

**UNIVERSIDAD NACIONAL AGRARIA**

**LA MOLINA**

**FACULTAD DE ZOOTECNIA**



**“IDENTIFICACIÓN BIOINFORMÁTICA DE POLIMORFISMOS DE NUCLEÓTIDO SIMPLE (PNS) EN GENES DE PROTEÍNAS ASOCIADAS A QUERATINAS EN ALPACAS (*Vicugna pacos*)”**

Presentada por:

**DEYANIRA ANTONELLA FIGUEROA VENEGAS**

Tesis para Optar el Título Profesional de:

**INGENIERO ZOOTECNISTA**

Lima – Perú

**2022**

---

La UNALM es la titular de los derechos patrimoniales de la presente investigación  
(Art. 24. Reglamento de Propiedad Intelectual)

**UNIVERSIDAD NACIONAL AGRARIA  
LA MOLINA**

**FACULTAD DE ZOOTECNIA**

**“IDENTIFICACIÓN BIOINFORMÁTICA DE POLIMORFISMOS  
DE NUCLEÓTIDO SIMPLE (PNS) EN GENES DE PROTEÍNAS  
ASOCIADAS A QUERATINAS EN ALPACAS (*Vicugna pacos*)”**

Presentada por:

**DEYANIRA ANTONELLA FIGUEROA VENEGAS**

Tesis para Optar el Título Profesional de:

**INGENIERO ZOOTECNISTA**

Sustentada y aprobada ante el siguiente jurado:

---

Ph.D. Humberto Blas Sevillano  
Presidente

---

Ph.D. Jorge Calderón Velásquez  
Miembro

---

Ph.D. Alberto Barrón López  
Miembro

---

Ph.D. Gustavo Gutiérrez Reynoso  
Asesor

---

Ph.D. Federico Abel Ponce de León Bravo  
Co Asesor

## **DEDICATORIA**

A Dios, por guiar siempre mi camino.

A mi mamá Katty y mi papá Beto, por ser mi mayor soporte y mi más grande apoyo.

A mi hermano Junior, por ser mi motor y mi razón de ser mejor cada día.

A las familias Figueroa, Venegas y Purizaga, por estar presentes en cada paso que doy y celebrar cada uno de mis logros.

A Edelmira Purizaga†, Enrique Venegas† y Leoncio Ortiz† que no pudieron enterarse de este logro.

A Damaris†, porque su compañía fue parte importante de mi vida.

## **AGRADECIMIENTOS**

Al Proyecto N° 028-2016 INIA-PNIA/UMPSI/IE, sus colaboradores e instituciones asociadas, por permitir que se realice esta investigación, por capacitarme y darme la oportunidad de conocer a estupendos investigadores.

Al Ph.D. Gustavo A. Gutiérrez y Ph.D. Abel Ponce de León, por su constante apoyo y su invaluable asesoramiento durante esta investigación.

A los miembros del jurado, por sus comentarios, correcciones y consejos.

Al Vicerrectorado de Investigación por el uso del Servidor HPC Bioinformática – UNALM y el Programa de Mejoramiento Animal de la Facultad de Zootecnia, por las facilidades brindadas.

A Manuel More y Álvaro Fernández, por compartir sus conocimientos para la realización de esta investigación.

A mi promoción de ingreso y la MDZ-2018 que llenaron de buenos momentos mi paso por la universidad.

A todos aquellos amigos, compañeros y profesores, que han formado parte de todo este proceso.

A la mejor facultad, Zootecnia y mi alma mater, la Universidad Nacional Agraria La Molina, por darme la oportunidad de conocer a un gran grupo humano y por permitirme gozar de todas esas experiencias/tradiciones que la caracterizan.

Familia, no sólo les dedico este trabajo, sino que también les agradezco por todo su apoyo, paciencia y comprensión. Los amo.

## ÍNDICE GENERAL

RESUMEN.....	vi
ABSTRACT.....	vii
I. INTRODUCCIÓN .....	1
II. REVISIÓN DE LITERATURA.....	3
2.1. LOS CAMÉLIDOS SUDAMERICANOS.....	3
2.2. LA ALPACA.....	3
2.3. LA FIBRA DE ALPACA.....	4
2.4. ALFA QUERATINAS.....	5
2.5. PROTEÍNAS ASOCIADAS A LAS QUERATINAS.....	5
2.6. GENES RELACIONADOS AL DESARROLLO DE LA FIBRA.....	6
2.7. POLIMORFISMOS DE NUCLEÓTIDO SIMPLE.....	8
2.8. POLIMORFISMOS DE NUCLEÓTIDO SIMPLE EN ALPACAS.....	9
2.9. CENTRO NACIONAL DE INFORMACIÓN BIOTECNOLÓGICA (NCBI).....	10
2.10. GENOMAS DE REFERENCIA EN ALPACAS.....	10
III. METODOLOGÍA.....	11
3.1. DATOS GENÓMICOS.....	11
3.2. REVISIÓN Y OBTENCIÓN DE UN LISTADO DE GENES DE PROTEÍNAS ASOCIADAS A LAS QUERATINAS EN ALPACAS.....	11
3.3. LOCALIZACIÓN DE GENES DE LAS PROTEÍNAS ASOCIADAS A LAS QUERATINAS EN EL VICPAC 3.1.....	12
3.4. IDENTIFICACIÓN DE POLIMORFISMOS DE NUCLEÓTIDO SIMPLE.....	12
3.5. SELECCIÓN DE PNSs.....	13
IV. RESULTADOS Y DISCUSIONES.....	16
4.1. GENES DE PROTEÍNAS ASOCIADAS A LAS QUERATINAS EN ALPACAS.....	16
4.2. POLIMORFISMOS DE NUCLEÓTIDOS SIMPLES (PNSs) EN <i>KRTAPs</i> EN ALPACAS.....	31
4.2.1. GENES QUE CODIFICAN PROTEÍNAS CON ALTO CONTENIDO DE SULFURO.....	31
4.2.2. GENES QUE CODIFICAN PROTEÍNAS CON MUY ALTO CONTENIDO DE SULFURO.....	31

4.2.3. GENES QUE CODIFICAN PROTEÍNAS CON ALTO CONTENIDO EN GLICINA Y TIROSINA .....	31
4.3. DISCUSIÓN .....	35
V. CONCLUSIONES .....	43
VI. RECOMENDACIONES .....	44
VII. BIBLIOGRAFÍA.....	45
VIII. ANEXOS.....	52

## ÍNDICE DE TABLAS

Tabla 1: Clasificación de los genes de queratina en los mamíferos. ....	7
Tabla 2: Clasificación de los genes codificantes de las proteínas asociadas a las queratinas en mamíferos. ....	8
Tabla 3: Lista de <i>KRTAPs</i> en alpacas .....	16
Tabla 4: Listado de genes <i>KRTAP</i> de alpacas, cabras y ovejas equivalentes en nomenclatura.....	17
Tabla 5: Comparación entre las secuencias de genes <i>KRTAP</i> de oveja que no tienen un equivalente en nomenclatura en alpaca.....	19
Tabla 6: Comparación entre las secuencias de genes <i>KRTAP</i> de cabra que no tienen un equivalente en nomenclatura en alpaca.....	20
Tabla 7: Comparación entre las secuencias de genes <i>KRTAP</i> de oveja y cabra que no tienen un equivalente en nomenclatura en alpaca.....	21
Tabla 8: Porcentaje de similitud de <i>KRTAPs</i> anotados en los genomas de referencia entre alpaca, cabra y oveja. ....	22
Tabla 9: Ubicación de genes, polimorfismos de nucleótido simple (PNSs) y secuencias flanqueantes a los PNSs (35pb) en el VicPac3.1 de <i>KRTAPs</i> que codifican proteínas con alto contenido de sulfuro. ....	25
Tabla 10: Ubicación de genes, polimorfismos de nucleótido simple (PNSs) y secuencias flanqueantes a los PNSs (35pb) en el VicPac3.1 de <i>KRTAPs</i> que codifican proteínas con muy alto contenido de sulfuro. ....	32
Tabla 11: Ubicación de genes, polimorfismos de nucleótido simple (PNSs) y secuencias flanqueantes a los PNSs (35pb) el VicPac3.1 de <i>KRTAPs</i> que codifican proteínas con alto contenido en glicina y tirosina.....	33

## ÍNDICE DE FIGURAS

Figura 1: Esquema de la formación de filamentos intermedios. ....	5
Figura 2: Polimorfismo de nucleótido simple. ....	9
Figura 3: Configuración de entrada para realizar la comparación del genoma de la alpaca con el gen <i>SHEEP-KRTAP1-1</i> haciendo uso de la herramienta BLAST. ....	12
Figura 4: Ubicación del exón.....	14



## ÍNDICE DE ANEXOS

Anexo 1: Valores de los parámetros considerados para cada PNS de los genes <i>KRTAPs</i> que codifican proteínas con alto contenido de sulfuro.....	53
Anexo 2: Valores de los parámetros considerados para cada PNS de los genes <i>KRTAPs</i> que codifican proteínas con muy alto contenido de sulfuro.....	57
Anexo 3: Valores de los parámetros considerados para cada PNS de los genes <i>KRTAPs</i> que codifican proteínas con alto contenido en glicina y tirosina. ....	58
Anexo 4: Alineamiento de las secuencias de los genes <i>KRTAP1-1</i> , <i>KRTAP3-2</i> y <i>KRTAP4-12L</i> , con la herramienta Clustal Omega. ....	59
Anexo 5: Alineamiento de la secuencia del gen <i>KRTAP5-4</i> con la herramienta Clustal Omega.....	61
Anexo 6: Alineamiento de la secuencia del gen <i>KRTAP9-7L</i> (LOC114116849) con la herramienta Clustal Omega.....	62
Anexo 7: Alineamiento de las secuencias del gen <i>KRTAP13-2L</i> y las dos copias del gen <i>KRTAP10-7L</i> (LOC114110483 y LOC101104456) , con la herramienta Clustal Omega.....	63
Anexo 8: Alineamiento de las secuencias de los genes <i>KRTAP1-4</i> , <i>KRTAP4-11L</i> y <i>KRTAP9-3L</i> con la herramienta Clustal Omega. ....	66
Anexo 9: Alineamiento de la secuencia del gen <i>KRTAP9-8</i> con la herramienta Clustal Omega.....	68
Anexo 10: Alineamiento de las secuencias de los genes <i>KRTAP10-9L</i> y <i>KRTAP16-4</i> con la herramienta Clustal Omega.....	69
Anexo 11: Alineamiento de las secuencias de los genes <i>KRTAP6-1L</i> , <i>KRTAP1-3</i> y <i>KRTAP9-9L</i> con la herramienta Clustal Omega.....	71
Anexo 12: Alineamiento de las secuencias de los genes <i>KRTAP4-8L</i> y <i>KRTAP4-9L</i> con la herramienta Clustal Omega.....	73

