

ANÁLISIS DE LA UTILIDAD PARA ESTUDIOS GENÓMICOS DEL “*Illumina Ovine SNP50BeadChip*” EN RAZAS OVINAS ESPAÑOLAS

Gutiérrez-Gil, B.¹, García-Gámez, E.¹, Suárez-Vega, A.¹, Sánchez, J.P.², Kijas, J.³, Calvo, J.H.⁴, *International Sheep Genomics Consortium (ISGC)*, y Arranz J.J.¹

¹Departamento de Producción Animal, Facultad de Veterinaria, Universidad de León, 24071 León. ²Departamento de Ciencia Animal, UPV, Valencia. ³CSIRO Livestock Industries (Australia). ⁴Unidad de Tecnología en Producción Animal, CITA, Zaragoza. E-mail: beatriz.gutierrez@unileon.es

INTRODUCCIÓN

El proyecto “Sheep HapMap” (www.sheepmap.org), desarrollado por el *International Sheep Genomics Consortium (ISGC)* tiene varios objetivos. Entre ellos podemos destacar los siguientes: el desarrollo y validación del *Ovine SNP50BeadChip*, el estudio de la diversidad genética entre razas, el análisis de la estructura poblacional, la medición del desequilibrio de ligamiento, la estimación del tamaño efectivo de las poblaciones ovinas, la detección de huellas de selección en las razas de diferente aptitud y el conocimiento del proceso de domesticación en esta especie animal. El desarrollo del *Ovine SNP50BeadChip* y su genotipado en muestras de 71 razas ovinas, de ellas cuatro españolas, y constituye un primer paso para llegar a cumplir los objetivos descritos y es, además, un instrumento fundamental para el desarrollo de las herramientas genómicas en el ganado ovino.

Para el desarrollo del Chip de SNPs se secuenció parcialmente el genoma de seis animales (uno por raza) pertenecientes a las razas Awassi, Merino Australiano, Poll Dorset, Romney, Scottish Blackface y Texel, utilizando técnicas de secuenciación de segunda generación que incluyen el 454-FLX y el *Illumina GA*. Las secuencias obtenidas fueron alineadas frente al genoma ovino generándose así el genoma virtual de la oveja (Dalrymple et al., 2007). De estas secuencias, aquellos lugares que presentaban un polimorfismo claro en varios de los animales y que presentaban una distribución uniforme a lo largo del genoma fueron seleccionados para ser incluidas en el Chip de 54.241 SNPs comercializado por *Illumina (Ovine SNP50BeadChip)*. En la elaboración de estas herramientas se ha detectado un sesgo que tiene que ver con las razas de animales en las que se utilizan. Así, tanto en el ganado porcino (Ramos et al., 2009) como en el bovino (McClure et al., 2011) se ha descrito que la utilidad de los marcadores puede estar claramente condicionada por la población animal que estamos analizando y su proximidad filogenética con las razas que fueron utilizadas en la secuenciación para la detección de SNPs y posterior elaboración del Chip. Con objeto de conocer el efecto del sesgo producido en el ganado ovino y de medir la utilidad de esta herramienta genómica de alto rendimiento en nuestras poblaciones ovinas se han analizado un total de 4 razas pertenecientes a diferentes troncos filogenéticos utilizando el *Ovine SNP50BeadChip de Illumina*.

MATERIAL Y MÉTODOS

Los animales utilizados en este estudio, a través de la participación de nuestro grupo de investigación en el proyecto *Sheep HapMap*, fueron: la raza Churra (CHU) (tronco Churro), razas Castellana (CAS) y Rasa Aragonesa (RAA) (tronco Entrefino), Ojalada (OJA) (tronco ibérico). Un total de 167 animales pertenecientes a las 4 razas, fueron seleccionados para el presente trabajo. La distribución por raza fue la siguiente: CHU (n = 100), CAS (n = 23), RAA (n = 20) y OJA (n = 24).

