

REGIONES CON ALTA PERMANENCIA DE DESEQUILIBRIO DE LIGAMIENTO ENTRE LAS POBLACIONES DE VACUNO DE CARNE ESPAÑOLAS

Mouresan, E. F., González-Rodríguez, A., Cañas-Álvarez, J. J., Munilla, S., Moreno, C., Altarriba, J., Baro, J. A., Díaz, C., Molina, A., Piedrafita, J. y Varona, L.
.Universidad de Zaragoza. c/ Miguel Servet 177. 50013. ZARAGOZA. mouresan@unizar.es

INTRODUCCIÓN

En los últimos años, la disponibilidad de genotipado masivo de marcadores SNP ha permitido el desarrollo de procedimientos de evaluación genómica (Meuwissen et al., 2001). La evaluación genómica se ha constituido en una herramienta imprescindible en vacuno de leche (Hayes et al., 2009), sin embargo su utilización en vacuno de carne es más discutible, debido al menor censo de las poblaciones y a la menor implantación de la inseminación artificial. Este problema podría ser solventado a través de la evaluación genómica de poblaciones múltiples, pero los resultados obtenidos tanto mediante simulación (Mouresan et al., 2014) como con datos experimentales (Kachman et al., 2013) son poco esperanzadores.

El éxito de la evaluación genómica a partir de varias poblaciones está ligado a la persistencia de fase de ligamiento entre ellas, de tal manera que el desequilibrio de ligamiento entre marcadores y QTL se mantenga. Varios autores han estudiado la persistencia de la fase entre poblaciones como una medida más de la diversidad genética (Cañas et al., 2014), pero apenas se ha estudiado la arquitectura genética de esta persistencia.

Por lo tanto, el objetivo de este trabajo es estudiar el patrón de la persistencia de fase entre las poblaciones de vacuno de carne españolas a lo largo de los autosomas bovinos.

MATERIAL Y MÉTODOS

Se realizó un genotipado de SNP de alta densidad mediante el BovineHD BeadChip (IlluminaInc, USA) a un total de 171 tríos de padre/madre/hijo pertenecientes a 7 poblaciones españolas de vacuno de carne, que incluían las razas Asturiana de los Valles (AV, n=25), Avileña-Negra Ibérica (ANI, n=24), Bruna dels Pirineus (BP, n=25), Morucha (Mo, n=25), Pirenaica (Pi, n=24), Retinta (Re, n=24) y Rubia Gallega (RG, n=24). Los progenitores fueron seleccionados con la menor relación de parentesco entre si. Los SNP que se utilizaron en el presente estudio pertenecen a los cromosomas autosómicos y no están en posiciones repetidas. Los requisitos adicionales fueron: una probabilidad de equilibrio de Hardy-Weinberg menor de 0,05, una tasa de genotipado ("call rate") entre SNP e individuos superior a 0,95 y una frecuencia del alelo menos frecuente (MAF) inferior a 0,01. El control de calidad se realizó mediante el software PLINK (Purcell et al., 2007), quedando 706.704 SNP, a lo largo de las 2.510.606 kb, con un marcador por cada 3.553 kb de promedio. Los haplotipos parentales se establecieron mediante el software Beagle (Browning y Browning, 2009).

A partir de los haplotipos parentales, se tomaron las poblaciones por pares y se separaron los marcadores que segregaban en ambas poblaciones con una MAF de 0,05. Para cada población se definieron ventanas de 1 Mb a lo largo del genoma y se calculó el desequilibrio de ligamiento entre todos los marcadores de cada ventana como

$$r = \frac{D}{\sqrt{p_i q_i p_j q_j}},$$

donde $D = f_{11}f_{22} - f_{12}f_{21}$ (Falconer y Mackay, 1996). En promedio, el número de marcadores utilizados en cada ventana osciló entre 201 (Pi-Re) y 215 (AV-BP), con desviaciones típicas de en torno a 60.

La persistencia del desequilibrio de ligamiento entre 2 poblaciones se calculó a partir de la correlación entre los valores de DL de cada ventana para cada pareja de marcadores SNP en cada una de las poblaciones.

Una vez obtenidos los valores de persistencia para todo el genoma y para todos los pares de poblaciones, se calculó la media y la desviación típica de estos valores y se seleccionaron las zonas con valores mayores que la media más 3 desviaciones típicas. Posteriormente, se utilizó www.biomart.org y se identificaron los genes localizados en las zonas seleccionadas y las principales rutas biológicas en las que participan.

