

## Identificación y validación de marcadores tipo SNPs en genes KRT33A (Keratine 33A) y FGF5 (Fibroblast Grown Factor 5)

William Salas<sup>1,\*</sup>, Diego Veliz<sup>2</sup>, Irene Delgado Flor<sup>2</sup>, Jorge Rodríguez<sup>2</sup>, José Espinoza<sup>2</sup>

<sup>1</sup> Facultad de Zootecnia, Universidad Nacional de Huancavelica, Avenida Universitaria 1420, Huancavelica, Perú.

<sup>2</sup> Unidad de Biotecnología Molecular de la Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430, San Martín de Porres, Lima, Perú.

### Resumen

El presente proyecto tuvo como objetivo identificar SNPs en el gen KRT33A y gen FGF5. Un total de 80 alpacas huacaya provenientes del fundo Lacchoc, departamento de Huancavelica fueron utilizadas para la identificación de SNPs mediante el análisis bioinformático, diseño y validación de cebadores específicos para la amplificación por PCR y posterior secuenciamiento. Los resultados sugieren la ausencia de polimorfismos en el gen FGF5 y la presencia de 2 SNPs en el gen de KRT33A. Se recomienda continuar con el análisis en un mayor número de genes e individuos.

Palabras clave: Alpacas; Genotipos; Queratina; Nucleótidos; Camélidos sudamericanos

### Identification and validation of SNPs markers in (Keratine 33A) and FGF5 (Fibroblast Grown Factor 5) genes

#### Abstract

The present project has to objective identificate SNPs in KRT33A and FGF5genes. A total of 80 alpacas huacaya from Lacchoc breeding center from Huancavelica province were used to SNP identification, primer design and validation using bioinformatic tools, PCR amplification and sequencing. The results suggest the absent of polymorphism in the FGF5 gene and the presence of 2 SNPs in the KRT33A gene. We suggest continue with the analysis of more genes and increase the alpaca sample.

Keywords: Alpaca; Genotype; Keratin; Nucleotides; South American camelids

## 1. Introducción

El Perú es el principal productor de fibra de alpaca en el mundo, porque cuenta con la mayor población de alpacas en el orden de los 3 millones de ejemplares [1]. Sin embargo, en las últimas décadas la alpaca peruana viene siendo sometida a una fuerte presión de selección para una mejora continua de las características textiles como el diámetro de fibra, principalmente [2]. La variabilidad genética en proteínas relacionadas con la constitución y calidad de la fibra podrían permitir realizar una selección para determinar características de interés económico, para lo cual estudios que identifiquen marcadores genéticos en poblaciones de alpacas constituyen una herramienta muy útil y primordial en la generación de programas de selección y mejoramiento genético. En la actualidad, se cuenta con un genoma de alpaca disponible en la web

(<https://www.ncbi.nlm.nih.gov/>), este recurso genético permite la búsqueda de potenciales marcadores genéticos para ser usados en procesos de selección asistida.

Las queratinas han sido considerados como buenos genes candidatos a estar asociados con distintas características de fibra en diferentes especies, incluyendo oveja y cabra, a los que se han asociado con alguna característica de fibra [3], mientras que genes involucrados en la regulación de la producción de fibra como el FGF5 ha sido asociada con la presencia de pelo largo en perros, gatos, conejos, ratón y ratas, habiéndose encontrado polimorfismos tipo SNPs en dicho gen en otras especies.

## 2. Metodología

### 2.1 Identificación de secuencias para los genes de queratinas KRT33A y FGF5 por

\*Correspondencia autor: [willysalasc@hotmail.com](mailto:willysalasc@hotmail.com)

### Análisis bioinformático

Información de tres genomas de alpaca pertenecientes a la Unidad de Biotecnología Molecular de la Universidad Peruana Cayetano Heredia (UPCH) fueron utilizados para la identificación de secuencias reguladoras y codantes y la presencia de potenciales polimorfismos presentes en los genes KRT33A y FGF5 mediante homología con secuencias reportadas para especies de ungulados.