Se han utilizado los criterios de calidad estándar en este tipo de trabajos: se han eliminado los animales con una tasa de genotipos inferior al 97%, y con tasas de heterocigosis observada por encima del 40% y por debajo del 20%; se han eliminado SNPs con una MAF (frecuencia del alelo raro) igual o inferior a 0,01, si no se encontraban en equilibrio Hardy Weinberg ($P < 4,5 \cdot 10^{-6}$) o si presentaban inconsistencias en el modo de herencia en tríos de paternidad y maternidad comprobada.

Los criterios de calidad fueron analizados con las opciones disponibles en el software PLINK (1.06) (<http://pngu.mgh.harvard.edu/~purcell/plink/>). Como medidas de la información aportada por el chip en las diferentes poblaciones se estimaron los siguientes parámetros poblacionales: proporción de alelos con MAF < 0,01 y > de 0,3, indicativos de marcadores poco informativos y muy informativos, respectivamente en cada población; heterocigosis media poblacional o diversidad génica y la heterocigosis individual (proporción de SNPs

heterocigotos en cada individuo); proporción media de alelos compartidos entre los individuos de cada población como D_{st} y definida como $(IBS2+0,5*IBS1)/N$, donde $IBS1$ y $IBS2$ es el número de loci con 1 o 2 alelos idénticos y N el número de individuos de la población.

RESULTADOS Y DISCUSIÓN

Los resultados de los parámetros poblacionales obtenidos al analizar los 49.034 marcadores que han superado las pruebas de calidad en un total de 169 animales, se muestran en la Tabla 1. En la Tabla 2 se presentan los parámetros de informatividad de los marcadores en las razas ovinas secuenciadas para la producción del *Ovine SNP50BeadChip de Illumina*. El dato que presenta una mayor diferencia entre las dos tablas es el MAF. Se puede apreciar que para los SNPs más y menos polimórficos, en las razas utilizadas para la elaboración del Chip, el número de SNPs con $MAF < 0.01$ es prácticamente inexistente, mientras que en nuestras razas alcanza un 4,4% de los SNPs. En cuanto a los más polimórficos, el hecho de que en las razas del Chip haya menos SNPs que cumplan la condición en todas las razas, puede ser debido a las diferencias existentes entre ellas (distintos orígenes, aptitudes productivas, etc.), mucho mayores que las existentes entre las razas ovinas españolas.

En la Figura 1 se muestra la distribución de MAF en el Chip de SNPs para la raza ovina Churra. Estos datos están en concordancia con los obtenidos en el conjunto de las poblaciones (datos no mostrados), siendo el rango 0,4-0,45 el que muestra una mayor proporción de SNPs.

Este estudio constituye la confirmación de que el chip existente es muy informativo en las razas españolas analizadas y que el único sesgo que hemos detectado ha sido en la proporción de marcadores con $MAF < 0,1$. Por todo ello, podemos concluir que este chip de SNPs constituye una herramienta muy útil en estudios de asociación y búsqueda de genes con importancia económica en las razas autóctonas.

REFERENCIAS BIBLIOGRÁFICAS

- Dalrymple, B. P., Kirkness, E. F., Nefedov, M., McWilliam, S., Ratnakumar, A., Barris, W., et al.; International Sheep Genomics Consortium. 2007. *Genome Biol.* 8: R152.
- McClure, M. C., Matakumali, L., Sonstegard, S., Curtis Van Tassell, P. & the Bovine HapMap Consortium. 2011. *Plant & Animal Genome XIX* P523.
- Ramos, A.M., Crooijmans, R. P. M. A., Affara, N. A., Amaral, A. J., Archibald, A. L. 2009. *PLoS ONE* 4: e6524.

Agradecimientos: Este trabajo forma parte del proyecto AGL2009-07000 financiado por el Ministerio de Ciencia e Innovación. BGG es investigadora contratada a través del programa Juan de la Cierva (Ministerio de Ciencia e Innovación). EGG y ASV son becarias FPU (Ministerio de Educación).

Tabla 1. Parámetros de informatividad en las razas españolas.