RESULTADOS Y DISCUSIÓN

Los valores de persistencia entre poblaciones (Tabla 1) y oscilaron entre 0,127 (Mo-Pi) y 0,999 (ANI-BP y BP-RG). Este resultado confirma la relación genética entre las poblaciones (Cañas et al., 2014). Pese a todo, se observó una importante heterogeneidad en la persistencia. Como ejemplo, en la Figura 1 se muestran los resultados obtenidos en el cromosoma 7. Como se puede observar, en este cromosoma hay dos regiones que presentan una alta persistencia (Mb 11 y 54), mientras que en otras zonas apenas se observa. Un aspecto a destacar es que los resultados de persistencia son muy consistentes entre poblaciones. Este hecho nos hace sospechar que ésto debe ser atribuido a la presencia de genes o regiones genómicas de especial relevancia, cuya estructura se mantiene evolutivamente en mayor medida que en otras regiones.

A lo largo del genoma autosómico, se detectaron 42 regiones que presentaron valores por encima del 99,9% de la distribución empírica. Entre ellas, 10 regiones fueron detectadas en, al menos, 10 pares de poblaciones (Tabla 2). En esta tabla también se muestran los genes localizados en cada región.

La búsqueda en las bases de datos genómicos (www.biomart.org) de estas regiones localizó en torno a 220 genes distintos. Un análisis de enriquecimiento mediante el software Enrichr (<http://amp.pharm.mssm.edu/Enrichr/>) identificó las siguientes categorías GO: Procesos biológicos (*cell adhesion*), funciones moleculares (*translation repressor activity*) y componentes celulares (*connexon complex*). Este resultado confirma la presencia de genes fundamentales en las regiones asociadas con una mayor persistencia de fase.

En futuros estudios se pretende estudiar la posibilidad de ponderar las regiones del genoma en función de la persistencia entre poblaciones en los procedimientos de evaluación genómica con varias razas.

REFERENCIAS BIBLIOGRÁFICAS

- Browning, B. L. & Browning, S. R. 2009. *Am. J. Hum. Genet.* 84(2):210-223.
- Cañas-Alvarez, J. J., González-Rodríguez, A., Munilla, S. 2014. 10th WCGALP
- Cañas-Alvarez, J. J., Mouresan, E. F., Varona, L., et al. 2014. 10th WCGALP
- Falconer & Mackay. 1996 4th edn. Longman Group UK, Harlow, England
- Hayes, B. J., Bowman, P. J., Chamberlain, A. J., et al. 2009. *J. Dairy Sci.* 92:433-443.
- Kachman, S. D., Spangler, M. L., Bennett, G.L., et al. 2013. *Genet Sel Evol* 45:30
- Meuwissen, T. H. E., Hayes, B. J., Goddard, M. E. 2001. *Genetics* 157:1819-1829.
- Mouresan, E. F., Munilla, S., Diaz, C., et al. 2014. 10th WCGALP.
- Purcell, S., Neale, B., Todd-Brown, K. et al. 2007. *Am. J. Hum. Genet.* 81:559-575.

Agradecimientos: Esta investigación ha sido posible gracias al soporte económico del proyecto AGL2010-15903. Aldemar González-Rodríguez y Jhon Jacobo Cañas-Álvarez agradecen el apoyo financiero aportado por la beca universitaria BES-2011-045434 y la beca de COLCIENCIAS Francisco José de Caldas 497/2009.

Tabla 1. Valor medio, s.d y valores mínimos y máximos de r en todos los pares de poblaciones.

	Pares de poblaciones	Medio (s.d) r	Min – Max r
1	AV - ANI	0,576 (0,120)	0,239 – 0,989
2	AV - BP	0,632 (0,110)	0,293 – 0,990
3	AV - Mo	0,596 (0,116)	0,266 – 0,987
4	AV - Pi	0,562 (0,122)	0,169 – 0,994
5	AV - Re	0,565 (0,121)	0,228 – 0,979
6	AV - RG	0,592 (0,122)	0,223 – 0,992
7	ANI - BP	0,522 (0,125)	0,162 – 0,999
8	ANI - Mo	0,621 (0,118)	0,227 – 0,998
9	ANI - Pi	0,489 (0,131)	0,140 – 0,992
10	ANI - Re	0,566 (0,127)	0,204 – 0,991
11	ANI - RG	0,525 (0,130)	0,194 – 0,994
12	BP - Mo	0,538 (0,120)	0,209 – 0,989
13	BP - Pi	0,579 (0,127)	0,227 – 0,997

