

## DETECCIÓN DE HUELLAS DE SELECCIÓN EN EL CROMOSOMA X DE LAS POBLACIONES DE VACUNO DE CARNE AUTOCTONO ESPAÑOL

Lopez, A., González-Rodríguez, A., Cañas-Álvarez, J. J., Díaz, C., Molina, A., Altarriba, J., Baro, J. A., Piedrafita, J., Varona, L<sup>1</sup>.

<sup>1</sup>Unidad de Mejora Genética Animal. Instituto Agroalimentario de Aragón (IA2). 50013. Zaragoza.

### INTRODUCCIÓN

Las poblaciones de ganado vacuno de carne autóctono se encuentran relacionadas genéticamente (Cañas-Álvarez et al., 2015) y es muy probable que surjan evolutivamente de un mismo tronco común a partir del cual se han diferenciado por procesos de deriva genética, de selección y adaptación. Estos procesos de selección y adaptación pueden dejar huellas en la estructura del genoma de las poblaciones. En la literatura, se han descrito un gran número de procedimientos para la detección de esta huella (Qanbari y Simianer, 2014), que, de una manera general, se pueden clasificar en función de la estrategia que utilizan para la identificación de estas señales en tres grandes grupos: 1) procedimientos basados en el estudio de la reducción de la variabilidad local, 2) métodos basados en el análisis de la diferenciación genómica entre poblaciones y 3) procedimientos basados en el análisis de la extensión del desequilibrio de ligamiento.

Un estudio previo (Gonzalez-Rodríguez et al., 2016) identificó varias regiones genómicas asociadas a estos procesos de selección en estas poblaciones. Sin embargo, el análisis se ciñó exclusivamente a los cromosomas autosómicos. Este estudio pretende complementar este trabajo con la identificación de huellas de selección en el cromosoma X.

### MATERIAL Y MÉTODOS

En este trabajo se han utilizados los genotipos del *Bovine HDBeardChip* de Illumina de 171 trios (individuo-padre-madre) pertenecientes a 7 poblaciones de bovino de carne autóctono español: AV – Asturiana de los Valles, n=25; ANI – Avelana Negra Ibérica-, n=24; BP – Bruna dels Pirineus-, n=25; Mo – Morucha-, n=25; Pi – Pirenaica-, n=24; Re – Retinta-, n=24; RG – Rubia Gallega-, n=24. Se realizó un control de calidad con el software PLINK (Purcell et al., 2007), seleccionando aquellos marcadores con un *call rate* superior al 95% a nivel de individuo y de SNP. Tras el proceso de filtrado se seleccionaron 36,732 SNP localizados en el cromosoma X. Posteriormente se realizó un proceso de imputación y de reconstrucción de fases haplotípicas mediante el programa informático BEAGLE (Browning y Browning, 2009) Los SNP resultantes se analizaron mediante tres procedimientos de detección de señales de selección basadas en distintas estrategias:

1. **Reducción de la variabilidad:** Se utilizó el procedimiento de identificación de segmentos de homocigosis de al menos 50 SNP utilizando los genotipados de las madres fundadoras de los trios. Para ello, se usó el programa *detectRUNS* (Biscarini et al., 2018) y se cuantificó el porcentaje de individuos que incluyen cada SNP dentro de los segmentos de homocigosis detectados.
2. **Diferenciación genómica entre poblaciones:** En esta estrategia se optó por el cálculo del estadístico  $F_{ST}$ , utilizando las frecuencias alélicas calculadas a partir de las fases haplotípicas de los machos (1 fase) y de las hembras (2 fases). En total se utilizaron 72 (AV, Pi, Re, RG) o 75 (ANI, BP y Mo) fases haplotípicas por población.
3. **Longitud de Haplotipos o Extensión del Desequilibrio de Ligamiento:** Se utilizó el procedimiento *Nsl* (Ferrer-Admetlla et al., 2014) mediante el software *Selscan* (Szpiech y Hernández, 2014) utilizando las mismas fases haplotípicas que en el apartado anterior.