#### 2.2 Diseño, síntesis y evaluación de cebadores

Se diseñaron cebadores de 18 a 24 nucleótidos con capacidad de amplificación de fragmentos entre 400 y 700 pares de bases, complementarios a los extremos de las secuencias que incluyan la presencia de SNPs putativos utilizando el programa *online* Oligo Analyzer™. La especificidad de los cebadores fue evaluada mediante la herramienta de NCBI 'Primer-BLAST' para las secuencias target de cada par de cebadores.

#### 2.3 Identificación de polimorfismos tipo SNPs en KRT33A y FGF5

Un total de 80 alpacas de la raza huacaya adultas no emparentadas fueron utilizadas para la identificación de polimorfismos de nucleótido único (SNPs) en los genes de KRT33A y FGF5.

ADN genómico fue extraído mediante GF-1 Nucleic Acid Extraction Kits - Vivantis Technologies siguiendo las instrucciones del fabricante. El ADN genómico obtenido fue diluido hasta una concentración final de 5 ng/uL. Fragmentos entre 500 y 700 pares de bases fueron amplificados mediante reacción en cadena de la polimerasa (PCR) utilizando los siguiente cebadores: para el exón 1 del gen FGF5: FGF5f 5'-GGC AAA CTGG AAG CAA GGT G-3' y FGF5r 5'-TTT CCC AAG GCT ATG TCC AC-3' y para el gen de KRT33A: 5'-AGA GCC GAG GAC TGC AAA CTC-5' y KRT33Ar 5'-GCCCTAAAGTAAAACCTGGAGC-3'

utilizando las siguientes condiciones: 5 nanogramos de ADN, 1X Buffer PCR, 1.5 mM de MgCl<sub>2</sub>, 0.2mM dNTPs 5 picomoles de cada cebador, 1U de ADN polimerasa (Thermo Scientific). Las condiciones termales para ambos genes fueron las siguientes: 95 °C por 3 minutos, 25 ciclos de 94 °C por 30 segundos, 60 °C por 45 segundos, 72 °C por 30 segundos.

Los productos de PCR fueron secuenciados por ambas direcciones usando ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit y un secuenciador automático ABI PRISM 3130 Genetic Analyser®. La edición de las secuencias y un alineamiento múltiple para la identificación de los SNPs se realizó mediante el Programa Mega v6.0 [4].

Una tabla de genotipos fue generada para cada gen y a partir de esa información se calcularon las frecuencias alélicas, número de alelos por locus, heterocigosidad observada y esperada, contenido de información polimórfica (PIC), probabilidad de exclusión [5] para los 2 genes identificados.

### 3. Resultados y Discusión

El exón 1 del gen FGF5 no presentó polimorfismos tipo SNPs en la totalidad de su secuencia, siendo la secuencia consenso la misma que la obtenida a partir del análisis de los genomas de la alpaca, observándose un alto nivel de conservación similar a lo reportado en otros mamíferos, en contraste con el gen KRT33A que presentó un total de 2 polimorfismos tipo SNPs (T/C, G/A) con frecuencias mayores al 30 %. La heterocigosidad esperada ( $H_E$ ) para los dos SNPs en KRT33A fueron 0.508 y 0.427. La probabilidad de exclusión varió entre 0.164 y 0.187 con una probabilidad de exclusión acumulada para el total de SNPs de 0.323. El número de alelos por locus, heterocigosidad observada y esperada, contenido de información polimórfica (PIC), probabilidad de exclusión para los SNPs del gen KRT33A se detallan en la Tabla 1.

**Tabla 1.** Alelos, frecuencia alélica, heterocigosidad observada (H<sub>O</sub>), heterocigosidad esperada (H<sub>E</sub>), contenido de información polimórfica (PIC) y probabilidad de exclusión (PE) para 2 SNPs del gen KRT33A.

<i>Locus</i>	<i>Nº alelos</i>	<i>Alelos</i>	<i>Frecuencia Alélica</i>	<i>H<sub>O</sub></i>	<i>H<sub>E</sub></i>	<i>PIC</i>	<i>PE</i>
K33-1	2	C/T	0.5;0.5	1.000	0,508	0,375	0,187
K33-2	2	A/G	0.7;0.3	0.000	0,427	0,332	0,164
Promedio	2			0,500	0,468	0,354	

Ambos SNPs en KRT33A se encontraron en desequilibrio de ligamiento ( $p < 0.05$ ) y desequilibrio de Hardy Weinberg ( $p < 0.05$ ), K33-1 debido a un exceso de heterocigotos ( $p < 0.05$ ) mientras que k33-2 debido a deficiencia de heterocigotos ( $p < 0.05$ ).