14	BP - Re	0,511 (0,123)	0,192 – 0,994
15	BP - RG	0,572 (0,126)	0,186 – 0,999
16	Mo - Pi	0,507 (0,127)	0,127 – 0,983
17	Mo - Re	0,574 (0,123)	0,215 – 0,986
18	Mo - RG	0,544 (0,126)	0,224 – 0,994
19	Pi - Re	0,480 (0,130)	0,155 – 0,998
20	Pi - RG	0,520 (0,133)	0,189 – 0,996
21	Re - RG	0,515 (0,131)	0,151 – 0,994

Tabla 2. Cromosoma y posición de las zonas con valores de r mayores que el 99,9% que aparecen más frecuentemente, y los genes encontrados en estas zonas.

Cromosoma	Mb	Genes
7	11, 54	Olfactory receptor, <i>CYSTM1</i> , <i>PCDHA5</i> , (y 27 genes más)
10	105	<i>EIF3J</i> , <i>SPG11</i> , <i>U6</i> , <i>U5</i> , <i>PATL2</i> , <i>B2M</i> , <i>TRIM69</i>
12	42, 92	<i>RPL22</i> , <i>CDC16</i> , bta-mir-2303, <i>UPF3A</i> , <i>CHAMP1</i>
15	1, 86	<i>OR9G1</i> , <i>GLB1L2</i> , <i>B3GAT1</i>
16	8	-
17	76	<i>DGCR8</i> , <i>TRMT2A</i> , <i>RANBP1</i> , <i>SNORA77</i> , <i>ZDHHC8</i>
23	1	<i>KHDRBS2</i>

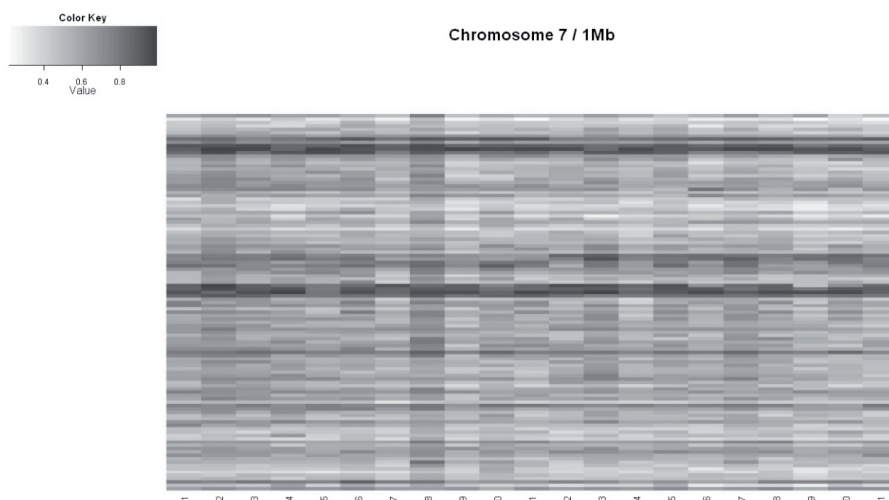


Figura 1. Heatmap de los valores de r por Mb del cromosoma 7 de todos los pares de poblaciones.

REGIONS WITH HIGH PERSISTENCY OF LINKAGE DISEQUILIBRIUM IN SEVEN SPANISH BEEF CATTLE POPULATIONS

ABSTRACT: This study analyzed the persistency of the linkage disequilibrium between 7 Spanish beef cattle populations. The persistency was calculated as the correlation between the LD values of the SNPs of regions of 1 Mb each, between two populations. The results yielded positive values in all cases of comparison and regions. The persistency values of 42 regions were found to be higher than the 99.9% of the empirical distribution and 10 of those regions were found to be consistent in more than 10 of the 21 comparison cases. These 42 regions were found to accommodate around 220 genes associated with biological processes (*cell adhesion*), molecular functions (*translation repressor activity*) and cell components (*connexon complex*).

Keywords: Linkage disequilibrium, beef cattle, persistency