### RESULTADOS Y DISCUSIÓN

El resumen de los resultados del análisis de segmentos de homocigosis se presenta en la Tabla 1. En ella se observa como el número de segmentos fue de entre 20,4 (Pi) y 29,8 (BP) por individuo. Por otra parte, el tamaño de los segmentos fue mayor en AV y menor en BP y

RG. Estos resultados se pueden asociar con la pérdida de variabilidad local si se cuantifica el porcentaje de individuos para los cuales un SNP se localiza dentro de un segmento de homocigosis. Los resultados se muestran en la Figura 1, donde se observa que la mayor parte del cromosoma X presenta porcentajes de inclusión en los segmento de homocigosis inferiores al 40%. Sin embargo, se puede observar que se pueden identificar 2 regiones donde el porcentaje de homocigosis es más alto de manera relevante en la mayor parte de las poblaciones. En concreto, estas regiones se localizan en torno a los Mb 52, 74 y 86. En la primera de las regiones se localiza el gen PCDH19 (*protocadherin 19*). Las protochaderinas juegan un papel importante en las conexiones del sistema nervioso (Sano et al., 1993) y se trata de proteínas con alto grado de conservación evolutiva. De hecho, Mouresan et al. (2017) identificaron una región en el cromosoma 7 altamente conservada que contenía los genes codificantes de otras proteínas de esta familia.

El segundo de los análisis estuvo basado en la identificación de regiones de alta diferenciación genómica entre poblaciones. El valor medio de  $F_{ST}$  fue de 0,018, muy inferior al obtenido por Cañas-Álvarez et al. (2015) con estas mismas poblaciones y con los marcadores SNP localizados en los cromosomas autosómicos. Este resultado es esperable, debido a que las posibilidades de recombinación son menores en los cromosomas sexuales que en los cromosomas autosómicos. De hecho, en la región genómica situada a partir de la 144 Mb, se puede ver que, en promedio, el grado de diferenciación en las poblaciones es mayor ( $F_{ST} = 0,042$ ) y similar al obtenido en las poblaciones autosómicas. La explicación a este fenómeno es que en esa región se localiza la región pseudoautosómica (Van Laere et al., 2008).

Por otra parte, los resultados del barrido genómico a partir del promedio del estadístico  $F_{ST}$  entre 10 marcadores consecutivos se muestran en la Figura 2, donde se puede observar como los valores promedio son más altos en la región pseudoautosómica (> 144 Mb aprox.). Sin embargo, en algunas regiones dentro de la región no pseudoautosómica se puede detectar un incremento considerable del grado de diferenciación entre poblaciones. En concreto, merecen ser destacadas dos regiones genómica situadas en torno al Mbs 57 y 67. Entre los genes localizados en estas regiones se encuentra el RNF128, que se encuentra implicado en procesos inmunitarios (Anandasabapathy et al., 2003), algunos genes de la familia BEX (BEX4, BEX2 y BEX3), ligados al proceso de apoptosis y que juegan un papel en el proceso de desarrollo neuronal (Kazi et al., 2005) o de la familia TCEAL (TCEAL7 y TCEAL9) que se han asociado a patologías de la reproducción en humana (Chien et al., 2005).

Finalmente, los resultados del análisis de extensión del desequilibrio de ligamiento (no mostrados) no fueron capaces de detectar ninguna región asociada a un proceso de selección reciente, corroborando los resultados obtenidos por González-Rodríguez et al. (2016).

Las conclusiones de este trabajo indican que existe heterogeneidad dentro de la variabilidad a nivel local dentro del cromosoma X en las poblaciones bovinas autóctonas españolas y que, además, se han detectado regiones con un mayor grado de diferenciación entre las poblaciones. En el futuro se pretende realizar un análisis más exhaustivo a través de la comparaciones de las poblaciones por pares y un análisis de enriquecimiento que permita identificar la asociación de las regiones genómicas poco variables o altamente diferenciadas con funciones biológicas específicas.