## RESUMEN

El objetivo del presente trabajo fue identificar polimorfismos de nucleótido simple (PNSs) en genes de proteínas asociadas a las queratinas (*KRTAPs*) en alpacas (*Vicugna pacos*). Los objetivos específicos fueron: a) identificar las secuencias de genes de proteínas asociadas a queratinas en el VicPac3.1, b) identificar PNSs en la secuencia de los genes (exones, intrones o regiones no traducidas), y c) seleccionar los PNSs identificados, considerando parámetros de calidad (frecuencia del alelo menor, tamaño de la secuencia flanqueante y ausencia de PNS en la secuencia flanqueante). Se realizó el estudio en un total de 34 genes de proteínas asociadas a las queratinas, que se encuentran anotados en la base de datos del National Center for Biotechnology Information (NCBI). Para hallar los PNSs se comparó el genoma de referencia VicPac3.1. con nueve genomas (almacenados en NCBI) y *reads* de 300 bibliotecas de representación reducida de ADN de alpacas, mediante el uso del BLASTN. Se halló la frecuencia del alelo menor (MAF) y la tasa de genotipificación (TG) con secuencias en los sitios del PNS, mediante el uso de los softwares KGD y PLINK, así como, el Illumina Score, mediante el Illumina Design Studio para cada PNS. Se seleccionaron marcadores con frecuencia de alelo menor (MAF)  $\geq 0.05$ , tasa de genotipificación (TG)  $\geq 45\%$  por PNS basado en secuencias *reads* y Illumina Score = 0.6. Se identificaron 67 PNSs ubicados en las regiones intrónicas, exónicas y no traducidas de los genes de proteínas asociadas a las queratinas, que cumplen con dichos parámetros. De estos, 35 PNSs fueron incluidos en una micromatriz de PNSs de Alpaca de los cuales 32 PNSs fueron confirmados.

**Palabras clave:** Alpaca, PNS, proteínas asociadas a queratinas.

## ABSTRACT

The aim of this study was to identify single nucleotide polymorphisms (SNPs) in keratin-associated protein genes, in alpacas (*Vicugna pacos*). The specific aims were a) identify the keratin-associated protein gene sequences in the VicPac3.1, b) identify SNPs in the gene sequences (exons, introns and/or untranslated regions), and c) select the identified SNPs, considering quality parameters (minor allele frequency, length of the flanking sequence and absence of additional SNPs in the flanking regions). The study was conducted on a total of 34 keratin-associated protein genes (*KRTAPs*), which are annotated in the database of the National Center for Biotechnology Information (NCBI). To identify the SNPs, the reference genome VicPac3.1 was compared to nine genomes (stored in NCBI) and reads from 300 alpaca reduce representation DNA libraries, using BLASTN. The minor allele frequency (MAF) and genotyping rate were calculated using KGD and PLINK software. While the Illumina Score was calculated through the Illumina Design Studio software for each SNP. Markers with  $MAF \geq 0.05$ , a genotyping rate  $> 45\%$  per SNP based on read sequences and an Illumina Score  $\geq 0.6$  were selected. Sixty-seven SNPs identified in intron, exon and/or untranslated regions of the keratin-associated protein genes, met these parameters. Of these, 35 SNPs were included in an alpaca SNP microarray and 32 SNPs were confirmed as true SNPs.

**Keywords:** Alpaca, SNP, keratin-associated proteins.

## I. INTRODUCCIÓN

Los Camélidos Sudamericanos pertenecen al Orden Artiodáctila y habitan a más de 4000 m.s.n.m., a temperaturas de entre -15 y 20°C. (Pérez-Cabal *et al.*, 2010). La alpaca, es uno de los dos Camélidos Sudamericanos domésticos y es criada principalmente para la producción de fibra. Las fibras de alpaca son flexibles y suaves al tacto, poco alergénicas y de bajo afieltramiento (Contreras, 2009).

La fibra de alpaca, está compuesta por alfa queratinas y proteínas asociadas a las queratinas. Estas últimas, son un grupo complejo de proteínas, que poseen altos niveles de cisteína, o glicina y tirosina, debido a ello, se cree que estas proteínas desempeñan un papel importante en la definición de las propiedades físico-mecánicas de la fibra, por lo que la variación en la expresión de los genes que codifican a las proteínas asociadas a las queratinas en la alpaca, puede estar relacionada a una variación en la calidad de la fibra. (Gong *et al.*, 2012; Powell y Rogers, 1997)

El mejoramiento genético de las alpacas a través del uso de la selección, es de interés de los criadores y de la industria textil asociada al procesamiento de la fibra de alpaca. El criador busca mejorar sus ingresos vía el incremento de la cantidad de fibra por animal esquilado y disminuyendo el diámetro de fibra de sus vellones, y la industria busca adquirir un mayor volumen de fibra de calidad extrafina y fina, las cuales destinará a la producción de textiles de mayor valor económico en el mercado (Gutiérrez, 2008).

Los marcadores moleculares son secuencias de bases nucleotídicas que están distribuidos a lo largo del ADN y se utilizan para seguir la transmisión de un segmento de cromosoma de una generación a otra, permitiendo identificar y/o diferenciar genotipos (FAO, 2007; Dunner y Cañon, 2001). Un tipo de marcador molecular de ADN, es el polimorfismo de nucleótido simple (PNS) que es usado en el mejoramiento genético en diferentes especies animales (selección genómica), ya que son muy frecuentes en el genoma y gracias a la constante evolución de las tecnologías moleculares, se está aumentando la automatización y disminuyendo el costo de genotipificación de PNSs (Balding, 2006; FAO, 2010).

Dentro de esta perspectiva, la presente investigación tiene como objetivo general, identificar marcadores moleculares del tipo PNS en genes que codifican las proteínas asociadas a queratinas (*KRTAPs*) en alpaca (*Vicugna pacos*), teniendo como objetivos específicos; (1) identificar las secuencias de genes de proteínas asociadas a queratinas en el VicPac3.1, (2) identificar PNSs en la secuencia de los genes (exones, intrones y regiones no traducidas) de genes de proteínas asociadas a queratinas, y (3) seleccionar los PNSs identificados, considerando parámetros de calidad (frecuencia del alelo menor, tamaño de la secuencia flanqueante, entre otros).

## II. REVISIÓN DE LITERATURA

### 2.1. LOS CAMÉLIDOS SUDAMERICANOS

Los camélidos llegaron al continente norteamericano por el estrecho de Bering hace 45 millones de años aproximadamente, a partir de un pequeño antecesor de 30cm de talla, un grupo de estos animales migró a Sudamérica formando así los camélidos del nuevo mundo o sudamericanos (Stanley *et al.*, 1994). No existe acuerdo en cuanto a la cronología y si existieron uno o varios centros de domesticación, aunque se acepta por la mayoría de los investigadores que el origen de la domesticación de los camélidos sudamericanos se restringe a los Andes (Mengoni, 2008). Actualmente, habitan en Sudamérica cuatro especies de camélidos; dos especies silvestres: guanaco (*Lama guanicoe*) y vicuña (*Vicugna vicugna*) y dos especies domésticas: alpaca (*Vicugna pacos*) y llama (*Lama glama*) (Bonavia, 1996; Gentry *et al.*, 2004).

Las alpacas, llamas y vicuñas habitan la zona Altoandina, por encima de 3000 m.s.n.m., del Perú, Bolivia, Argentina y Chile, mientras que los guanacos habitan predominantemente zonas más bajas y desérticas, como la Patagonia Argentina y Chilena (Quispe *et al.*, 2009).

### 2.2. LA ALPACA

Basados en estudios de ADN mitocondrial, la alpaca (*Vicugna pacos*) ha sido clasificada taxonómicamente (Kadwell *et al.*, 2001; Gentry *et al.*, 2004; Marín *et al.*, 2007), de la siguiente manera:

Clase:	<i>Mamalia</i>
Orden:	<i>Artiodáctyla</i>
Suborden:	<i>Tylópoda</i>
Familia:	<i>Camelidae</i>
Tribu:	<i>Lamini</i>
Género:	<i>Vicugna</i>
Especie:	<i>Vicugna pacos</i>

La alpaca es un mamífero que se adapta a altitudes extremas, tolera bien las heladas y tiene muchas ventajas sobre el ganado introducido como el ovino y el vacuno. Por lo general, se cría principalmente para la producción de fibra y de manera complementaria para la carne. La industria textil se refiere a las fibras de alpaca como fibras especiales y los artículos confeccionados con ellas, están clasificados como artículos de lujo (Wang *et al.*, 2003). Quienes llevan prendas de alpaca, las eligen debido a sus cualidades térmicas, así como por otros atributos como su impermeabilidad y anti inflamabilidad (Czaplicki, 2012).

Hay dos razas de alpacas, la Huacaya y la Suri. La alpaca Huacaya se caracteriza por tener un vellón compacto y esponjoso, que se asemeja al vellón del ovino Corriedale que le atribuye una apariencia más voluminosa, con fibras finas suaves y onduladas, mientras que la alpaca Suri se caracteriza por tener fibras de gran longitud organizadas en rizos colgantes, que se asemeja a los rizos del ovino Lincoln, lo cual le atribuye al animal una apariencia angulosa (Hoffman y Fowler, 1995; Antonini *et al.*, 2004).

### **Población de alpacas**

La población de alpacas a nivel nacional es estimada en 3'685,500 cabezas, la que representa el 80% de la población mundial de alpacas. La mayor parte de la población pertenece a la raza Huacaya (80.4%), seguida de la raza Suri (12.2%) y finalmente las de fenotipo cruzado (7.3%) (FAO, 2005). De todos los departamentos del Perú, Puno es el que concentra la mayor población de alpacas (50%), seguida de otros departamentos como: Cusco, Arequipa, Huancavelica, Ayacucho y Apurímac (INEI, 2012). El Perú es considerado el principal productor de fibra de alpaca en el mundo. El 85% de la producción nacional es para el mercado internacional y representa el 1.35% de las exportaciones peruanas (FAO, 2005).

