

## IDENTIFICACIÓN DE LA SECUENCIA DEL TERMINADOR TRANSCRIPCIONAL DEL GEN GAP DE *Pichia pastoris* (*Komagataella* spp.)

Nancy Pentón-Piña, Ana Lucía Herrera-Estala, Martha Guerrero-Olazarán, José María Viader-Salvadó  
Universidad Autónoma de Nuevo León, Facultad de Ciencias Biológicas, Instituto de Biotecnología  
San Nicolás de los Garza, N. L., México C.P.66455. jose.viadersl@uanl.edu.mx

*Palabras clave:* *Pichia pastoris*, terminador transcripcional, gen GAP.

**Introducción.** Los avances en bioinformática, el empleo de las anotaciones de los genomas y el estudio transcriptómico han dado la posibilidad de identificar regiones promotoras y más reciente las terminadoras transcripcionales que tienen potencial uso en la expresión de genes heterólogos (1). La definición adecuada de estas regiones, en ocasiones se dificulta por existir diferencias entre las anotaciones para diferentes cepas del mismo microorganismo o porque las secuencias conservadas de las regiones 3'UTR no estén estudiadas en el organismo donde se quieren identificar. Este es el caso de la región terminadora del gen GAP de *Pichia pastoris*.

Es por ello que el objetivo del presente trabajo fue identificar la secuencia de la posible región terminadora del gen GAP.

**Metodología.** Se escogió la secuencia nucleotídica entre las dos regiones codificantes (CDS) del gen GAP y del gen NAB6 (gen contiguo localizado río abajo del gen GAP). Esta secuencia conformada de 490 pb, con posible función de terminador transcripcional del gen GAP y promotor de la expresión del gen NAB6, constituyó la región diana para identificar la secuencia del terminador transcripcional del gen GAP. Se mapearon tres transcriptomas de la cepa KM71 de *P. pastoris*, determinados por RNA-seq, contra el ADN genómico de la cepa CBS7435 de *P. pastoris*, utilizando la herramienta HISAT2 de la plataforma Galaxy, con el fin de localizar una zona sin presencia de transcritos en la región diana, y hacerla corresponder con la posible región del terminador del gen GAP. El resultado del mapeo se visualizó con el visualizador genómico integrativo (IGV). Además, se buscaron cinco secuencias conservadas descritas en la literatura (2, 3) para genes de *Saccharomyces cerevisiae* que participan en el proceso de maduración del extremo 3' de un ARNm: el elemento de eficiencia donde se une el factor de corte 1B (CFIB) que potencia la eficiencia del factor de corte 1A (CFIA); el elemento de posicionamiento de CFIA; dos regiones poli-U (en el ADN poli-T) que flanquean el sitio de corte; y el sitio de corte y poliadenilación. Por último, se buscaron motivos de secuencias de regiones UTR de eucariotas mediante la herramienta UTRscan de la plataforma ITBtools.

**Resultados.** En los tres mapeos de transcriptomas de la cepa KM71 contra el genoma de la cepa CBS7435 se observó la presencia de transcritos en toda la región diana seleccionada para identificar la secuencia del terminador transcripcional del gen GAP; probablemente por la presencia de transcritos crípticos que son propios de genes involucrados en las rutas metabólicas del catabolismo de la glucosa, como lo es el gen GAP que codifica para la gliceraldehído-3-fosfato deshidrogenasa (4). La presencia de estos transcritos dificultó la identificación de una región con posible función terminadora a través del análisis de transcriptomas. En el análisis de la presencia de secuencias conservadas se identificaron cinco secuencias que correlacionaron con elementos que participan en el proceso de maduración del extremo 3' de un ARNm. Se identificó la secuencia GTATGT como elemento de eficiencia por tener dos residuos T en las posiciones 2 y 6, que son considerados esenciales en este tipo de elemento (2); la secuencia AAATAG como elemento de posicionamiento, por ser rica en residuos A y encontrarse río abajo del elemento de eficiencia (2, 3); las secuencias TTCATT y TTTATT como las secuencias poli-T que flanqueen al sitio de corte y poliadenilación; y el probable sitio de corte y poliadenilación. UTRscan predijo en la última sección de la secuencia diana, una región "upstream ORF" (uORF) característica de regiones 5'UTR que podría pertenecer a la región 5'UTR del gen NAB6, lo que indicó que el terminador del gen GAP debería finalizar antes de esta región, lo cual daría una longitud similar a la reportada en la literatura para un terminador transcripcional de genes de levaduras (5).

**Conclusiones.** Se identificó una posible región terminadora del gen GAP de *P. pastoris*.

**Agradecimientos.** Los autores agradecen al proyecto CB-2016-01 (286093). N.P.-P. agradece además el apoyo del CONACYT por la beca otorgada.

### Bibliografía.

1. Vogl T *et al.* (2016) *ACS Synth. Biol.* 5:172-186.
2. Van Helden J *et al.* (2000) *Nucleic Acids Res.* 28:1000-1010.
3. Graber JH *et al.* (2003) *Trends Genet.* 19:473-476.
4. Neil H *et al.* (2009) *Nature* 457:1038-1042.
5. Tuller T *et al.* (2009) *BMC Genomics* 10:391.