Estos resultados son similares a los observados para KRT33A y otros genes de queratinas y proteínas asociadas a queratinas en otras especies [6] siendo necesario realizar análisis de asociación para estas variantes, para determinar si se encuentran asociadas con características como diámetro de fibra.

Los resultados confirman la presencia de polimorfismos tipo SNPs en el gen KRT33A, con un fuerte desequilibrio de ligamiento entre ellos y desequilibrio de Hardy Weinberg, lo cual limita la capacidad de información que se puede obtener de ellos. Marcadores que se encuentran en un mismo gen presentan de moderado a altos niveles de desequilibrio de ligamiento debido a la cercanía entre ellos, así mismo la presencia de desequilibrio de Hardy Weinberg nos indica que alguna fuerza evolutiva está fuertemente afectada. Una fuerza evolutiva muy fuerte es la selección; sin embargo, el limitado número de muestras constituye un problema en el análisis de la importancia del polimorfismo e incluso en el análisis de equilibrio de Hardy Weinberg.

Se requiere de más estudios y un número muestral mucho mayor para caracterizar la variabilidad disponible en el gen de KRT33A y conocer las implicancias de este gen en las características de la fibra de alpaca.

#### 4. Conclusiones

Se identificaron 2 polimorfismos tipo SNPs en el gen KRT33A de alpaca, ambos en desequilibrio de ligamiento y desequilibrio de Hardy Weinberg.

El exón 1 del gen FGF5 presentó un alto nivel de conservación de secuencias sin la presencia de polimorfismos tipo SNPs.

#### 5. Agradecimientos

El presente proyecto de investigación fue financiado por el proyecto FOCAM-UNH: Identificación de marcadores genéticos moleculares SNPs para las características de importancia económica en alpaca huacaya. Agradecemos a la Dra. Patricia Herrera, Mg. Teresa Barreto, Mg (c) Francesco Foppiano, de la Unidad de Biotecnología Molecular de la Universidad Peruana Cayetano Heredia y al Proyecto: Una aproximación de genómica de poblaciones aplicada a la filogenia y mapeo de rasgos productivos en alpaca” Contrato 19-FINCYT-IA-2013, por la disponibilidad de los genomas y secuencias así como por el apoyo en el análisis de los datos.

#### 6. Bibliografía

- [1]. Ministerio de Agricultura y Riego. [Homepage]. Camélidos Sudamericanos. [acceso 2013]. Disponible en: <http://minagri.gob.pe/portal/40-sector-agrario/situacion-de-las-actividades-de-crianza-y-produccion/298-camelidos-sudamericanos?start=4>
- [2]. Brenes ER, Madrigal K, Pérez F, Valladares K. El Cluster de los Camélidos en Perú: Diagnóstico Competitivo y Recomendaciones Estratégicas. Documento de Trabajo. Instituto Centroamericano de Administración de Empresas INCAE; 2001. Disponible en: [http://infoalpacas.com.pe/wp-content/uploads/2015/12/cluster\\_camelidos\\_peru.pdf](http://infoalpacas.com.pe/wp-content/uploads/2015/12/cluster_camelidos_peru.pdf)
- [3]. Liu GF, Tian KC, Zhang EP, Huang, XX, Zhang YH Candidate gene analysis of high quality merino sheep. Yi Chuan. 2007; 29(1): 70-74.
- [4]. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular evolutionary genetics

analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*. 2007; 24: 1596-1599.

[5]. Jamieson A, Taylor SC. Comparisons of three probability formulae for parentage exclusion. *Animal Genetics*. 1997; 28: 397-400.

[6]. Wang X, Zhao ZD, Xu HR, Qu L, Zhao HB, Li T, Zhang ZY. Variation and expression of KAP9.2 gene affecting cashmere trait in goats. *Molecular Biology Reports*. 2012; 39(12): 10525-10529.