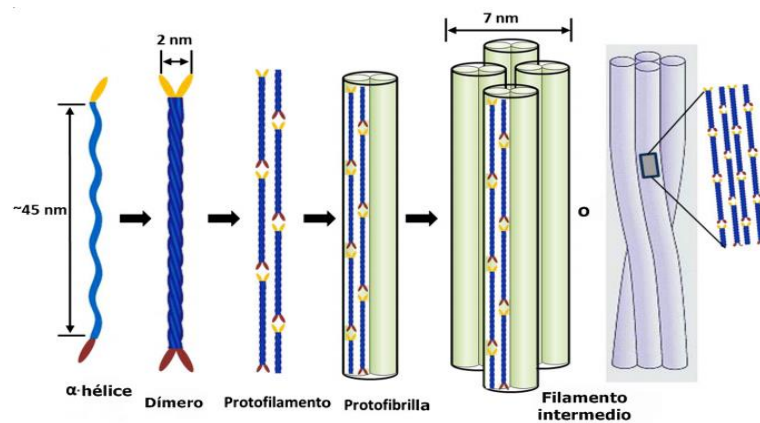
### **2.3. LA FIBRA DE ALPACA**

Cada fibra de alpaca se desarrolla a partir de una estructura llamada folículo piloso (Orwin, 1979). Esta estructura es muy conservada en los distintos taxones de los mamíferos (Rogers, 2004). Sin embargo, el pelo o la fibra pueden variar considerablemente de especie a especie, estas diferencias se deben en parte a las familias de proteínas estructurales que la componen (Khan *et al.*, 2014). Dos familias de proteínas (alfa queratinas y proteínas asociadas a queratinas (KAPs)), se expresan principalmente en las células estructurales de la fibra, (Powell y Rogers, 1997).

## 2.4. ALFA QUERATINAS

Las alfa queratinas (queratinas para el resto del texto) pertenecen a la superfamilia de proteínas capaces de formar filamentos intermedios y se pueden clasificar en dos grupos, las del tipo I que son pequeñas y de carácter ácido y las del tipo II que son grandes y de carácter básico o neutro (Powell y Rogers, 1997; Rogers *et al.*, 2006).

Para la formación de los filamentos intermedios, se agrupan una cantidad equimolar de queratina del tipo I y otra del tipo II, de manera antiparalela y se mantienen unidas por efecto hidrofóbico, luego se forman los protofilamentos (dos heterodímeros), que forman las protofibrillas (dos protofilamentos) y finalmente cuatro protofibrillas constituyen un filamento intermedio (Figura 1) (Höcker, 2002; Wang *et al.*, 2016).



**Figura 1:** Esquema de la formación de filamentos intermedios.  
**FUENTE:** Modificado de Wang *et al.*, 2016.

## 2.5. PROTEÍNAS ASOCIADAS A LAS QUERATINAS

Las proteínas asociadas a las queratinas (KAPs), constituyen la matriz que rodea a los filamentos intermedios de queratina en las células corticales y de la cutícula (Rogers *et al.*, 2006). Son un grupo de hasta más de 100 proteínas diferentes, que presentan alta divergencia, lo que es consistente con la diversidad de fenotipos de pelo que se pueden observar en mamíferos (Powell y Rogers, 1997; Wu *et al.*, 2008). Esta familia se describió por primera vez al aislar proteínas del folículo de oveja (Gillespie y Broad 1972; Gillespie, 1972).

Las KAPs interactúan con los filamentos intermedios durante el desarrollo de la fibra y se unen covalentemente a éstos, a través de un enlace cruzado extenso entre la cisteína en las KAP y en la cabeza y la cola de las queratinas (Powell y Rogers, 1997). Mientras que las



KAPs pueden tener poco, o ningún efecto perceptible en la estructura de los filamentos intermedios, su efecto en el ensamblaje de éstos en arreglos más grandes, es considerado como crucial. Por lo tanto, se cree que las KAPs desempeñan un papel importante en la definición de propiedades físico-mecánicas de la fibra (Plowman, 2003; Gong *et al.*, 2016).

Las KAPs se dividen en dos grupos principales, las que presentan un alto contenido de sulfuro y las que presentan un alto contenido de glicina y tirosina. El primero se puede separar en dos subgrupos, las proteínas con alto contenido de sulfuro con menos de 30% de residuos de cisteína y las de muy alto contenido de sulfuro que presentan más del 30% de residuos de cisteína (Gillespie y Broad 1972; Gillespie, 1972).

## **2.6. GENES RELACIONADOS AL DESARROLLO DE LA FIBRA**

Las fibras están compuestas principalmente de proteínas de queratina, que son codificadas por genes de queratina (*KRT*) y genes de proteínas asociadas a las queratinas (*KRTAP*). Estos genes se expresan de manera altamente regulada durante el crecimiento del folículo piloso y fibra, reflejado por la distribución ordenada temporal y espacial de su ARNm y proteínas traducidas (Powell y Rogers, 1997; Langbein y Schweizer, 2005; Yu *et al.*, 2009).

La nomenclatura actual de los genes de queratina, es la propuesta por Schweizer *et al.* (2006), siguiendo la forma:

*KRT#*\$

Siendo “#” un número que depende del tipo y ubicación del gen (Tabla 1); y “\$” una letra que representa la variante o alelo.

**Tabla 1:** Clasificación de los genes de queratina en los mamíferos.

Ubicación	Filamentos intermedios	
	Tipo I	Tipo II
Epiteliales		<i>KRT1</i> al <i>KRT5</i>
	<i>KRT9</i>	<i>KRT6a</i> , <i>KRT6b</i> ,
	<i>KRT12</i> al <i>KRT20</i>	<i>KRT6c</i>
	<i>KRT23</i> al <i>KRT25</i>	<i>KRT7</i> , <i>KRT8</i>
		<i>KRT71</i> al <i>KRT80</i>
Fibra del folículo	<i>KRT31</i> , <i>KRT32</i>	<i>KRT81</i> al <i>KRT87</i>
	<i>KRT33a</i> , <i>KRT33b</i>	
	<i>KRT34</i> al <i>KRT40</i>	

**FUENTE:** Modificado de Schweizer *et al.*, 2006

Por otro lado, los genes de proteínas asociadas a queratinas (*KRTAPs*) que son pequeños en masa atómica (10-30 kDa) y en su mayoría sin intrones (Rogers, 2006), se encuentran divididos en tres grupos (Tabla 2).

La nomenclatura actual de los genes de las proteínas asociadas a las queratinas es la propuesta por Gong *et al.* (2012), siguiendo la forma:

$$\text{sp-KRTAP}_{m-n(p/L)*x}$$

Siendo “sp” un código único basado en letras, obtenidas de la base de datos Uniprot (<https://www.uniprot.org/docs/speclist>), por ejemplo “SHEEP” para *Ovis aries*, “VICPA” para *Vicugna pacos* y “CAPHI” para *Capra hircus*, “m” es el número que identifica a la familia, “n” es el número del componente dentro de la familia, “p” significa pseudogen, mientras que “L” significa “like (parecido)” que hace referencia a una nomenclatura temporal hasta que se pueda comprobar que el gen pertenece a dicha familia o componente. Por último, “x” representa el número de variante o alelo.

**Tabla 2:** Clasificación de los genes codificantes de las proteínas asociadas a las queratinas en mamíferos.

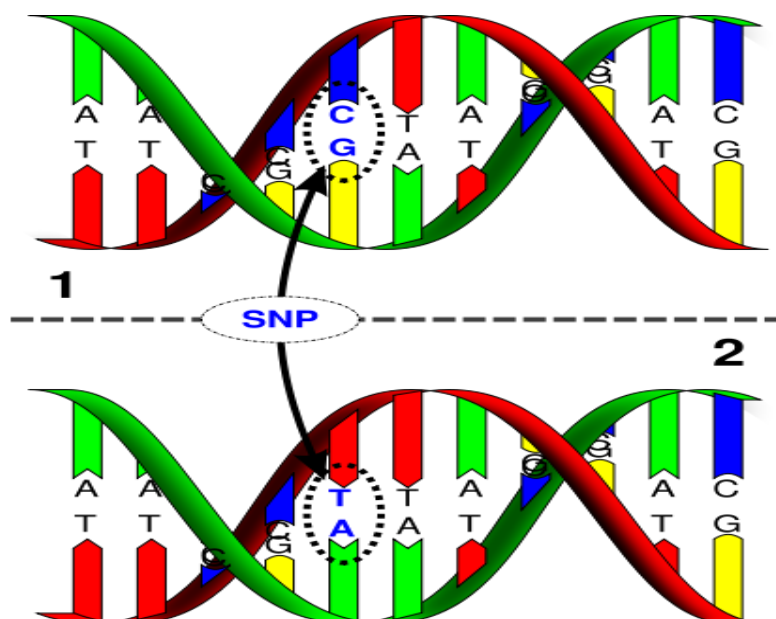
<b>Grupo</b>	<b>Familia</b>
Alto contenido de sulfuro (<30 mol% de cisteína)	<i>KRTAP1</i> al <i>KRTAP3</i> <i>KRTAP10</i> al <i>KRTAP16</i> <i>KRTAP23</i> al <i>KRTAP35</i>
Muy alto contenido de sulfuro (>30 mol% de cisteína)	<i>KRTAP4</i> , <i>KRTAP5</i> <i>KRTAP9</i> , <i>KRTAP17</i>
Alto contenido en glicina y tirosina (35-60 mol% de glicina y tirosina)	<i>KRTAP6</i> al <i>KRTAP8</i> <i>KRTAP19</i> al <i>KRTAP22</i>

**FUENTE:** Adaptado de Gong *et al.*, 2012 y Khan *et al.*, 2014.

## 2.7. POLIMORFISMOS DE NUCLEÓTIDO SIMPLE

Los polimorfismos de nucleótido simple (PNSs), consisten en diferencias de base única en la secuencia de ADN (Figura 2) entre individuos de una especie, son numerosos y están presentes a lo largo de los genomas. Los PNSs son valiosos como marcadores en los estudios de asociación debido a su variabilidad y a la incidencia amplia a lo largo del genoma; y también son útiles para determinar las relaciones de parentesco ya que la mayoría de los PNSs son únicos dentro de una población. (Pierce, 2006)

Los PNSs están distribuidos por todo el genoma y se pueden nombrar según la ubicación que tengan: iPNS, si están localizados en regiones intrónicas; cPNS, si están localizados en regiones codificantes (exones); rPNS, si están localizados en regiones reguladoras; y gPNS, si están localizados en regiones intergenómicas. Además, los cPNS pueden ser sinónimos (sPNS) o no sinónimos (nsPNS). (Cargill M, *et al.* 1999)



**Figura 2:** Polimorfismo de nucleótido simple.  
**FUENTE:** <https://knowgenetics.org/snps/>

El interés en los PNSs proviene y continúa dependiendo de la fusión de varias áreas de investigación, como el análisis del genoma a gran escala y la bioinformática (Brookes, 1999).

