

Predicción Computacional de Tríplex Híbridos de ADN-ARN: Nuevos Vínculos entre lncRNAs y Genes Relacionados con la Aptitud Reproductiva en Caballos de Pura Raza Española

María Ángeles Vargas-Pérez^{1*}, Nora Laseca², María Yuzhi Arjona-Delgado¹, Chiraz Ziadi¹, Irene Zayas-Arroyo¹, Mercedes Valera², Antonio Molina¹, Sebastián Demyda-Peyrás¹

¹ Departamento de Genética, Universidad de Córdoba, 14071 Córdoba, España

² Departamento de Agronomía, Escuela Técnica Superior de Ingeniería Agronómica. Universidad de Sevilla, Ctra. Utrera, Km 1, 41013 Sevilla, España

* Corresponding author: z42vapem@uco.es

Resumen

La aptitud reproductiva es un factor clave para el éxito de los programas de mejora genética equina. Sin embargo, los mecanismos moleculares que regulan la fertilidad en yeguas no han sido completamente caracterizados, en parte debido a la elevada complejidad de su arquitectura y a las limitaciones en la disponibilidad de registros fenotípicos. Es por ello que la identificación de nuevos marcadores moleculares que expliquen parcialmente la variabilidad fenotípica representa una estrategia clave para mejorar la precisión de los programas de selección.

En este contexto, los ARN largos no codificantes (lncRNAs), transcritos de al menos 200 pb que adoptan estructuras secundarias complejas, han sido asociados con la regulación génica en diversos procesos biológicos y fenotipos en los animales. Uno de los mecanismos implicados en esta regulación es la formación de tríplex híbridos de ADN-ARN, en los que los lncRNAs se unen a regiones específicas del ADN, generalmente promotores ricos en islas CpG. Estas estructuras, a su vez, inducen el reclutamiento de complejos remodeladores de la cromatina, regulando la expresión de sus genes dianas mediante mecanismos epigenéticos.

Nuestro objetivo fue predecir, mediante análisis *in silico*, la formación de tríplex híbridos entre genes relacionados con la aptitud reproductiva en yeguas Pura Raza Española y secuencias asociadas a transcritos de lncRNAs colindantes al sitio de inicio de la transcripción de los genes candidatos, empleando para ello el software LongTarget. Asimismo, analizamos la estructura de los lncRNAs identificados usando RNAplfold y el solapamiento de los potenciales sitios de unión al ADN con islas CpG.

El dataset analizado incluyó 15 genes seleccionados en base a un estudio de asociación de genoma completo (GWAS) realizado por nuestro grupo con una población de ~900 caballos. Identificamos 11 loci de lncRNAs situados en la proximidad de cuatro de estos genes (*HTRA3*, *ERCC1*, *FOXA3* y *PRSS21*), a una distancia media de 21,2 kb, la cual es consistente con un modo de acción en *cis*. Además, en tres de estos loci se identificaron polimorfismos de un solo nucleótido (SNPs) en el GWAS. Posteriormente, llevamos a cabo una predicción de las posibles interacciones tipo tríplex entre estos lncRNAs y los genes candidatos, seleccionando aquellos que cumplieran los siguientes criterios: 1) enriquecimiento en el número de tríplex en comparación con los controles, indicativo de que son interacciones de alta estabilidad; 2) solapamiento con islas CpG presentes en la región promotora de los genes candidatos, relacionadas con la regulación a nivel epigenético; y 3) accesibilidad de la zona de unión al ADN dentro de la estructura secundaria de los lncRNAs, ya que la exposición es un prerequisite para el establecimiento de este tipo de interacciones.

Este enfoque integrativo nos permitió identificar un conjunto de pares gen-lncRNA de alta confianza que podrían estar implicados en la regulación de la aptitud reproductiva a nivel molecular: *PRSS21*-ENSECAT00000140240, *HTRA3*-ENSECAT00000082066 y *ERCC1*-ENSECAT00000129946. Una vez confirmada experimentalmente su relevancia biológica, estos marcadores podrían incorporarse en programas de selección genómica con el objetivo de mejorar la eficiencia reproductiva de los caballos.

Keywords: ARN largos no codificantes; Tríplex híbridos de ADN-ARN; Aptitud reproductiva; Pura Raza Española

Agradecimientos: Este trabajo ha sido parcialmente financiado por las ayudas CNS2024-154334 (Agencia Estatal de Investigaciones), RYC2021-031781-I (Universidad de Córdoba), y JDC2023-051241-I (Agencia Estatal de Investigaciones Científicas y Fondo Social Europeo Plus). También agradecemos a la Asociación Nacional de Criadores de Caballos de Pura Raza Española (ANCCE) por su colaboración.