

## EXPRESIÓN MUSCULAR DE RNAs LARGOS NO-CODIFICANTES EN RESPUESTA A LA INGESTIÓN DE ALIMENTOS

Cardoso<sup>1,2</sup>, T.F., Quintanilla<sup>3</sup>, R., Tibau<sup>4</sup>, J., Mármol-Sánchez<sup>1</sup>, E., González-Rodríguez<sup>3</sup>, O., González-Prendes<sup>1</sup>, R., Ballester<sup>3</sup> y Amills<sup>1</sup>, M.

<sup>1</sup>Departament de Genètica Animal, Centre de Recerca en Agrigenòmica (CSIC-IRTA-UAB-UB), Campus de la Universitat Autònoma de Barcelona, Bellaterra 08193. <sup>2</sup>Fundación CAPES, Ministerio de Educación de Brasil, Brasilia - DF 70040-020, Brasil. <sup>3</sup>Genètica y Mejora Animal, IRTA, Torre Marimon, Caldes de Montbui 08140. <sup>4</sup>IRTA-Monells, Finca Camps i Armet s/n, Monells 17121. taina.figueiredo@cragenomica.es

### INTRODUCCIÓN

La secuenciación masiva del transcriptoma ha revelado la expresión de millares de RNA largos no-codificantes (lncRNA) cuya función se desconoce. Los lncRNAs tienen un tamaño mínimo de 200 nucleótidos y en humano representan un 68% del catálogo de genes expresados (Iyer et al., 2015). Estudios en humano y ratón han revelado que los lncRNA tienen un efecto muy importante sobre la regulación de la expresión génica a través de una variedad de mecanismos que incluyen la unión a los promotores de los genes, la remodelación de la cromatina y la modificación de las histonas, el reclutamiento de proteínas capaces de unirse a regiones reguladoras y el enmascaramiento de lugares de *splicing* en el mRNA (Jeon and Lee, 2011; Kino et al., 2010; Schmitz et al., 2010).

La expresión diferencial de lncRNAs se ha investigado en porcino con el fin de adquirir nuevos conocimientos sobre sus potenciales consecuencias fenotípicas. Zhao et al. (2015) identificaron lncRNAs que podrían desempeñar un papel clave en el desarrollo del músculo fetal porcino y demostraron que en general los lncRNAs están menos conservados que los genes que codifican mRNAs. Zhou et al (2014) usaron datos de RNA-seq y anotaciones de Unigene (<https://www.ncbi.nlm.nih.gov/unigene>) para identificar un total de 6621 RNAs largos no-codificantes e intergénicos (lincRNA) en porcino, y dicho catálogo fue ampliado a 12103 lincRNAs que actualmente están almacenados en la base de datos Animal lncRNA database (ALDB, <http://www.ibiomedical.net/aldb/>).

En un experimento previo, analizamos la expresión diferencial de mRNAs en el músculo *gluteus medius* porcino en respuesta a la ingestión de alimento (Cardoso et al., 2016). Dicho trabajo demostró que el aporte de nutrientes al músculo esquelético afecta a la expresión de numerosos factores de transcripción, algunos de ellos relacionados con el mantenimiento de los ritmos circadianos, y genes relacionados con el metabolismo, el estrés oxidativo y la formación de vasos sanguíneos. En el presente trabajo, nos hemos planteado extender estos análisis a la fracción lncRNA del transcriptoma.

### MATERIAL Y MÉTODOS

**Muestras biológicas y diseño experimental:** Para el experimento de análisis de la expresión lncRNA se obtuvieron muestras biológicas de cerdas sometidas a los siguientes tratamientos: T0 (ayuno), T1 (5 h post-ingestión de alimento) y T2 (7 h post-ingestión), conforme a lo descrito en Cardoso et al. (2016). Cada uno de estos puntos temporales estuvo representado por un total de 12 hembras de la raza Duroc, cuya fase de transición y engorde se había llevado a cabo en la granja experimental de porcino de IRTA-Monells, en idénticas condiciones de alojamiento, manejo y alimentación *ad libitum*. Cuando las cerdas alcanzaron un peso de 80 kg (edad aproximada de 120 días) fueron sacrificadas en el matadero experimental de IRTA-Monells.