## REFERENCIAS BIBLIOGRÁFICAS

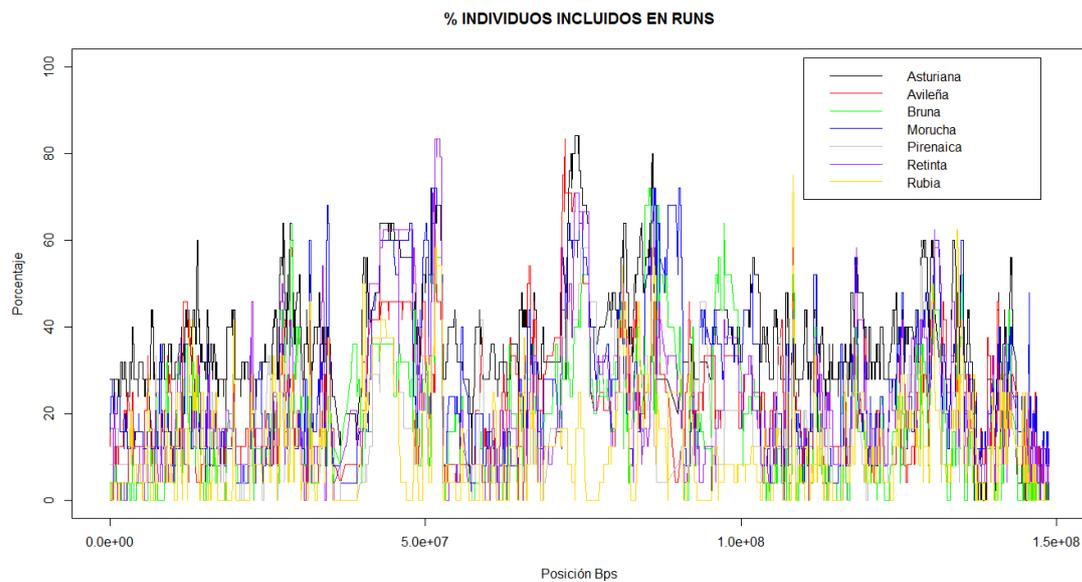
- Anandasabapathy, N et al. 2003. *Immunity* 18:535.
- Biscarini F. et al, 2018. <https://cran.r-project.org/web/packages/detectRUNS/detectRUNS.pdf>
- Browning, B. L. y Browning S: R. 2009. *Am. J. Hum. Genet.* 84:210-223.
- Cañas-Álvarez, J. J. et al., 2015. *J. Anim. Sci.* 93:5164-5174.
- Chien, J. et al. 2005. *Oncogene* 24:5089-5100.
- Ferrer-Admetlla, A et al. 2014. *Mol. Biol. Evol.* 31:1275-1291.
- González-Rodríguez, A. et al. 2016. *Genet. Sel. Evol.* 48:81.
- Kazi, J. U. et al. 2015. *Biochim. Biophys. Acta* 1856:226-223.
- Mouresan, E. F. et al.,

2017. Archivos de Zootecnia 66:59-65. • Purcell, S et al. 2007. Am. J. Human Genet. 81:559-575. • Qanbari, S y Simianer, H. 2014. Livest. Sci. 166:133-143. • Sano, K. et al. 1993. EMBO J. 12:2249-2256. • Szpiech, Z. A. et al. 2014. Mol. Biol. Evol. 31:2824-2827 • Van Laere et al. 2008. Genome Res. 18:1884-1895.

