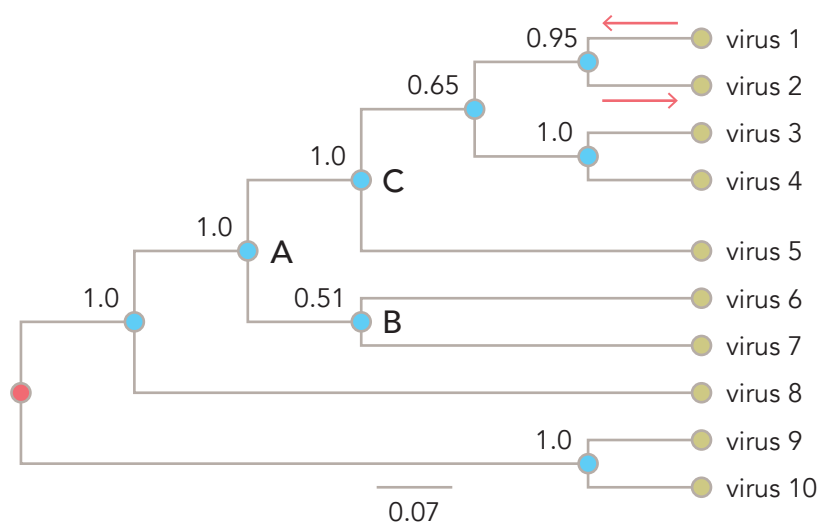


Figura 2: ejemplo de un árbol filogenético enraizado. El círculo rojo es el ancestro originario común al resto de virus del árbol. Esta figura procede del tutorial sobre cómo leer un árbol filogenético de ARTIC Network (<http://artic.network/how-to-read-a-tree.html>). La barra de escala marca la distancia genética en sustituciones por nucleótido.

La Figura 2 muestra un ejemplo esquemático de un árbol filogenético viral. Las líneas horizontales se llaman ramas y representan el cambio genético acumulado a lo largo de sucesivas generaciones de organismos, de antepasados a descendientes. La longitud de la rama es proporcional al grado de cambio genético. Las líneas verticales no transmiten información, sino que se usan para separar las ramas y hacer que el árbol se lea mejor.

Los círculos se llaman *nodos* y representan los antepasados que marcan el final de cada rama. Los círculos verdes en los extremos del árbol reflejan las secuencias presentes en el conjunto de datos. Desplazarse desde esas «puntas» hacia la izquierda equivale a retroceder en el tiempo a través de la historia evolutiva de estas muestras. Cada círculo azul en las líneas horizontales representa un organismo del que no se ha comprobado su existencia en el pasado, sino que se ha deducido. Si nos remontamos lo suficiente en su pasado, los organismos muestreados empiezan a compartir antepasados y características, de la misma forma en la que los primos tienen padres diferentes, pero abuelos comunes (y bisabuelos, etc.). Por ejemplo, B es el antepasado común más reciente de los organismos 6 y 7. Cualquier «corte» en la filogenia resulta en un *clado*, esto es, una secuencia antepasada junto con todos sus descendientes. El nodo interno marcado con un círculo rojo oscuro es la «raíz», que representa el ancestro más reciente que comparten todos los virus del árbol.

En la Figura 2, la distancia genética entre los virus 1 y 2 es la suma de las longitudes de las dos ramas marcadas con flechas rojas. La leyenda en la parte inferior indica la escala para estas longitudes, en este caso 0.07, en unidades de sustituciones por nucleótido. Esto significa que una rama de la misma longitud que la escala representa 7 mutaciones acumuladas por cada 100 nucleótidos en el genoma.

En algunos árboles, el eje x mide unidades de tiempo en lugar de distancia genética. Las mutaciones no siempre ocurren al mismo ritmo, ya que la evolución puede ser más rápida o más lenta. No es fácil convertir un árbol filogenético basado en distancia genética en otro árbol basado en tiempo, pero existen varios modelos matemáticos para llevar a cabo esta tarea.

