

Desafíos en la reconstrucción de haplotipos para el análisis de cuasiespecies virales del Virus de la Fiebre Aftosa

González Mora, RD(1,2); Cacciabue, M(2, 3); Jatón, J(1,2,3); König, G(1, 2); Gismondi, MI(1, 2,3).

(1) Instituto de Agrobiotecnología y Biología Molecular (IABIMO, INTA-CONICET); (2) CONICET; (3) Universidad Nacional de Luján, Departamento de Ciencias Básicas.

Contacto: gonzalezmora.roman@inta.gob.ar

La estructura poblacional de los virus de ARN está vinculada con la alta tasa de error de su polimerasa. Debido a esto los virus son descritos como "nubes de mutantes" o "enjambres de mutantes", denominadas cuasiespecies virales. Estas cuasiespecies son distribuciones complejas de variantes genómicas estrechamente relacionadas que actúan como una unidad de selección. En la era de la secuenciación masiva, la reconstrucción de cuasiespecies a partir de secuencias cortas aún constituye un desafío. El objetivo de este trabajo fue reconstruir cuasiespecies a partir de *datasets* de secuencias obtenidas por tecnología Illumina, con especial énfasis en la identificación de las limitaciones del algoritmo de análisis. Para ello, se analizaron 3 muestras de cultivos celulares y 2 muestras de bazos de ratones infectados con variantes del virus de la fiebre aftosa (VFA). A partir del ARN total de las muestras, se amplificaron dos fragmentos parcialmente solapados, abarcando el 88 % de la longitud del genoma viral. Ambos amplicones se secuenciaron a partir de bibliotecas *paired end* Nextera XT o DNA prep. Las lecturas filtradas (profundidad promedio de entre 1000X y 10000X) se procesaron utilizando el algoritmo MultiQuas, desarrollado por nuestro grupo, que realiza una correlación entre el número de sitios polimórficos identificados con el programa LoFreq y la reconstrucción de haplotipos a partir del programa QuRe. Asimismo, el algoritmo permite utilizar múltiples genomas de referencia. A partir de los analizados fue posible reconstruir los haplotipos que conforman las distintas cuasiespecies y estimar la estructura poblacional en cada muestra utilizando el *software* BAPS (*Bayesian Analysis of Population Structure*). Se observó que la reconstrucción de haplotipos puede variar según la longitud y las características del segmento genómico estudiado. Dado que los haplotipos reconstruidos a partir de lecturas cortas pueden no coincidir exactamente con las secuencias de los genomas que conforman las cuasiespecies, el análisis de los agrupamientos generados por BAPS contribuye a una descripción más robusta de las poblaciones virales. Además, observamos que el algoritmo utilizado para la reconstrucción de haplotipos priorizó diferencias sinónimas de baja frecuencia, lo que plantea un desafío significativo, ya que dichas variaciones no siempre se correlacionan con las características fenotípicas del virus que podrían estar asociadas a distintos de secuencias. Este análisis subraya la importancia de ajustar los métodos bioinformáticos para reconstruir haplotipos, asegurando que las posiciones

relevantes sean adecuadamente priorizadas, y ofrece una base para futuros estudios de evolución viral.