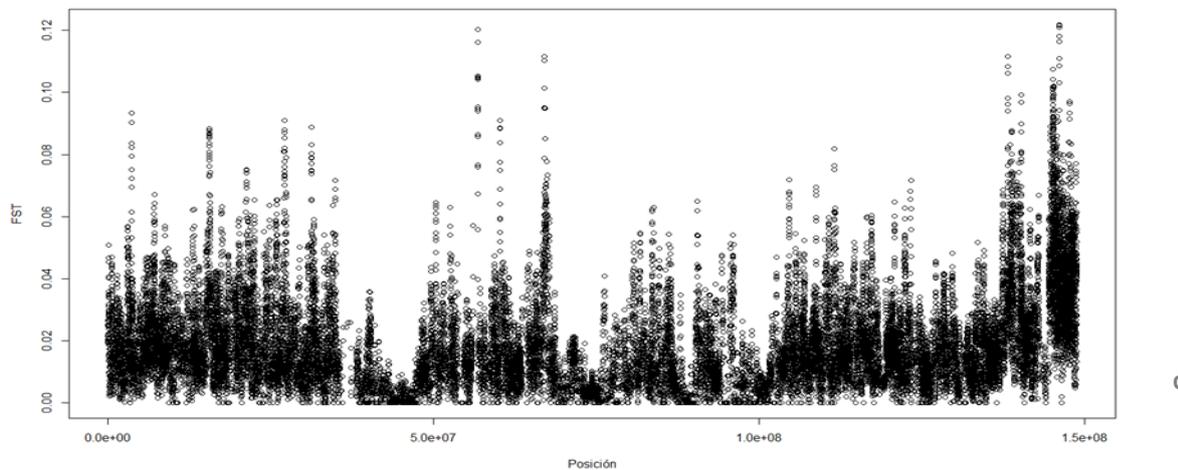
**Agradecimientos:** El material utilizado en este trabajo fue financiado por el proyecto AGL2010-15903. También agradecemos la colaboración de las asociaciones de ganaderos y de Juanjo Arranz por la obtención de las muestras.

**Tabla 1.** Número de segmentos de homocigosis identificados y media (y desviación típica) del tamaño de los mismos en número de SNP y pares de bases

Población	Segmentos	Tamaño (SNP)	Tamaño (bps)
AV	631 (25,2 por ind.)	186,50 (275,30)	2128761 (3110569)
ANI	503 (20,9 por ind.)	136,47 (251,56)	1596667 (2909254)
BP	746 (29,8 por ind.)	80,45 (53,28)	908341 (1001075)
Mo	715 (28,6 por ind.)	130,97 (285,67)	1354434 (2730605)
Pi	490 (20,4 por ind.)	102,21 (173,36)	1196808 (2037519)
Re	657 (27,4 por ind.)	112,25 (165,62)	1192527 (1806342)
RG	612 (25,5 por ind.)	81,13 (67,21)	703532 (819595)



**Figura 1.** Porcentaje de individuos incluidos en segmentos de homocigosis



**Figura 2.** Estadístico  $F_{ST}$  promedio en ventanas deslizantes de 10 SNP

## DETECTION OF SIGNATURES OF SELECTION ON THE X CHROMOSOME OF SPANISH AUTOCHTHONOUS BEEF CATTLE POPULATIONS

**ABSTRACT:** This work is an extension of a previous study to detect signatures of selection in the Spanish Autochthonous Beef Cattle Populations. Here, we analyze them in the X chromosome. We have applied 3 alternative procedures with a data set of 171 trios (individual-sire-dam) from 7 populations (Asturiana de los Valles –AV-, Avileña-Negra Ibérica –ANI-, Bruna dels Pireneus –BP-, Morucha –Mo-, Pirenaica -Pi-, Retinta – Re- and Rubia Gallega –RG-). The first method ( $F_{ST}$ ) is based on the differentiation between populations, the second (ROH) is focused on the local reduction of genetic variation and the third (nSL) on the extension of the linkage disequilibrium. The results of the first method identify relevant signals around the 57 and 67 Mb, whereas the second identifies the genomic regions around the 52, 74 and 86 Mb. In addition, the third method was not able to detect any relevant signals. Finally, the genes located in the genomic regions of relevance were discussed.

**Keywords:** Beef cattle, selection signatures, X chromosome