## 2.8. POLIMORFISMOS DE NUCLEÓTIDO SIMPLE EN ALPACAS

De los estudios realizados de PNSs en alpacas, es importante mencionar que Salas (2019) encontró 34 variantes tipo PNS en genes involucrados en la estructura y desarrollo de la fibra de alpaca (*KRT31*, *HOXC13* y *EDAR*); sin embargo, no se encontró asociación entre el diámetro de la fibra de alpaca y los fragmentos de los genes *KRT31*, *HOXC13* y *EDAR* analizados. Asimismo, Foppiano (2016) identificó 27 marcadores PNSs en cinco genes de la familia de proteínas asociadas a la queratina (*KRTAP1-2*, *KRTAP6-1*, *KRTAP9-2*, *KRTAP11-1*, *KRTAP13-1*), y realizó un estudio de asociación entre el diámetro de fibra y dos PNSs del gen *KRTAP11-1*, pero no encontró una asociación significativa en muestras de 152 alpacas huacaya blancas de la región Puno. Además, Fernández *et al.* (2019) identificó en alpacas un total de 48 PNSs en 22 genes de queratina de fibra, realizando un análisis bioinformático basándose en genes ortólogos de ovino, caprino y humano. Por último, Delgado de la Flor (2014) caracterizó marcadores PNSs en el gen de Tricohialina e identificó dos PNSs que se encuentran en desequilibrio de ligamiento con el gen Tricohialina, pero al analizar su posible asociación con el fenotipo Suri y/o Huacaya, no se encontró asociación de estos PNSs con dichos fenotipos. En 2021, Calderon *et al.* reportó el desarrollo de una

micromatriz de 76k de PNSs, de alpaca, en la cual incluyó 302 PNSs que se encontraban en genes candidatos para calidad de fibra y color.

## **2.9. CENTRO NACIONAL DE INFORMACIÓN BIOTECNOLÓGICA (NCBI)**

El Centro Nacional de Información Biotecnológica (National Center for Biotechnology Information o NCBI) se creó en 1988 para desarrollar sistemas de información que puedan ser utilizados en la investigación de la biología molecular. Mantiene bases de datos de secuencias de ácidos nucleicos y bases de datos de secuencias de nucleótidos que recibe de tres fuentes: presentaciones directas de investigadores externos; colaboraciones o alianzas nacionales e internacionales con proveedores de datos y consorcios de investigación; y esfuerzos internos de conservación de datos. Además, el NCBI proporciona sistemas de recuperación de datos y recursos computacionales para el análisis de datos como PubMed, PubMed Central, Gene, the NCBI Taxonomy Browser, BLAST, SPLIGN, PRIMER-BLAST, etc. Todos los recursos con los que cuenta el NCBI se encuentran en [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) (NCBI Resource Coordinators, 2013).

## **2.10. GENOMAS DE REFERENCIA EN ALPACAS**

El desarrollo de genomas de referencia de alta calidad cromosómicamente asignados, constituye una característica clave de la arquitectura del genoma de una especie y es fundamental para el descubrimiento de las huellas genéticas de los rasgos de importancia biológica. (Richardson *et al.*, 2019)

En la actualidad, hay tres conjuntos de secuencia anotados para la alpaca que están disponibles en navegadores como NCBI y Ensembl: VicPac1 (versión 1.0), VicPac2 (versión 2.0.1) y VicPac3.1. Todos usaron el ADN de la misma hembra Huacaya individual. El primero se generó en el Broad Institute y tiene una cobertura de 2.51X, el segundo fue montado en la Universidad de Washington y tiene una cobertura de 22X, el tercero es un nuevo montaje que contiene el 90% del genoma de alpaca en 103 scaffolds y el 76% de todos los scaffolds han sido asignados a los 36 pares de autosomas de alpaca y el cromosoma X. (NCBI, 2019; Richardson *et al.*, 2019). El genoma de referencia de alpaca VicPac3.1 presenta una mejora significativa en integridad, contigüidad y precisión sobre VicPac2 y es una herramienta importante para el avance de la investigación genómica en todos los camélidos del Nuevo Mundo (Richardson *et al.*, 2019).

### **III. METODOLOGÍA**

#### **3.1. DATOS GENÓMICOS**

La información genómica utilizada en el presente trabajo fue:

- Un genoma de referencia: VicPac3.1, (Número de acceso GenBank: GCA\_000164845.4).
- Tres genomas (Número de acceso: PRJNA340289) publicados por la Universidad Peruana Cayetano Heredia (UPCH).
- Seis genomas (Número de acceso: PRJEB42148) y 150 bibliotecas de representación reducida de ADN elaboradas con la enzima ApeKI y 150 de doble digestión con las enzimas PstI y MspI realizados por el proyecto N° 028-2016-INIA-PNIA/UPMSI/IE ejecutado por la Universidad Nacional Agraria La Molina (UNALM).

#### **3.2. REVISIÓN Y OBTENCIÓN DE UN LISTADO DE GENES DE PROTEÍNAS ASOCIADAS A LAS QUERATINAS EN ALPACAS**

Se buscó genes de proteínas asociadas a las queratinas (*KRTAPs*) en alpacas en la base de datos del National Center for Biotechnology Information (NCBI). Asimismo, se realizó un análisis comparativo con una especie productora de fibra (*Capra hircus*) (Genoma de referencia: ARS1) y una de lana (*Ovis aries*) (Genoma de referencia: ARS-UI\_Ramb\_v2.0), para comprobar la existencia de otros genes de proteínas asociadas a las queratinas en alpacas y que estos no se encuentren anotados en el NCBI. Esto se realizó utilizando las herramientas BLAST y Clustal Omega, mediante un alineamiento entre las secuencias de las otras especies y el genoma de referencia de la alpaca (VicPac3.1) (Figura 3). Además, para sustentar y comprobar la similitud existente entre genes de proteínas asociadas a las queratinas (*KRTAP*) en alpacas y las especies mencionadas, se realizó un análisis comparativo entre las secuencias de los *KRTAPs* anotados en el genoma de referencia de la alpaca (VicPac3.1) en el NCBI con los genomas de cabra (ARS1) y oveja (ARS-UI\_Ramb\_v2.0), esto se realizó con el mismo procedimiento, utilizando las herramientas

BLAST y Clustal Omega, mediante un alineamiento entre las secuencias de las otras especies y el genoma de referencia de la alpaca (VicPac3.1).

The image shows the NCBI BLAST search interface. The main input field contains the accession number 'NM\_001159760.1'. Below it, there is an option to 'Or, upload file' which is currently set to 'Ningún archivo seleccionado'. The 'Job Title' field is filled with 'NM\_001159760:Ovis aries keratin associated...'. The 'Database' dropdown menu is set to 'VicPac3.1 RefSeq assembly [GCF\_000164845.3]'. Under the 'Program Selection' section, the 'Optimize for' options are: 'Highly similar sequences (megablast)' (selected), 'More dissimilar sequences (discontiguous megablast)', and 'Somewhat similar sequences (blastn)'. At the bottom, the 'BLAST' button is prominent, and the search parameters are summarized as 'Search database genomic/30538/GCF\_000164845.3 using Megablast (Optimize for highly similar sequences)'. There is also a checkbox for 'Show results in a new window'.

**Figura 3:** Configuración de entrada para realizar la comparación del genoma de la alpaca con el gen *KRTAPI-1* de oveja, haciendo uso de la herramienta BLAST.  
FUENTE: NCBI, 2021

### 3.3. LOCALIZACIÓN DE GENES DE LAS PROTEÍNAS ASOCIADAS A LAS QUERATINAS EN EL VICPAC 3.1

Para localizar las secuencias de nucleótidos de los genes de proteínas asociadas a las queratinas en el VicPac 3.1, se necesitaba obtener la posición inicial y final de cada exón, así como, de los intrones y/o regiones no traducidas (UTRs). Para ello, usando la base de datos del NCBI se colocó marcadores en la gráfica de cada gen y en cada marcador aparece un número que indica la posición (Figura 4).

### 3.4. IDENTIFICACIÓN DE POLIMORFISMOS DE NUCLEÓTIPO SIMPLE

Se utilizaron los genes que ya se encontraban anotados en el NCBI para una mayor precisión del estudio y habiendo obtenido las posiciones de cada exón, se generó un archivo de coordenadas (bed file) de los exones, intrones y UTRs de cada gen. Se identificó los polimorfismos de nucleótido simple (PNSs) localizados en dichas regiones, utilizando la herramienta vcftools y la base de datos (archivos vcf) de los PNSs identificados a partir de diferentes fuentes de información:

- Alineamiento entre datos de secuenciamiento de 150 bibliotecas de representación reducida obtenidas a partir de la digestión con enzima ApeKI y genoma de referencia VicPac3.1.
- Alineamiento entre datos de secuenciamiento de 150 bibliotecas de representación reducida obtenidas a partir de la digestión doble con las enzimas MspI-PstI y genoma de referencia VicPac3.1.
- Alineamiento entre datos de secuenciamiento de 6 genomas completos UNALM (Número de accesoión: PRJEB42148) y genoma de referencia VicPac3.1.
- Alineamiento entre datos de secuenciamiento de 3 genomas completos UPCH (Número de accesoión NCBI: PRJNA340289) y genoma de referencia VicPac3.1.

### **3.5. SELECCIÓN DE PNSs**

Para que los PNSs sean confiables y elevar la probabilidad de éxito al incorporar dichos PNSs en una micromatriz, se consideró, la frecuencia del alelo menor (MAF) y la tasa de genotipificación, porque los PNSs raros o poco comunes no se pueden detectar con una frecuencia razonable en la población. Además, para evitar errores en el diseño de cebadores y sondas para la genotipificación, se tomó en cuenta el Illumina Score y que no existan otros PNSs en las secuencias flanqueantes. Para seleccionar un PNS estos tuvieron que pasar por cuatro filtros:

- a. Que no existan otros PNSs en las secuencias flanqueantes del PNS (35 pb).
- b. Que la frecuencia del alelo menor (MAF) sea mayor o igual al 0.05.

Se creó un archivo vcf con la información de genotipos usando el software KGD (<https://github.com/AgResearch/KGD>), y se calculó el MAF para cada PNS usando el software PLINK.