	CAS		CHU		OJA		RAA		Razas Españolas	
Nº Animales	23		100		24		22		169	
MAF media	0.283		0.276		0.284		0.289		0.283	
	%	SNP	%	SNP	%	SNP	%	SNP	%	SNP
MAF < 0.01	2.4	1158	1.7	834	1.9	909	1.5	732	4.4	2168
MAF > 0.3	50.6	24799	46.5	22808	48.6	23839	48.8	23920	22.8	11187
Heterocigosis Poblacional										
Observada	0.378		0.354		0.373		0.384		0.372	
Esperada	0.367		0.359		0.369		0.375		0.367	
Heterocigosis Individual										
Media	0.380		0.363		0.378		0.386		0.370	
Máximo	0.398		0.385		0.391		0.393		0.398	
Mínimo	0.354		0.277		0.356		0.363		0.277	
D_{st} (media)	0.704		0.713		0.702		0.697		0.698	

Tabla 2. Parámetros de informatividad en las razas utilizadas en la elaboración del Ovine SNP50BeadChip.

	Awassi		Merino		Poll Dorset		Romney		Scottish Blackface		Texel		Razas secuenciadas para la elaboración del Chip	
Nº Animales	23		50		108		24		56		24		285	
MAF media	0.224		0.286		0.262		0.266		0.274		0.260		0.262	
	%	SNP	%	SNP	%	SNP	%	SNP	%	SNP	%	SNP	%	SNP
MAF < 0.01	13.9	6828	1.3	625	4.4	2135	4.5	2208	3.2	1582	5.2	2550	0.0	15
MAF > 0.3	37.8	18511	48.5	23806	44.2	21695	44.8	21986	47.1	23113	43.1	21121	2.5	1236
Heterocigosis Poblacional														
Observada	0.311		0.358		0.337		0.348		0.353		0.345		0.348	
Esperada	0.299		0.371		0.344		0.347		0.358		0.341		0.352	
Heterocigosis Individual														
Media	0.311		0.358		0.338		0.340		0.353		0.341		0.343	
Máximo	0.330		0.380		0.397		0.364		0.368		0.357		0.397	
Mínimo	0.266		0.289		0.272		0.297		0.329		0.290		0.266	
Dst (media)	0.762		0.697		0.724		0.713		0.712		0.721		0.672	

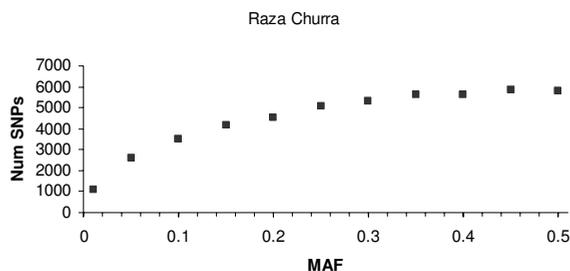


Figura 1. Distribución de los SNPs en función del MAF observado en la raza Churra.

ANALYSIS OF THE USEFULNESS of Ovine SNP50BeadChip (Illumina) FOR GENOMIC STUDIES IN SPANISH SHEEP BREEDS.

ABSTRACT: During the development of the Sheep HapMap project, ewes from 6 breeds (Awassi, Merino, Poll Dorset, Romney, Scottish Blackface and Texel) were used to obtain the sequence of the ovine genome and look for variability. The Ovine SNP50BeadChip (Illumina) was then genotyped in 71 breeds, four of which were Spanish breeds. In some species (pig and cow) a bias has been described when testing a SNP Chip in different breeds. The aim of this study was to analyse this possible bias due to the sheep breed in the Spanish populations included in the Sheep HapMap project. For this purpose, we estimated for both, the breeds used in the Chip development and the Spanish populations the following parameters: MAF within and between breeds, gene diversity, individual heterozygosity and average proportion of alleles shared within population. The results obtained in Spanish breed were similar to those observed in the six breeds sequenced during the SNP-chip development. These findings support the value of Ovine SNP50BeadChip for genomic analysis in Spanish sheep breeds.

Keywords: sheep, ovine SNP-chip, genetic variability