

miRBac: generación y evaluación de un pipeline bioinformático para la predicción de genes productores de microRNA en Baculoviridae

Cuccovia Warlet, FU(1); Motta, LF(1); Lassalle, FM(1); Belaich, MN(1); Cerrudo, CS(1).

(1) Laboratorio de Ingeniería Genética y Biología Celular y Molecular –Área Virosis de Insectos (LIGBCM-AVI), Instituto de Microbiología Básica y Aplicada (IMBA), Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Roque Sáenz Peña 352 (B1876BXD). Bernal, Provincia de Buenos Aires. Argentina.

Contacto: francocuccovia@gmail.com

Baculoviridae es una familia de patógenos virales infectivos de insectos, que portan genomas de doble hebra de DNA y con potencial para múltiples aplicaciones biotecnológicas. Estos virus se dividen en 4 géneros (*Alpha*-, *Beta*-, *Gamma*- y *Deltabaculovirus*), y codifican entre 89 y 183 genes de proteínas. Además de estos, los genomas baculovirales constan de genes productores de RNA no codificantes, como microRNA (miRNA) o, incluso, potenciales RNA de transferencia (tRNA). Los miRNA son moléculas de RNA pequeños, con una marcada estructura de horquilla en el estadio de transcrito primario y roles regulatorios de silenciamiento génico cuando han madurado, ya sea por acelerar la degradación de transcritos, o por impedimento estérico sobre la maquinaria traduccional. Los miRNA virales pueden afectar directamente los procesos celulares o regular sus propias cascadas de expresión para contrarrestar la respuesta celular, siendo crucial su evaluación en contextos de infección. Hasta la fecha, se han reportado experimentalmente solo 13 genes productores de miRNA baculovirales (bac-miR) en los genomas de los nucleopoliedrovirus de *Bombyx mori* (BmNPV), *Autographa californica* (AcMNPV), *Anticarsia gemmatalis* (AgMNPV) y *Spodoptera litura* (SpltNPV). Esto podría sugerir que los bac-miR se encuentran subnotados en esta familia viral. Por ello, diseñamos una herramienta bioinformática con potencial para predecir nuevos bac-miR, donde los previamente reportados se caracterizaron bioinformáticamente para ser utilizados como set de entrenamiento para descubrir características compartidas. Estos bac-miR expresan pre-miRNA de aproximadamente 90 nucleótidos, que presentan patrones estructurales característicos (longitud y estabilidad termodinámica conservadas), y que llevan a la producción de miRNA de alrededor de 20-23 nucleótidos, caracterizadas por un contenido de GC más alto que el del genoma en el que se encuentran (con un valor de $1,51 \pm 0,33$ en la relación % GC miRNA/% GC genoma). Estos resultados iniciales prosiguieron con el desarrollo de un pipeline bioinformático basado en un análisis por ventana deslizante, sobre las cuales se realiza una predicción de estructura (con el software RNAstructure) y se genera un score en función de varias características: similitud estructural con los miRNA reportados; presencia de motivos promotores; motivos terminadores de transcripción; predicción de miRNA maduro; dG y relación conservada de porcentaje de GC en

el RNA maduro y en el pre-miRNA. La aplicación de este pipeline sobre los genomas de BmNPV, AcMNPV, AgMNPV, SpltNPV permitió la evaluación de 213364 ventanas, seleccionando 7431 con un score significativo y otras 4000 con máximo score. Así, se pudo predecir la existencia de aproximadamente 1000 potenciales bac-miR por genoma, lo que permitió, luego de una caracterización del target y la funcionalidad del potencial miRNA producido, postular un conjunto de bac-miR prometedores para ser validados experimentalmente.