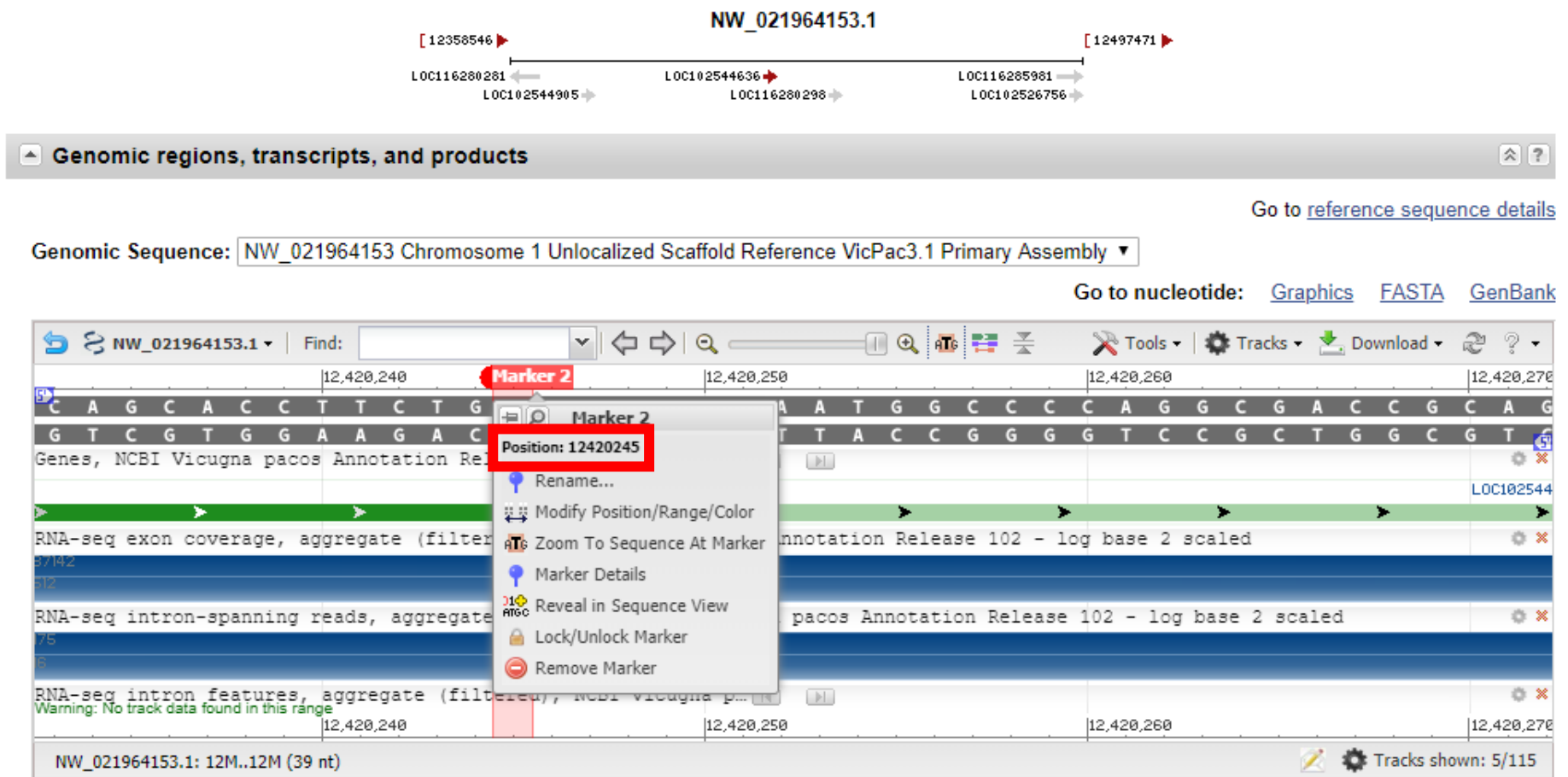
- c. Que la tasa de genotipificación sea mayor al 45 %.

Se creó un archivo vcf con la información de genotipos usando el software KGD (<https://github.com/AgResearch/KGD>), y se calculó el parámetro usando el software PLINK.

- d. Que el Illumina Score sea mayor o igual a 0.6.

Se utilizó el Illumina Design Studio (<https://designstudio-array.illumina.com>) para calcular el Illumina Score para cada PNS





**Figura 4:** Ubicación del exón  
 Ubicación de la posición final del exón del gen *VICPA-KRTAP7-1*, colocando un marcador en el NCBI.  
**FUENTE:** NCBI, 2021

Una vez identificados los PNSs, se creó una tabla de datos donde se registró cada PNS con su secuencia total de 71 nucleótidos, ubicándose el polimorfismo en el medio.

Se realizó un análisis de los PNSs de los genes *KRTAP* incluidos en el chip (Calderon et al., 2021), haciendo uso de la herramienta Genome Studio ([https://support.illumina.com/array/array\\_software/genomestudio/downloads.html](https://support.illumina.com/array/array_software/genomestudio/downloads.html)) para analizar el MAF en una población de 150 alpacas.

## IV. RESULTADOS Y DISCUSIONES

### 4.1. GENES DE PROTEÍNAS ASOCIADAS A LAS QUERATINAS EN ALPACAS

Se obtuvo la localización y secuencias de 34 genes *KRTAP* anotados en el genoma de referencia de alpacas de los cuales según la clasificación de Gong et al. (2012), 19 son genes con alto contenido de sulfuro, 6 son genes con muy alto contenido de sulfuro y 9 son genes con alto contenido en glicina y tirosina (Tabla 3).

**Tabla 3:** Lista de *KRTAPs* en alpacas

Grupo de <i>KRTAPs</i>	Identificación del Locus	Gen	# de exones
Alto contenido de sulfuro	LOC102528693	<i>KRTAP2-4L</i>	1
	LOC102528725	<i>KRTAP3-1</i>	1
	LOC102528468	<i>KRTAP3-3L*</i>	1
	LOC102528209	<i>KRTAP3-3L*</i>	2
	LOC107034113	<i>KRTAP10-1L</i>	3
	LOC116284217	<i>KRTAP10-10</i>	5
	LOC102544905	<i>KRTAP11-1</i>	1
	LOC102528691	<i>KRTAP12-1L*</i>	1
	LOC116283958	<i>KRTAP12-1L*</i>	2
	LOC116280071	<i>KRTAP12-2L*</i>	1
	LOC102530622	<i>KRTAP12-2L*</i>	1
	LOC102533350	<i>KRTAP13-1L*</i>	1
	LOC102532866	<i>KRTAP13-1L*</i>	1
	LOC102532037	<i>KRTAP13-1L*</i>	1
	LOC102544939	<i>KRTAP16-1</i>	2
	LOC102532787	<i>KRTAP24-1</i>	1
	LOC107034115	<i>KRTAP26-1</i>	1
	LOC102533595	<i>KRTAP27-1</i>	1
	LOC102544666	<i>KRTAP29-1</i>	3
Muy alto contenido de sulfuro	LOC102543867	<i>KRTAP4-1</i>	2
	LOC102529268	<i>KRTAP4-3L</i>	2
	LOC107032861	<i>KRTAP4-7L*</i>	1

Continuación ...

Grupo de <i>KRTAPs</i>	Identificación del Locus	Gen	# de exones
<b>Muy alto contenido de sulfuro</b>	LOC102529530	<i>KRTAP4-7L</i> *	12
	LOC102544126	<i>KRTAP9-2L</i>	1
	LOC102534338	<i>KRTAP9-4L</i>	2
<b>Alto contenido en glicina y tirosina</b>	LOC102544636	<i>KRTAP7-1</i>	1
	LOC116280298	<i>KRTAP8-1</i>	1
	LOC116276605	<i>KRTAP19-1L</i>	1
	LOC116280314	<i>KRTAP19-3L</i>	1
	LOC116286138	<i>KRTAP20-2L</i> *	1
	LOC116286182	<i>KRTAP20-2L</i> *	1
	LOC116276588	<i>KRTAP20-2L</i> *	1
	LOC116276577	<i>KRTAP20-2L</i> *	1
	LOC116286002	<i>KRTAP21-1L</i>	1

\*Genes que cuentan con más de una copia.

**FUENTE:** Información tomada de NCBI (<https://www.ncbi.nlm.nih.gov/>) y clasificación adaptada de Gong *et al.* (2012) y Khan *et al.* (2014).

En la Tabla 4 se presenta el listado de genes *KRTAP* de alpacas, cabras y ovejas equivalentes en nomenclatura anotada en el NCBI.

**Tabla 4:** Listado de genes *KRTAP* de alpacas, cabras y ovejas equivalentes en nomenclatura.

Genes en oveja	Genes en cabra	Genes en alpaca
<i>KRTAP1-1</i>		
<i>KRTAP1-3</i>	<i>KRTAP1-3</i>	
	<i>KRTAP1-4</i>	
		<i>KRTAP2-4L</i>
	<i>KRTAP3-1</i>	<i>KRTAP3-1</i>
<i>KRTAP3-2</i>		
<i>KRTAP3-3L</i>	<i>KRTAP3-3</i>	<i>KRTAP3-3L</i>
		<i>KRTAP4-1L</i>
<i>KRTAP4-3</i>		<i>KRTAP4-3L</i>
<i>KRTAP4-7L</i>		<i>KRTAP4-7L</i>
<i>KRTAP4-8L</i>	<i>KRTAP4-8L</i>	
<i>KRTAP4-9L</i>	<i>KRTAP4-9</i>	
<i>KRTAP4-11p</i>	<i>KRTAP4-11L</i>	
<i>KRTAP4-12L</i>		
<i>KRTAP5-4</i>		
<i>KRTAP6-1L</i>	<i>KRTAP6-1</i>	
<i>KRTAP7-1L</i>	<i>KRTAP7-1</i>	<i>KRTAP7-1</i>
<i>KRTAP8-1</i>	<i>KRTAP8-1</i>	<i>KRTAP8-1</i>

Continuación ...

