



ANÁLISIS TRANSCRIPCIONAL DE *Pichia pastoris* DURANTE LA PRODUCCIÓN DE UNA PROTEÍNA HETERÓLOGA

Karla Fernández-Cano, José María Viader-Salvadó, Miguel Castillo-Galván, Martha Guerrero-Olazarán
Universidad Autónoma de Nuevo León, UANL, Facultad de Ciencias Biológicas, Instituto de Biotecnología
San Nicolás de los Garza, N.L., México C.P. 66455. martha.guerrero@uanl.edu.mx

Palabras clave: Pichia pastoris, RNA-seq, análisis transcripcional

Introducción. *Pichia pastoris* es ampliamente empleada para la producción de proteínas heterólogas, generalmente induciendo la expresión del gen heterólogo con metanol.

Con el fin de comprender los mecanismos de regulación transcripcional que afectan a la fisiología de *P. pastoris* por la sobreproducción de la proteína recombinante y las condiciones de cultivo, en este trabajo se realizó un análisis de expresión diferencial de genes empleando la tecnología de secuenciación masiva de mRNA (RNA-seq) de una cepa Mut^s de *P. pastoris* crecida bajo condiciones que conducen a altos (CM) y bajos (CP) niveles de producción extracelular de una proteína recombinante (1).

Metodología. La cepa KM71FTEII portadora de una secuencia codificante para una fosfohidrolasa y regulada por el promotor AOX1 fue crecida en un bioreactor de 5 L bajo las condiciones CM que condujeron a una $\mu = 0.013 \text{ h}^{-1}$, $Y_{p/x} = 0.052 \text{ g/g}$, biomasa = 122 g/L en peso seco, y las condiciones CP que resultaron en una $\mu = 0.003 \text{ h}^{-1}$, $Y_{p/x} = 0.022 \text{ g/g}$, biomasa = 87 g/L en peso seco (1). El proceso empleado fue en tres etapas: lote y lote alimentado con glicerol y lote alimentado con metanol (1). Se tomaron muestras antes del cambio de glicerol a metanol (C0) y a las 47 horas después de la inducción. Se obtuvieron preparaciones de RNA con valores RIN comprendidos entre 7.8 y 8.5. Las librerías de DNAC se generaron y secuenciaron en LANGE BIO C INVESTAV, empleando la plataforma Miseq de Illumina. Se empleó un valor de puntaje Phred mayor de 20. Se realizó el mapeo en base a un genoma de referencia con la herramienta TopHat2 v0.6 (2) y el análisis de expresión diferencial con Cuffdiff v0.0.7 (2), ambos en la plataforma Galaxy. El análisis de enriquecimiento se realizó mediante ClueGO (3), de la plataforma Cytoscape. La expresión diferencial de 13 genes fue validada por RT-qPCR empleando SYBR Green y se determinó la correlación entre las técnicas de RNA-seq y RT-qPCR.

Resultados. Se detectó la expresión de 4,950 genes, el 93% de los genes totales anotados. CM condujo a la sub- y sobre-expresión de 470 y 526 genes, respectivamente, respecto a CP, mientras que por el cambio de glicerol a metanol se sub- y sobre-expresaron 380 y 427 genes para CM, y 467 y 422 genes para CP, respectivamente. En CM vs CP se sobre-expresaron significativamente genes relacionados con la mitosis, glicosilación de

proteínas y biosíntesis de aminoácidos, indicando una mayor actividad anabólica en CM. El gen heterólogo y los genes de la ruta de desasimilación del metanol se sobre-expresaron sin diferencias entre CM y CP debido al cambio a metanol, sin embargo las enzimas claves (DAS1 y DAS2) de la ruta de asimilación del metanol fueron sobre-expresadas significativamente en CM vs CP, indicando que en CM está favorecida la producción de biomasa y la generación de energía a través de esta vía, explicando los valores más altos para la μ y biomasa obtenidos en CM respecto a CP (4.3 y 1.4, veces respectivamente).

De 112 genes analizados involucrados en la vía de secreción, 21 se sobre-expresaron en CM vs CP, 14% relacionados con el transporte y translocación al retículo endoplásmico (RE), 24% con chaperonas implicadas en el plegamiento, 29% y 33% relacionados con el procesamiento en el RE y en el aparato de Golgi. Además se sobre-expresaron 31 genes, de 341 analizados, relacionados a la traducción, la mayoría que codifican para proteínas ribosomales (61%). De 57 genes analizados relacionados indirectamente con la secreción, 11 y 26 se sobre-expresaron y sub-expresaron respectivamente, en CM vs CP, de los cuales el 100% de los genes de autofagia, 27 y 29% de la respuesta ERAD y UPR y 55% de otros genes relacionados con el estrés celular se sub-expresaron en CM. Finalmente se observó una correlación entre los métodos de RNA-seq y RT-qPCR ($r^2=0.7$).

Conclusiones. El análisis de la expresión diferencial señala que los factores ambientales en CM condujeron a la regulación de la expresión de genes del proceso de secreción, estrés y traducción que condujeron a valores de $Y_{p/x}$, 2.4 veces más altos en CM que en CP. La regulación de la ruta del metanol hacia la asimilación y una mejor adaptación al estrés en CM generó un mayor crecimiento y producción de biomasa en CM que en CP.

Agradecimientos. Los autores agradecen el apoyo del CONACYT (CB-2010-156719) y las becas otorgadas.

Bibliografía.

1. Viader-Salvadó, J.M., et al., (2013). *Biotechnol. Prog.* 29: 1377–1385.
2. Trapnell, C., et al., (2012). *Nat. Protoc.* 7: 562–578.
3. Bindea, G., et al., (2009). *Oxf. Engl.* 25: 1091–1093.