Debido a que existen diferentes técnicas para crear árboles filogenéticos, puede haber varios resultantes para una misma muestra. En una publicación científica, un árbol filogenético suele mostrar el árbol más probable identificado por un algoritmo. Además, puede incluir información que resuma los muchos árboles posibles que se generaron durante el análisis. A veces, se coloca un número cerca de un nodo interno del árbol para indicar la fracción de árboles generados en la que se agrupan los nodos terminales que descienden de este. En la Figura 2, el 1.0 que se encuentra junto al nodo interno A indica que en todos los árboles construidos aparecía el mismo nodo A que dividía las dos ramas que dieron lugar a los nodos B y C. El 0.51 junto al nodo B señala que este se daba solo en el 51% de todos los árboles construidos. Cuando se comparan secuencias más largas, como el genoma completo del VIH, es recomendable construir múltiples árboles a partir de diferentes partes del genoma para asegurarse de que los árboles resultantes sean coherentes.

Diversidad intrahuésped

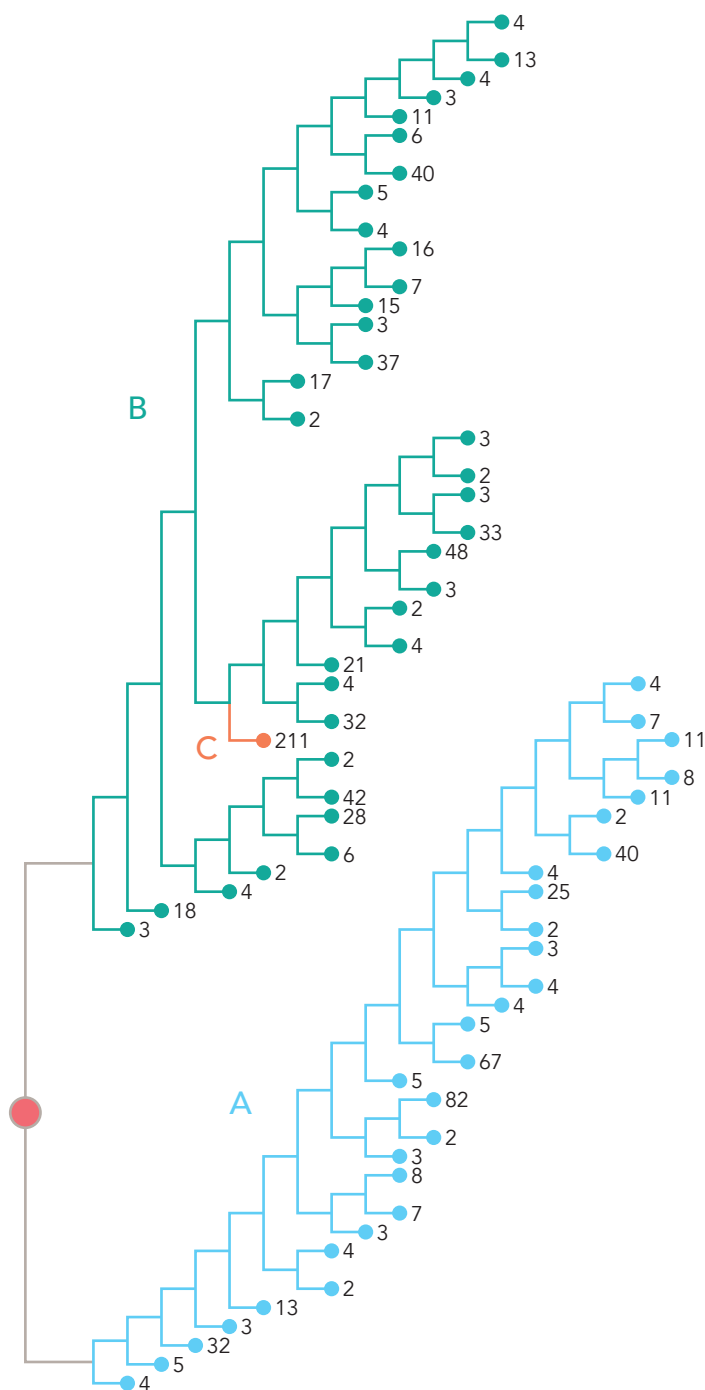


Figura 3: árbol filogenético que muestra la variación dentro de las muestras de una persona (un solo color) y la similitud entre muestras de dos personas (cambio de color). Los números en cada nodo terminal indican el número de copias del mismo genoma en una muestra.