Genes en oveja	Genes en cabra	Genes en alpaca
<i>KRTAP9-2</i>	<i>KRTAP9-2</i>	<i>KRTAP9-2L</i>
	<i>KRTAP9-3L</i>	
		<i>KRTAP9-4L</i>
<i>KRTAP9-7L</i>		
	<i>KRTAP 9-8L</i>	
<i>KRTAP9-9L</i>	<i>KRTAP 9-9L</i>	
	<i>KRTAP 10-1</i>	<i>KRTAP10-1L</i>
<i>KRTAP10-7L</i>		
<i>KRTAP10-8</i>	<i>KRTAP 10-8</i>	<i>KRTAP10-8L</i>
	<i>KRTAP 10-9L</i>	
		<i>KRTAP10-10</i>
<i>KRTAP10-11L</i>	<i>KRTAP 10-11L</i>	
	<i>KRTAP 11-1</i>	<i>KRTAP11-1</i>
<i>KRTAP12-1L</i>	<i>KRTAP 12.1</i>	<i>KRTAP12-1L</i>
<i>KRTAP12-2</i>	<i>KRTAP 12-2</i>	<i>KRTAP12-2L</i>
<i>KRTAP13-1</i>	<i>KRTAP 13-1</i>	<i>KRTAP13-1L</i>
<i>KRTAP13-2L</i>		
<i>KRTAP13-4L</i>		
<i>KRTAP16-1</i>	<i>KRTAP 16-1</i>	<i>KRTAP16-1</i>
	<i>KRTAP 16-4</i>	
<i>KRTAP17-1</i>	<i>KRTAP 17-1</i>	
		<i>KRTAP19-1L</i>
<i>KRTAP19-3L</i>		<i>KRTAP19-3L</i>
<i>KRTAP20-2L</i>		<i>KRTAP20-2L</i>
<i>KRTAP21-1L</i>		<i>KRTAP21-1L</i>
<i>KRTAP24-1</i>	<i>KRTAP24-1</i>	<i>KRTAP24-1</i>
<i>KRTAP26-1</i>	<i>KRTAP26-1</i>	<i>KRTAP26-1</i>
<i>KRTAP27-1</i>	<i>KRTAP27-1</i>	<i>KRTAP27-1</i>
<i>KRTAP29-1</i>	<i>KRTAP29-1</i>	<i>KRTAP29-1</i>

En la Tabla 5 se presentan los resultados del análisis comparativo entre las secuencias de los genes *KRTAP* de oveja que no tienen un equivalente en nomenclatura en alpaca, con el genoma de referencia (VicPac 3.1). Los porcentajes de identidad sombreados en amarillo son menores a 70%, para fines de la investigación estos genes no serán considerados como genes equivalentes a los genes de alpaca por su bajo porcentaje de identidad.

**Tabla 5:** Comparación entre las secuencias de genes *KRTAP* de oveja que no tienen un equivalente en nomenclatura en alpaca.

OVEJA		ALPACA			
Gen	Cromosoma	Cromosoma	Scaffold	Posición en el scaffold	% de identidad
<i>KRTAP1-1</i>	OAR11	VPA16	ABRR 03077387.1	500435- 500917	90.68%
<i>KRTAP3-2</i>	OAR11	VPA16	ABRR 03077387.1	514734- 515030	85.86%
<i>KRTAP4-12L</i>	OAR11	VPA16	ABRR 03077387.1	411149- 411999	71.01%
<i>KRTAP5-4</i>	OAR21	na	ABRR 03077255.1	84630- 85800	74.11%
<i>KRTAP9-7L</i> (LOC114116843)	OAR11	VPA16	ABRR 03077387.1	379100- 379872	54.39%
<i>KRTAP9-7L</i> (LOC114116851)	OAR11	VPA16	ABRR 03077387.1	379376- 379867	84.87%
<i>KRTAP9-7L</i> (LOC114116849)	OAR11	VPA16	ABRR 03077387.1	379376- 380328	77.20%
<i>KRTAP10-7L</i> (LOC101104456)	OAR1	VPA1	ABRR 03000004.1	1577928- 1578885	75.79%
<i>KRTAP10-7L</i> (LOC114110483)	OAR1	VPA1	ABRR 03000004.1	1621761- 1622791	81.63%
<i>KRTAP13-2L</i>	OAR1	VPA1	ABRR 03000004.1	12905032- 12905861	80.75%
<i>KRTAP13-4L</i>	OAR1	VPA1	ABRR 03000004.1	12911768- 12924277	49.69%

na: no asignado.

OAR: Cromosoma de oveja (*Ovis aries*).

VPA: Cromosoma de alpaca (*Vicugna pacos*).

En la Tabla 6 se presentan los resultados del análisis comparativo entre las secuencias de los genes *KRTAP* de cabra que no tienen un equivalente en nomenclatura en alpacas, con el genoma de referencia (VicPac 3.1). Los porcentajes de identidad sombreados en amarillo son menores a 70%, para fines de la investigación estos genes no serán considerados como genes equivalentes a los genes de alpaca por su bajo porcentaje de identidad.

**Tabla 6:** Comparación entre las secuencias de genes *KRTAP* de cabra que no tienen un equivalente en nomenclatura en alpaca.

CABRA		ALPACA			
Gen	Cromosoma	Cromosoma	Scaffold	Posición en el scaffold	% de identidad
<i>KRTAP1-4</i>	CHI19	VPA16	ABRR 03077387.1	497592- 498462	86.90%
<i>KRTAP4-11L</i>	CHI19	VPA16	ABRR 03077387.1	424685- 425393	75.04%
<i>KRTAP9-3L</i>	CHI19	VPA16	ABRR 03077387.1	379442- 379867	75.53%
<i>KRTAP9-8L</i>	n/a	VPA16	ABRR 03077387.1	459928- 460936	75.73%
<i>KRTAP10-9L</i>	CHI1	VPA1	ABRR 03000004.1	1621743- 1622784	81.83%
<i>KRTAP16-4</i>	CHI1	VPA1	ABRR 03000004.1	12496847- 12497483	78.06%

n/a: no asignado.

CHI: Cromosoma de cabra (*Capra hircus*).

VPA: Cromosoma de alpaca (*Vicugna pacos*).

En la Tabla 7 se presentan los resultados del análisis comparativo entre las secuencias de los genes *KRTAP* de oveja y cabra que no tienen un equivalente en nomenclatura en alpacas, con el genoma de referencia (VicPac 3.1). Los porcentajes de identidad sombreados en amarillo son menores a 70%, para fines de la tesis estos genes no serán considerados como genes ortólogos a los genes de alpaca por su bajo porcentaje de identidad.

**Tabla 7:** Comparación entre las secuencias de genes *KRTAP* de oveja y cabra que no tienen un equivalente en nomenclatura en alpaca.

Gen	Especie	Cromosoma	Cromosoma en alpaca	Posición en el scaffold de alpaca	% de identidad
<i>KRTAP1-3</i>	OVEJA	OAR11	VPA16	ABRR03077387.1: 497652-498119	83.29%
	CABRA	CHI19			84.19%
<i>KRTAP4-8L</i>	OVEJA	OAR11	VPA16	ABRR03077387.1: 455442-455951	80.00%
	CABRA	CHI19			87.68%
<i>KRTAP4-9L</i>	OVEJA	OAR11	VPA16	ABRR03077387.1: 455442-455951	81.33%
	CABRA	CHI19			82.35%
<i>KRTAP6-1L</i>	OVEJA	OAR1	VPA1	ABRR03000004.1: 12655780-12655996	80.00%
	CABRA	CHI1			87.68%
<i>KRTAP9-9L</i>	OVEJA	OAR11	VPA16	ABRR03077387.1: 379361-379864	78.66%
	CABRA	CHI19			79.26%
<i>KRTAP10-11L</i>	OVEJA	OAR1	VPA1	ABRR03000004.1: 1621806-1622875	65.89%
	CABRA	CHI1			66.02%
<i>KRTAP17-1</i>	OVEJA	OAR11	VPA16	ABRR03077387.1: 420026-420349	53.77%
	CABRA	CHI9			53.70%

OAR: Cromosoma de oveja (*Ovis aries*).

CHI: Cromosoma de cabra (*Capra hircus*).

VPA: Cromosoma de alpaca (*Vicugna pacos*).

En la Tabla 8 se presentan los resultados del análisis comparativo entre las secuencias de los genes *KRTAP* de alpaca anotados en el NCBI (Tabla 3) y los genomas de referencia de oveja y cabra. Para sustentar y comprobar la similitud existente entre los genes de estas especies. Los porcentajes de identidad sombreados en amarillo son menores a 70%, para fines de la tesis estos genes no serán considerados como genes ortólogos a los genes de alpaca por su bajo porcentaje de identidad.



**Tabla 8:** Porcentaje de similitud de *KRTAPs* anotados en los genomas de referencia entre alpaca, cabra y oveja.

Gen en Alpaca	Cromosoma en Alpaca	Especie	Gen	Cromosoma	% de identidad	% de identidad entre cabra y oveja
<i>KRTAP2-4L</i>	n/a	CABRA	<i>Keratin IIIA3L</i>	19	86.75%	84.75%
		OVEJA	<i>Keratin IIIA3A</i>	11	85.54%	
<i>KRTAP3-1</i>	16	CABRA	<i>KRTAP3-1</i>	19	80.74%	96.88%
		OVEJA	<i>Keratin BIIIB2</i>	11	79.15%	
<i>KRTAP3-3L*</i> (LOC102528468)	16	CABRA	<i>Keratin IIIB3</i>	19	79.08%	96.26%
		OVEJA	<i>Keratin IIIB3</i>	11	80.29%	
<i>KRTAP3-3L*</i> (LOC102528209)	16	CABRA	<i>Keratin IIIB3</i>	19	68.64%	76.72%
		OVEJA	<i>Keratin BIIIB4</i>	11	78.96%	
<i>KRTAP10-1L</i>	1	CABRA	<i>KRTAP13-1* (LOC102181431)</i>	1	64.35%	79.75%
		OVEJA	<i>KRTAP13-1L* (LOC101115597)</i>	1	80.68%	
<i>KRTAP10-10</i>	18	CABRA	<i>Late cornified envelope like proline-rich protein 1</i>	25	67.14%	99.11%
		OVEJA	<i>KRTAP10-2</i>	24	88.25%	
<i>KRTAP11-1</i>	1	CABRA	<i>KRTAP11-1</i>	1	86.45%	98.90%
		OVEJA	<i>KRTAP11-1</i>	1	87.32%	
<i>KRTAP12-1L*</i> (LOC102528691)	n/a	CABRA	<i>KRTAP12-1</i>	1	75.77%	66.30%
		OVEJA	<i>KRTAP 12-2L</i>	1	70.24%	
<i>KRTAP12-1L*</i> (LOC116283958)	1	CABRA	<i>KRTAP 12-1</i>	1	67.55%	66.30%
		OVEJA	<i>KRTAP 12-2L</i>	1	63.55%	
<i>KRTAP12-2L*</i> (LOC116280071)	1	CABRA	<i>KRTAP 12-2* (LOC102189437)</i>	1	75.54%	86.54%
		OVEJA	<i>KRTAP 12-2L</i>	1	77.16%	
<i>KRTAP12-2L*</i> (LOC102530622)	n/a	CABRA	<i>KRTAP 12-2* (LOC102189437)</i>	1	73.55%	86.54%
		OVEJA	<i>KRTAP 12-2L</i>	1	76.66%	

Continuación ...