**Medición de la expresión de lncRNAs mediante RNA-seq:** De cada individuo, se obtuvo una muestra del músculo *gluteus medius* que fue almacenada en RNAlater. El RNA total se extrajo con el kit RiboPure (Ambion, Austin, TX), se comprobó su calidad y cantidad, y fue secuenciado en el Centre Nacional d'Anàlisi Genòmica (CNAG, <http://www.cnag.cat>), mediante la aproximación RNA-seq, con un aparato HiSeq 2000. La calidad de las secuencias se verificó con el software *FASTQC*. Posteriormente, las secuencias fueron mapeadas con el programa *STAR* (Dobin et al., 2013), teniendo en cuenta el genoma de referencia porcino Sscrofa10.2. Para anotar los lncRNAs porcinos, se utilizaron dos bases de datos diferentes, Ensembl (<http://www.ensembl.org>) y ALDB (<http://www.ibiomedical.net/aldb>). Posteriormente, la expresión de los lncRNAs se cuantificó con el programa

*FeatureCounts v1.24.1* (Liao et al., 2014). Los análisis de expresión diferencial se realizaron con el paquete bioinformático *DESeq2* (Anders and Huber, 2010). El logaritmo de la tasa de cambio o "fold-change" (FC) se tuvo en cuenta para valorar si la expresión de un gen aumenta (valor positivo) o disminuye (valor negativo) respecto a un punto temporal de referencia, ya sea T0 (en las comparaciones T0-T1 y T0-T2) o T1 (en la comparación T1-T2). La búsqueda de lncRNA ortólogos en otras especies de mamíferos, así como la identificación de loci codificantes de proteínas situados cerca (< 30 kb) de los genes lncRNAs, se llevó a cabo con la herramienta *BioMart* de Ensembl (<http://www.ensembl.org/biomart/martview>) y la base de datos ALDB.

## RESULTADOS Y DISCUSIÓN

Mediante las bases de datos Ensembl y ALDB, se identificó un total de 95 y 2688 lncRNAs expresados en el músculo *gluteus medius* porcino, respectivamente. Solamente 5 (Ensembl) y 71 (ALDB) lncRNAs mostraron expresión diferencial en respuesta a la ingestión de alimento ( $P$ -valor  $\leq 0,01$  y  $|\log_2FC| \geq 1$ ). Se llevó a cabo un segundo análisis de expresión diferencial más estricto ( $q$ -valor  $\leq 0,05$  y  $|\log_2FC| \geq 1$ ), y ninguno de los lncRNAs identificados como expresados diferencialmente con Ensembl superó la corrección del  $P$ -valor para múltiples contrastes, mientras que 10 lncRNAs presentes en ALDB continuaron siendo significativos. Los resultados más significativos (ALDB y Ensembl) se muestran en la Tabla 1.

Se observó que la identificación de los lncRNA porcinos dependía fuertemente de la base de datos elegida para anotarlos. Uno de los motivos es que en Ensembl se ha anotado un total de 321 lncRNAs porcinos, mientras que ALDB contiene 12103 lncRNAs. No se ha podido establecer el nivel de concordancia de resultados obtenidos con ambas bases de datos, ya que han sido anotadas de forma independiente y no existe correspondencia entre ambas. Para aquellos lncRNAs expresados diferencialmente, se analizó la existencia de secuencias ortólogas en otras especies. En general, el nivel de conservación evolutiva fue muy bajo. Solamente el lncRNA *7SK.32-207*, presente en Ensembl, tiene genes ortólogos en mamíferos placentarios, mientras que ALDBSSCG0000003715 y ALDBSSCG0000005454, disponibles en ALDB, presentan loci ortólogos en humano y en ratón. Sin embargo la baja conservación evolutiva de la secuencia primaria de los lncRNAs no significa que carezcan de una función biológica importante, pues podría existir conservación a un nivel estructural superior (Johnsson et al., 2014).

Puesto que los lncRNAs a menudo regulan la expresión de genes cercanos (Guil and Esteller, 2012), hemos investigado la expresión de genes mRNA situados cerca (< 30 kb) de los lncRNA expresados diferencialmente. De este modo, hemos observado que nueve genes mRNA presentan expresión diferencial ( $P$ -valor  $< 0,05$  y  $|\log_2(FC)| > 0,6$ ), destacando *ND1* (T0-T1:  $P$ -valor = 0,01 y  $\log_2(FC) = -0,8$ ), *PTGES2* (T0-T2:  $P$ -valor = 0,04 y  $\log_2(FC) = 1,31$ ; T1-T2:  $P$ -valor = 0,02 y  $\log_2(FC) = 1,36$ ), *CYCS* (T1-T2:  $P$ -valor = 0,0009 y  $\log_2(FC) = 1,96$ ) y *GABPB2* (T1-T2:  $P$ -valor = 0,04 y  $\log_2(FC) = 1,25$ ). En definitiva, nuestros resultados indican que la ingestión de alimento tiene efectos sobre la expresión muscular de lncRNAs, lo cual sugiere que podrían regular el metabolismo y la homeostasis energética de dicho tejido.