La secuenciación profunda ha permitido a los investigadores ir más allá de la secuencia de consenso y considerar la variación genética presente en la muestra de VIH de una sola persona, no sólo la variación entre diferentes personas. Los virus van acumulando mutaciones a medida que se propagan en la población. Aquellos que causan infecciones de corta duración, es decir, que se resuelven rápidamente, mutan poco en cada individuo afectado. Sin embargo, dado que el VIH provoca una infección crónica, este virus multiplica sus mutaciones en cada persona hasta que se ataja su replicación mediante un tratamiento antirretroviral eficaz. En el caso del VIH, se trata de infecciones provocadas por una única partícula viral, lo que significa que todos los virus de un individuo suelen proceder de una secuencia común. Con cada ciclo de replicación viral, se introducen mutaciones, y en pocas semanas, los virus cambian tanto que la secuenciación revela un conjunto de virus estrechamente relacionados, pero no idénticos. A este conjunto se lo denomina *cuasiespecie*. Muchas de estas mutaciones resultarán en virus que ya no son infecciosos, pero muchas otras también darán lugar a variantes del VIH original que son ligeramente diferentes del virus parental y que, en algunos casos, pueden tener un funcionamiento distinto. Por ejemplo, podrían desarrollar resistencia a un tratamiento antirretroviral concreto o no verse detenidas por algunas defensas inmunitarias del huésped humano. Dado que el VIH provoca una infección crónica, la cuasiespecie tiene tiempo para cambiar progresivamente y diversificarse. Esta diversidad también puede mostrarse en un árbol filogenético (Fig. 3). Las secuencias de los virus encontrados en una persona (indicadas con el mismo color) están, en la mayoría de los casos, más estrechamente relacionadas entre sí que las secuencias de los virus encontrados en personas diferentes (indicadas con colores distintos). Esta variación entre los virus de una misma persona se denomina *diversidad intrahuésped*.

La diversidad intrahuésped aumenta con el tiempo desde que ocurrió la infección. Se han desarrollado métodos para calcular dicho tiempo transcurrido a partir de la diversidad genética del VIH. Estos métodos funcionan razonablemente bien a nivel de población, pero son de precisión muy incierta a nivel individual. Se tratarán con más detalle en el segundo informe de esta serie, *Practical Uses of Phylogenetics in Public Health*.

La diversidad intrahuésped también puede utilizarse para identificar pares de individuos en un conjunto de datos

que probablemente correspondan a casos de transmisión directa. En la Figura 3, se observó que la muestra C presentaba 211 secuencias idénticas. Esta escasa diversidad intrahuésped sugiere que la muestra se le tomó a un individuo poco después de haber contraído la infección por el VIH. Las 211 secuencias están estrechamente vinculadas con las secuencias de la muestra B, de las que son evolutivamente posteriores. Esto sugiere que el participante B ocupa una posición precedente en la cadena de transmisión que llevó al participante F a contraer la infección por VIH, probablemente por transmisión directa. Por otro lado, los virus de la muestra A están mucho menos relacionados con los de las muestras B y C, lo que hace improbable que haya habido transmisión directa. Los análisis de este tipo solo miden la probabilidad de que dos secuencias estén directamente relacionadas, y, por lo tanto, nunca pueden utilizarse fehacientemente para demostrar la transmisión entre dos individuos. Puesto que se basa en un modelo probabilístico, inferir que ha habido transmisión es sumamente delicado y presenta implicaciones éticas que se abordan en el tercer informe de esta serie, *Ethics Considerations for HIV Phylogenetic Analyses*.