Gen en Alpaca	Cromosoma en Alpaca	Especie	Gen	Cromosoma	% de identidad	% de identidad entre cabra y oveja
<i>KRTAP13-1L</i> * (LOC102533350)	1	CABRA	<i>KRTAP 13-1</i> *( <i>LOC102181160</i> )	1	68.25%	98.15%
		OVEJA	<i>KRTAP13-3</i>	1	73.54%	
<i>KRTAP13-1L</i> * (LOC102532866)	1	CABRA	<i>KRTAP13-1</i> *( <i>LOC102181431</i> )	1	64.64%	72.98%
		OVEJA	<i>KRTAP13-1L</i> * ( <i>LOC101106046</i> )	1	79.43%	
<i>KRTAP13-1L</i> * (LOC102532037)	1	CABRA	<i>KRTAP13-1</i> * ( <i>LOC102181160</i> )	1	78.82%	98.11%
		OVEJA	<i>KRTAP13-1</i> * ( <i>LOC101115849</i> )	1	79.63%	
<i>KRTAP16-1</i>	1	CABRA	<i>KRTAP16-1</i>	19	76.78%	98.52%
		OVEJA	<i>KRTAP16-1</i>	11	76.61%	
<i>KRTAP24-1</i>	1	CABRA	<i>KRTAP24-1</i>	1	77.69%	96.66%
		OVEJA	<i>KRTAP24-1</i>	1	79.05%	
<i>KRTAP26-1</i>	1	CABRA	<i>KRTAP 24-1</i>	1	49.76%	96.66%
		OVEJA	<i>KRTAP24-1</i>	1	49.94%	
<i>KRTAP27-1</i>	1	CABRA	<i>KRTAP27-1</i>	1	79.03%	98.47%
		OVEJA	<i>KRTAP27-1</i>	1	73.76%	
<i>KRTAP29-1</i>	16	CABRA	<i>KRTAP29-1</i>	19	68.30%	97.56%
		OVEJA	<i>KRTAP29-1</i>	11	69.42%	
<i>KRTAP4-1</i>	16	CABRA	<i>KRTAP4-8L</i> *( <i>LOC102178767</i> )	19	48.59%	69.60%
		OVEJA	<i>KRTAP4-3</i>	11	54.52%	
<i>KRTAP4-3L</i>	16	CABRA	<i>Keratin B2D</i>	19	82.88%	97.61%
		OVEJA	<i>Keratin B2D</i>	11	83.20%	
<i>KRTAP4-7L</i> * (LOC107032861)	16	CABRA	<i>KRTAP 4-8L</i> * ( <i>LOC106503216</i> )	19	64.75%	93.02%
		OVEJA	<i>KRTAP 4-8L</i>	11	64.86%	
<i>KRTAP4-7L</i> * (LOC102529530)	16	CABRA	<i>KRTAP 4-9L</i>	19	73.29%	89.84%
		OVEJA	<i>KRTAP 4-8L</i>	11	73.96%	

Continuación ...

Gen en Alpaca	Cromosoma en Alpaca	Especie	Gen	Cromosoma	% de identidad	% de identidad entre cabra y oveja
<i>KRTAP9-2L</i>	16	CABRA	<i>KRTAP 9-8L</i>	ASM170441v1	74.41%	96.94%
		OVEJA	<i>KRTAP 9-2L</i>	11	71.49%	
<i>KRTAP9-4L</i>	n/a	CABRA	<i>KRTAP 9-3L</i>	19	75.76%	75.96%
		OVEJA	<i>KRTAP 9-2L</i>	11	70.18%	
<i>KRTAP7-1</i>	1	CABRA	<i>KRTAP 7-1</i>	1	78.08%	97.97%
		OVEJA	<i>KRTAP7 (KAP7)</i>	1	78.09%	
<i>KRTAP8-1</i>	1	CABRA	<i>KRTAP8-1 (KAP8)</i>	1	76.10%	97.30%
		OVEJA	<i>KRTAP8 (KAP8)</i>	1	76.10%	
<i>KRTAP19-1L</i>	1	CABRA	<i>KRTAP16-3</i>	1	66.32%	74.20%
		OVEJA	<i>KRTAP19-3L</i>	1	51.48%	
<i>KRTAP19-3L</i>	1	CABRA	<i>KRTAP16-3</i>	1	58.33%	74.20%
		OVEJA	<i>KRTAP19-3L</i>	1	50.55%	
<i>KRTAP20-2L*</i> (LOC116286138)	1	CABRA	<i>KRTAP20-1</i>	1	63.91%	98.96%
		OVEJA	<i>KRTAP20-2</i>	1	68.48%	
<i>KRTAP20-2L*</i> (LOC116286182)	1	CABRA	<i>KRTAP22-2</i>	1	75.66%	77.89%
		OVEJA	<i>KRTAP22-1 (KAP22-1)</i>	1	74.88%	
<i>KRTAP20-2L*</i> (LOC116276588)	1	CABRA	<i>KRTAP22-2</i>	1	76.95%	77.89%
		OVEJA	<i>KRTAP22-1 (KAP22-1)</i>	1	77.33%	
<i>KRTAP20-2L*</i> (LOC116276577)	1	CABRA	<i>KRTAP22-2</i>	1	76.40%	77.89%
		OVEJA	<i>KRTAP22-1 (KAP22-1)</i>	1	77.70%	
<i>KRTAP21-1L</i>	1	CABRA	<i>KRTAP6-2 (KAP6-2)</i>	1	72.43%	86.43%
		OVEJA	<i>KRTAP21-1L</i>	1	75.78%	

\*Genes que cuentan con más de una copia.

**Tabla 9:** Ubicación de genes, polimorfismos de nucleótido simple (PNSs) y secuencias flanqueantes a los PNSs (35pb) en el VicPac3.1 de *KRTAPs* que codifican proteínas con alto contenido de sulfuro.

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP3-1</i>	ABRR03077387.1	16	Exón único	1	ATCAGTCTCCTTCAGCCTACCTGCTG CGACTGCTG[T/C]CCCCACCCTGCT GTGAGCCTGACACCTACGTGCC	514886
<i>KRTAP3-3L</i> (LOC102528468)	ABRR03077387.1	16	Exón único	2	GATAAGTCCTGCCGGTGTGGAGTCT GCCTGCCAG[T/C]ACCTGCCACAC ACGGTTTGGTTACTGGAGCCAAC	526383
					TCACAACCTACActcagccctgctctgagccctg C[G/A]TCCCAAGATGCTGCTGACCGA TGACTGCTTCGCTC	526561
<i>KRTAP3-3L</i> (LOC102528209)	ABRR03077387.1	16	Exón 2	2	ATTCCCAAGGCCATGGCTTGCTGTG TTTCCTGCGA[C/T]GGCTGCAGTGTT CGCACCGGCCCGCCACCACCAT	532094
					GCCTGCCAGCACCTGCCACACAC GGTTTGGTTA[C/T]TGGAGCCAACCT GCTGTGACAActgccccacc	532200
<i>KRTAP10-1L</i>	ABRR03000004.1	1	Exón 1	10	CTTTGATCTCAATACTCAGAAGGAT CTAGACCCTC[A/G]GCAAGTTGAATA ATAGAAACCAGAGCTAGAGATTG	12853114
			Exón 3		ATGACTGGCAGCTGCTGGAGGTAAA GCAGGTTGGA[C/G]GGCAGAATCTG GATCCACAGCTCAGGGAGGGGAAG	12853197
					ATCAATAGAAGCCAGACTGACAGGT TGGTCTGTAA[C/T]AGGATGATTGGC AGCTCCTAGGGGCAAAGTAGGTT	12888037

Continuación ...

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP10-1L</i>	ABRR03000004.1	1	Exón 3	10	TGATTGGCAGCTCCTAGGGGCAAAG TAGGTTGGAC[G/A]GTAGAAACCGA ATCCATGGCTCAGAGGAGGGAAGC	12888077
					AAACCGAATCCATGGCTCAGAGGAG GGAAGCCACA[G/T]ACTCCACAGCC CAGGGGTCTGAAGCCATGGGACCC	12888117
					TCTGAAGCCATGGGACCCAAAGCCC AGTGAGTAGC[A/G]GCTTCTGGACCC AGCGCCCAGGAAGCAGCCTCTTC	12888170
					TCTTCCGGACCCAGAGCCTGGAGAC CTAGCATAAG[T/C]TTTCTGGAAGGG ACTGCAGAGCATGGAGGTCCTTG	12888236
					CTGGAAGGGACTGCAGAGCATGGA GGTCCTTGGGC[G/A]GTAGCAGGAT GTCTGGCAGGGGCTGGATAACCAT	12888275
					CAGGATGTCTGGCAGGGGCTGGATA CCACATAAGA[C/T]GTCTGGCACCTG ATAGGCTTACAACAGGTCTCCTG	12888315
					CACCTGATAGGCTTACAACAGGTCT CCTGACAGCC[G/A]CTGTGAAGAGA GGAGTCCTGctggcaggtgctggg	12888357
					CAAGCTCCGATGATGACCCAGTGTC CCTGCCCTT[G/T]GGGGAGCGGCCA CCCCCGGGGCCACATGCCCTG	8124973
					ACATGTCCTACAACCTGCTCCACAAG GAATTGCTCT[T/G]CCAGGCTGATTG GGGACAATACTCTGTCCCCGTG	12376109

Continuación ...