## REFERENCIAS BIBLIOGRÁFICAS

- Anders & Huber. 2010. *Genome Biol.* 11: R106.
- Cardoso, TF. et al. 2016. XVIII Reunión Nacional de Mejora Genética Animal, Valencia, España
- Dobin, A. et al. 2013. *Bioinformatics* 29: 15–21.
- Guil, S. et al. 2012. *Nat. Struct. Mol. Biol.* 19: 1068–75.
- Iyer, M. et al. 2015. *Nat. Genet.* 47: 199–208.
- Jeon & Lee 2011. *Cell* 146: 119–133.
- Johnsson, P. et al. 2014. *Biochim. Biophys Acta* 1840: 1063–1071.
- Kino, T. et al. 2010. *Sci. Signal.* 3: ra8-ra8.
- Liao, Y. et al. 2014. *Bioinformatics* 30: 923–930.
- Schmitz, K.M. et al. 2010. *Genes Dev.* 24: 2264–2269.
- Zhao, W. et al. 2015. *Sci. Rep.* 5: 8957.
- Zhou et al. 2014. *Genome Biol. Evol.* 6: 1387-1392.

**Agradecimientos:** Este trabajo ha sido financiado con los proyectos AGL2013-48742-C2-1-R y AGL2013-48742-C2-2-R concedidos por el Ministerio de Economía y Competitividad (MINECO). Los autores agradecen el apoyo del MINECO por la acreditación del Centre de Recerca en Agrigenòmica como *Centro de Excelencia Severo Ochoa 2016-2019* (SEV-

2015-0533). Nuestro reconocimiento a la empresa Batallé por su apoyo en la generación del material animal.

**Tabla 1.** Principales RNA largos no-codificantes diferencialmente expresados en el músculo gluteus medius porcino antes de comer (T0) vs 5 horas (T1) y 7 horas (T2) después de ingerir alimento.

Antes (T0) vs 5 h después de comer (T1)							
Base de datos	Gene	log <sub>2</sub> (FC)	P-valor	q-valor	Expr. media* T0	Expr. media* T1	Chr
ALDB	ALDBSSCG0000000258	-1,28	7,36E-08	1,76E-04	48,48	18,63	1
Ensembl	RNASEL	-1,44	0,02	0,98	3,62	1,3	9
Antes (T0) vs 7 h después de comer (T2)							
Base de datos	Gene	log <sub>2</sub> (FC)	P-valor	q-valor	Expr. media* T0	Expr. media* T2	Chr
ALDB	ALDBSSCG00000003483	7,7	2,28E-24	1,53E-21	0,27	58,17	18
Ensembl	ENSSSCG000000018063	-1,1	3,62E-03	0,31	85,77	39,55	MT
5 h (T1) vs 7 h después de comer (T2)							
Base de datos	Gene	log <sub>2</sub> (FC)	P-valor	q-valor	Expr. media* T1	Expr. media* T2	Chr
ALDB	ALDBSSCG00000003483	7,57	3,61E-23	8,53E-20	0,31	56,88	18
Ensembl	ENSSSCG000000018063	-1,09	1,64E-03	0,14	82,05	39,55	MT

\* Valores medios de expresión por conteos normalizados.

## MUSCLE EXPRESSION OF LONG NON-CODING RNAs IN RESPONSE TO FOOD INGESTION

**ABSTRACT:** We have used a RNA-seq based approach to compare the lncRNA expression patterns in the *gluteus medius* muscle of fasted pigs vs fed pigs (5 h and 7 h after food intake). By considering the Ensembl and ALDB databases, we have detected differential expression for 5 and 71 lncRNAs ( $P\text{-valor} \leq 0.01$  y  $|\log_2\text{FC}| \geq 1$ ), respectively. In general, differentially expressed (DE) lncRNAs happened to be poorly conserved. We have also observed that in the vicinity of several DE lncRNAs, there are mRNA-encoding genes that are also DE at different timepoints. This result is interesting because there are evidences that lncRNAs might modulate the expression of neighbouring genes. In summary, our results indicate that the ingestion of food drives changes in the expression of a number of lncRNAs in the pig skeletal muscle, a feature that suggests that they might be involved in the physiological response of the muscle towards nutrient supply.

**Keywords:** RNA-Sequencing, lncRNA, food intake.