Hasta ahora, los análisis de la transmisión del VIH se han realizado principalmente en contextos de investigación. Las agencias de salud pública pueden optar por utilizar análisis de conglomerados (también conocidos como análisis de clústeres) en lugar de análisis filogenéticos de transmisión, ya que son más rápidos, más fáciles de usar y generan datos menos sensibles. Los análisis de conglomerados comparan secuencias en función de su similitud genética y agrupan las secuencias similares en conglomerados, pero no hacen suposiciones sobre la dirección de la transmisión. En varios países, los análisis de conglomerados se han utilizado en contextos de salud pública para identificar subepidemias de rápido crecimiento. En el segundo informe de esta serie, *Practical Uses of Phylogenetics in Public Health*, se tratarán con más detalle los usos de la filogenética y los enfoques basados en conglomerados en la salud pública.

De la muestra al análisis

En un estudio filogenético típico del VIH, se toman muestras de sangre de los participantes, se centrifugan para eliminar las células sanguíneas humanas, se congelan y se envían al laboratorio. Una vez allí, se extrae el ARN de las muestras y se convierte en ADN. En esta etapa, las muestras todavía contienen una gran cantidad de ADN humano, por lo que se enriquecen en ADN derivado del VIH. Durante este proceso, el ADN se amplifica mediante PCR para generar la cantidad suficiente para el proceso de secuenciación. El éxito de este procedimiento depende en gran medida de la *carga viral* del participante, es decir, del número de partículas virales por unidad de volumen sanguíneo. El secuenciador proporciona archivos de datos sin procesar que pueden contener cientos o millones de fragmentos de secuencia, dependiendo de la cantidad de ARN del VIH presente en la muestra original. A continuación, los fragmentos se alinean con una posición específica en el genoma del VIH y se determina el nucleótido más común en cada posición para construir una secuencia de consenso. También se almacenan los archivos de datos que contienen todas las lecturas de secuenciación.

Hasta la fecha, la mayoría de los análisis filogenéticos del VIH han utilizado secuencias de consenso. Estos análisis pueden utilizar el genoma completo o solo una parte. Un ejemplo es el gen de la polimerasa (*pol*), cuya secuenciación se realiza de forma sistemática para estudiar la resistencia a medicamentos en muchos países. Los análisis con secuencias de consenso han servido para descubrir cómo se propagó el VIH dentro y fuera de África durante el siglo XX, así como para estudiar las propiedades y la evolución de los diferentes *subtipos de VIH* y para comprender mejor los riesgos de transmisión.

Es práctica habitual compartir secuencias de consenso en bases de datos públicas, sobre todo una vez publicados los análisis en revistas científicas. Cualquiera puede utilizarlas libremente, pero generalmente contienen información adicional limitada, como el país o el año de obtención. Algunos estudios también tienen en cuenta la información de los fragmentos individuales, como en los análisis de transmisión. Por lo general, estos datos no se ponen a disposición del público, ya que contienen información más sensible, y también debido al gran tamaño de los archivos. La filogenética de patógenos utiliza únicamente secuencias patogénicas. Normalmente, se elimina todo resto de genoma humano que pueda haber en las muestras antes de analizarlas.

Beneficios y riesgos de la filogenética del VIH

Los análisis filogenéticos del VIH pueden utilizarse para entender la dinámica de las epidemias de este virus, incluyendo cómo se ha propagado por todo el mundo en los últimos 50 años. También permiten entender cómo han surgido y evolucionado los distintos subtipos, cómo ha influido la migración en estos patrones y cómo afecta la transmisión del VIH actualmente a diferentes grupos de edad.

Los análisis filogenéticos del VIH pueden generar datos altamente sensibles, lo que supone que su uso no siempre está justificado y que la información obtenida debe manejarse con mucho cuidado. Y lo que es más importante, la filogenética no puede implicar de manera definitiva a un individuo en una cadena de transmisión, pero sí puede excluirlo por completo.

Limitaciones

La realización de estudios filogenéticos acarrea costes significativos, tanto en capital como en gastos corrientes, y procesar datos de secuenciación profunda puede requerir un tiempo computacional considerable. La mayoría de los análisis se pueden llevar a cabo en una computadora estándar, pero el ensamblaje de un genoma y el análisis filogenético de varios miles de muestras pueden demorar semanas en un centro de computación de alto rendimiento, dependiendo de las preguntas de investigación, los análisis realizados y los métodos utilizados.

Conclusión

La filogenética es una herramienta poderosa que ayuda a responder preguntas de investigación virológica, clínica y epidemiológica. Sin embargo, para obtener el máximo beneficio posible, necesita combinarse con otras fuentes de información relevantes. En cualquier caso, conforme las técnicas mejoren y los costos disminuyan, la filogenética probablemente pasará a ser una fuente habitual de datos en los estudios epidemiológicos y clínicos. En este informe, se han expuesto los enfoques principales que se utilizan hoy en día para generar y analizar datos de secuencias del VIH. Por su parte, el segundo informe de esta serie describe ejemplos de cómo se han usado y se siguen utilizando estos enfoques filogenéticos en la salud pública y en la investigación destinada a mejorar la eficiencia y la efectividad de los programas de respuesta al VIH. Finalmente, el tercer informe explora las consideraciones éticas de estos enfoques y examina cuál sería la mejor manera de reducir los riesgos y evitar los daños.

Glosario

Ácido desoxirribonucleico (ADN)

Molécula presente en las células que contiene la información genética encargada del desarrollo y el funcionamiento de un organismo. El ADN está compuesto por cuatro bases: adenina (A), citosina (C), guanina (G) y timina (T).

Ácido ribonucleico (ARN)

Molécula presente en las células que desempeña diversas funciones en la traducción del ADN en proteínas. Algunos virus almacenan su información genética en el ARN en lugar de en el ADN. El ARN está formado por cuatro bases: adenina (A), citosina (C), guanina (G) y uracilo (U).

Antirretrovirales (medicamentos antirretrovirales)

Fármacos que interfieren en el ciclo de la vida del VIH dentro de la célula y evitan que el virus genere nuevas partículas virales. Un tratamiento antirretroviral eficaz para el VIH es un «cóctel» compuesto por tres medicamentos diferentes combinados en una sola pastilla, que generalmente debe tomarse a diario.

Base o par de bases

Componente básico del ADN o el ARN, también llamado nucleótido.

Carga viral

Densidad de partículas virales en un fluido; frecuentemente referida al número de partículas virales por mililitro de sangre.

Clado

Un organismo (o virus) y todos sus descendientes.

Cuasiespecie

Colección de virus ligeramente diferentes (o haplotipos) presentes en la misma persona.

Diversidad intrahuésped

Diversidad genética que existe entre los virus de una misma persona. La diversidad intrahuésped se desarrolla a medida que la partícula viral transmitida inicialmente acumula mutaciones y forma una cuasiespecie.

Evolución

Cambio en las características hereditarias de determinadas poblaciones biológicas a lo largo de varias generaciones.

Genoma

Conjunto completo de toda la información genética de un organismo, codificada por ADN en la mayoría de los seres vivos y por ADN o ARN en los virus.

Haplotipo

Conjunto de genes heredados de un solo progenitor. En el caso de los virus, la secuencia genética contenida en una sola partícula viral. Una cuasiespecie está compuesta por varios haplotipos diferentes.

Mutación

Cambio en la secuencia de A, C, G y T/U en un genoma. El cambio puede ser una sustitución de uno o más nucleótidos por otros, una pérdida de uno o más nucleótidos (delección), una ganancia de uno o más nucleótidos (inserción) o una combinación de los tres.

Nodo

Extremo de una rama en un árbol filogenético, que representa una secuencia genética observada o inferida. Existen dos tipos de nodos: los nodos terminales, que son las secuencias presentes en el conjunto de datos, y los nodos internos, que son los puntos de ramificación en el árbol. La raíz es el nodo interno que representa el ancestro más reciente compartido por todas las secuencias del conjunto de datos.

Nucleótido

Componente básico del ADN y el ARN, también llamado base en un contexto genético.

Provirus

Genoma del virus que se integra en el ADN de una célula huésped.

Reacción en cadena de la polimerasa (PCR)

Técnica rápida y económica para crear copias de pequeños fragmentos de ADN. La mayoría de los análisis genéticos requieren más ADN del que se encuentra en la muestra original y solo se han podido llevar a cabo desde que se inventó la PCR en 1985. El descubrimiento se consideró tan importante que implicó la concesión del premio Nobel de Química en 1993.

Retrovirus

Miembro de una familia de virus que inserta una copia de ADN de su genoma de ARN en el ADN de una célula huésped. Esto hace prácticamente imposible que el sistema inmunitario elimine por completo una infección retroviral.

Secuencia de consenso

Secuencia genética generada a partir de múltiples fragmentos de la misma muestra, en la que se señala el nucleótido más común en cada posición. Es posible que ninguna molécula de la muestra sea idéntica a la secuencia de consenso, pero ofrece un resumen útil de las secuencias que están presentes en la muestra.

Subtipo de VIH

Grupo de variantes del VIH relacionadas filogenéticamente entre sí de manera más estrecha que con otros subtipos. La variación genética dentro de un subtipo puede ser del 15 al 20 %, mientras que la variación entre subtipos suele estar entre el 25 y el 35 %.

Variante minoritaria

Variante genética, como la producida por polimorfismo de un solo nucleótido, que no está representada en la secuencia de consenso.

© Programa Conjunto de las Naciones Unidas sobre el VIH/sida (ONUSIDA), 2024

Algunos derechos reservados. Esta obra está disponible bajo licencia Creative Commons Atribución-NoComercial-CompartirIgual 3.0 Intergovernmental Organization (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo/>).

En virtud de las condiciones de esta licencia, puede copiar, redistribuir y adaptar el trabajo para fines no comerciales, siempre que se mencione adecuadamente el trabajo, como se indica a continuación. Sea cual sea el uso que se dé a este trabajo, bajo ninguna circunstancia puede insinuarse que ONUSIDA respalde a ninguna organización, producto o servicio específico. No se permite el uso del logotipo de ONUSIDA. Si adapta la obra, debe obtener autorización bajo la misma licencia o una licencia equivalente de Creative Commons. Si crea una traducción de este trabajo, debe añadir el siguiente descargo de responsabilidad junto con la mención sugerida: «La traducción no ha sido creada por ONUSIDA. ONUSIDA no se hace responsable del contenido ni de la precisión de esta traducción. La edición original en inglés será la edición vinculante y auténtica».

Cualquier mediación relacionada con disputas derivadas de la licencia se llevará a cabo de acuerdo con las normas de mediación de la Organización Mundial de la Propiedad Intelectual (<https://www.wipo.int/amc/es/mediation/rules/index.html>).

Cita sugerida. Aplicación de la filogenética al VIH: perspectivas sobre su biología y epidemiología. Ginebra: Programa Conjunto de las Naciones Unidas sobre el VIH/sida; 2024. Licencia: CC BY-NC-SA 3.0 IGO.

Material proporcionado por terceros. Si desea reutilizar material de este trabajo atribuido a un tercero, como tablas, cifras o imágenes, es su responsabilidad determinar si se necesita permiso para dicha reutilización y obtener el permiso del titular de los derechos de autor. El riesgo de reclamaciones derivadas de la infracción sobre cualquier componente de terceros en la obra recae exclusivamente en el usuario.

Las denominaciones empleadas y la presentación del material utilizado en esta publicación no implican la expresión de ninguna opinión por parte de ONUSIDA con relación al estado legal de ningún país, territorio, ciudad, área ni a sus autoridades. Las líneas discontinuas en los mapas representan líneas fronteras aproximadas para las que puede no haber pleno acuerdo aún.

La mención de empresas mercantiles específicas o de nombres comerciales de ciertos productos no implica que ONUSIDA los apruebe o recomiende por encima de otros similares. Salvo errores y omisiones, los nombres de los productos propietarios se distinguen por letras mayúsculas iniciales.

ONUSIDA ha adoptado todas las precauciones razonables para verificar la información incluida en esta publicación. No obstante, el material publicado se distribuye sin garantía de ningún tipo, ya sea explícita o implícita. La responsabilidad de la interpretación y el uso del material recae en el lector. En ningún caso ONUSIDA será responsable de los daños derivados de su uso.

ONUSIDA/JC3106a5



ONUSIDA
Programa Conjunto
de las Naciones Unidas
sobre el VIH/SIDA

20 Avenue Appia
1211 Geneva 27
Suiza

+41 22 791 3666

unaids.org/es