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP12-1L</i> (LOC102528691)	ABRR03009911.1	na	Exón único	1	CACACACACAGGGGCTTGCAGCTCA CGGGCAGGCG[C/T]GTGGTCAGCTTG CAGCTCACAGGCACGCACACGAC	764
<i>KRTAP12-1L</i> (LOC116283958)	ABRR03000004.1	1	Exón 1	1	GGAGCCCAACCCTGCGGAAGAACCT CCTGGATCAC[C/T]GAGTTTCCGCAT CAAAGTTGGGGACGCTGTGCTTG	3777628
<i>KRTAP12-2L</i> (LOC116280071)	ABRR03000004.1	1	UTR	1	GGGAACCAGGGCTTCCAGAGGAGA TGCGCAGATT[C/A/G]TAGAAAGATC CAGATCGTCCTGCCACGCTCTTGCC	3775478
					attaaaaaaaaatggaggacACACTCCATTGA GAA[A/T]GCATGATTTATTTGTCACA CATGGGATGTGGCTTC	807
<i>KRTAP12-2L</i> (LOC102530622)	ABRR03001991.1	na	UTR	3	GGGAACCAGGGCTTCCAGAGGAGA TGCGCAGATT[C/A/G]TAGAAAGATC CAGATCGTCCTGCCACGCTCTTGCC	902
					GAGGGCGGTGATGTCTGCAGCCTCC CTGCCCCGGG[C/G/T]GCCTTTATACCC GGGCCGTGGGCGTCCCAGCAGCA	1519
					TGCTGCTCTGGAaacttctctctctctctccttgg [G/A]GGCCACTTGCACTACCCAAGCT CTTCTGTGGCTC	12923762
<i>KRTAP13-1L</i> (LOC102533350)	ABRR03000004.1	1	Exón único	4	TGTGGCTCCTCCTACCCAGCAACC TGATCTACAC[T/C]ACGGAacctgctctcc cagcactgccagCTGAG	12923825
					GTACAGTGGCTGTCAGGAGACCTGC TGGGAGCCCA[C/G]CAGGTGCCAGA TGTCCCATGGTGTGTCCAGCCCCT	12923905

Continuación ...

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP13-1L</i> (LOC102533350)	ABRR03000004.1	1	Exón único	1	CTCTGGGCTCTGGGTCCGGCAAAGG CTACTTCCTG[G/T]TCTATGGATCCA GAAGCTGTGGGTCCAGTGGATCT	12924045
<i>KRTAP13-1L</i> (LOC102532866)	ABRR03000004.1	1	Exón único	1	TACCGCTGGAAGACCTCCACGCTCT GCCGTCCCTG[G/C]CAGACGACTTAC TCTGGGTCTCTGGGCTTTGGCTC	12846534
<i>KRTAP13-1L</i> (LOC102532037)	ABRR03000004.1	1	Exón único	2	GTCAGGAGACCAGCTGTGAGCCTAT CAGATGCCAG[G/T]CATCCTGCTGCC CCCAGAGGACCTCCAAGCTCTGC	12859532
					TGGGTCCAATTTCTGGTACCCAATC AACTTTCCTT[C/G]CAGAAGTTTCTG TTCATCTTGTTACTGGCCAATTT	12859761
<i>KRTAP16-1</i>	ABRR03077387.1	16	Exón 1	4	AGGGCCTGTTTGCCTGCCAGTTCCT GCCGGAGCC[G/A]GACTTGGCAACT GGTGACTCAAGATAGCTGCAGAT	353702
			Exón 2		AGTGTGCCCCACGCCTAGCTGCTGT TCATCTGTCC[G/A]CTCCGTGGCCAA TGGCCGCCAGTCTGTCTGCTGTG	354957
			Exón 2		GCTGCCGCCAGGGTCTTCTGCATC TGTCATCTGC[C/A]GGCCAATTTGCT CTCGAACTTTCTACATACCCAGC	355358
			UTR		AGAAAATGTGACTGCCCTCCCCAAA GCTCATCTGA[C/T]ttagaatcttttcttctgcac tATCACTACCA	355869

Continuación ...

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP24-1</i>	ABRR03000004.1	1	Exón único	4	ATATGCTACAGAACCCACTGTATTA TCCCAGTGAC[T/G]CCTTCTGTTGCT CTTTGCTCCAGCGATGTAAGCCC	12979560
					TACCAAGGAACTCTCTGGCTTCTGG ATAACTGCCA[A/T]GAAACCTGTGGT GAAGCACCAATCTGTGAATCTCC	12979659
					AGTGGGCAAAATACGCAGTGCCTGT GAAACTACCA[A/G]TGTCGGACCCA GCCCCAGCTGCAACCCATGCACTC	12979802
					CATACCAAATGGCTTCTCACCCCTCA TCTTGTATTG[C/T]CAACAGCTGCCG ATTCCaaaattatttaagaagaa	12980054
<i>KRTAP26-1</i>	ABRR03000004.1	1	Exón único	4	GTCatatttctctcctcctccattGCCCTCTGC[T/C] CTACCAATGTGAGCTGTGGAGAT GTCCTCTGCTTG	12947455
					GAGACTTCCAGCTGCCCTTCCACTG CTTGCTATGT[G/T]CCCAGACCCTGC CAAGGGACCAACCTTCTTCCTGC	12947628
					ctctggetcctgcctcccagtTCCTGCAGACCT C[T/A]GAGCTATGCGTCCAGCAGCTG CCGACCCCTGAGCC	12947708
					GATGCCAACCTCTGACACAAGTGTT CAGTCCTTGT[C/T]GTCCATCCTGCT CTGCATTGGGAGGCCAGTAGCTT	12947881



Continuación ...

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP27-1</i>	ABRR03000004.1	1	Exón único	3	ggatTTTTTaccCAGCAGCTGCCACAGTA GGAC[C/T]TGGCTCCTGGACAACTTT CAAGAAACCTGTTGTGA	12931105
					TGTCCACAGGGGATAGCTGTGTGCA AACTGCCTGC[C/G]TCCCCGAGTTG TCCAAACGAATTGTTCTAATTCC	12931214
					ACTTATGAGCCAACTTGCTGTGTTA CTGGTGGTTT[G/A]CAGTTGCCTAGT GAATgaagaatgaaaaatgtg	12931576
<i>KRTAP29-1</i>	ABRR03077387.1	16	Exón 2	3	GTCTGCTACTTTTTTAAGTCTTGTC ATCAGTTCC[C/T]TGCGTGCCTGTGC CCCACCAGCTGTGTCcttgtgt	361421
					GTAAAGAACCCTTGCAAACCAGCTT CCTGCAGCAC[C/T]GTCCCTTCTGGC CAACCAACTTGTGGTGAACCTAC	361604
					TCCAGTGGAGGTTCCAATCGCTTCC GAACCACTGC[T/C]CCAAGTCTGCCA GGCCAGCACCTGCTAGCCAACCTC	361748

na: no asignado.

## **4.2. POLIMORFISMOS DE NUCLEÓTIDOS SIMPLES (PNSs) EN *KRTAPs* EN ALPACAS**

### **4.2.1. GENES QUE CODIFICAN PROTEÍNAS CON ALTO CONTENIDO DE SULFURO**

Se ubicó un total de 48 PNSs en los genes *KRTAP* que codifican proteínas con alto contenido de sulfuro. La ubicación del gen (scaffold y cromosoma), la región en la que se encuentra el PNS, el número de PNS por gen, la secuencia flanqueante del PNS (35pb) y su posición dentro del scaffold, se encuentran disponibles en la Tabla 9 y los valores de los parámetros considerados para cada PNS se encuentran disponibles en el Anexo 1. No se encontró PNSs en el gen *KRTAP2-4L*.

### **4.2.2. GENES QUE CODIFICAN PROTEÍNAS CON MUY ALTO CONTENIDO DE SULFURO**

Se ubicó un total de ocho PNSs en los genes *KRTAP* que codifican proteínas con muy alto contenido de sulfuro. La ubicación del gen (scaffold y cromosoma), la región en la que se encuentra el PNS, el número de PNS por gen, la secuencia flanqueante del PNS (35pb) y su posición dentro del scaffold, se encuentran disponibles en la Tabla 10 y los valores de los parámetros considerados para cada PNS se encuentran disponibles en el Anexo 2. No se encontró ningún PNS en la copia del gen *KRTAP4-7L* (LOC102529530).

### **4.2.3. GENES QUE CODIFICAN PROTEÍNAS CON ALTO CONTENIDO EN GLICINA Y TIROSINA**

Se logró ubicar once PNSs en los genes *KRTAP* que codifican proteínas con alto contenido en glicina y tirosina. La ubicación del gen (scaffold y cromosoma), la región en la que se encuentra el PNS, el número de PNS por gen, la secuencia flanqueante del PNS (35pb) y su posición dentro del scaffold, se encuentran disponibles en la Tabla 11 y los valores de los parámetros considerados para cada PNS se encuentran disponibles en el Anexo 3.

De los 67 PNSs identificados, 35 fueron incluidos en un chip de alpaca (Calderon et al., 2021), que se analizó en una población de 150 alpacas (Anexo 1, 2 y 3), en la cual se obtuvo que tres PNSs fueron monomórficos (MAF=0), ocho presentaron MAF entre 0.007 y 0.044, y 24 presentaron  $MAF \geq 0.05$ . Lo que indica que un 91.4% de los PNSs incluidos en dicha matriz son polimórficos para esa población, por lo que se consideran PNSs verdaderos.

**Tabla 10:** Ubicación de genes, polimorfismos de nucleótido simple (PNSs) y secuencias flanqueantes a los PNSs (35pb) en el VicPac3.1 de *KRTAPs* que codifican proteínas con muy alto contenido de sulfuro.

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP4-1</i>	ABRR03077387.1	16	Intrón	3	TGTCGCCCCAGGTGCTGCAACTCCA GCTGCTGCCG[C/G]CCCATCTGCTGA GGATCCTCTTTTGGCTGAACTTC	400197
					gaCCACCAAGTTTATTATCCTGTGAC ACTACCAAG[G/C]CACTGGAGAAAG AGCCTCCATCTGCGTTCAGTACC	403985
			UTR			GCTGTTACACTTGCCTTGACCTTCAT AACTGCATG[T/C]TAACTATTTCTCA ATAATACTATCTTGATATCATA
<i>KRTAP4-3L</i>	ABRR03077387.1	16	Exón 2	1	CAGCTGTGAGACTGGCTGTGGCATT GGTGGTAGCA[T/C]TGGCTGTGGCCA GGAGGGTGGCAGCGGAGCTCTGA	500673
<i>KRTAP4-7L</i> (LOC107032861)	ABRR03077387.1	16	Exón único	1	CTGCTGCAGGCCACCTGCTGCATC TCTAGCTGCT[G/A]TCGTCCCAGCTG CTACCAGACCACGTGCTGCCGCC	411461
<i>KRTAP9-2L</i>	ABRR03077387.1	16	UTR	1	GTCCAGCTGCTGCCAGCCCTCCTGC TGATCACCTC[A/T]CCAAGAGCCATC CCCTGCATCCAACAATCTGTC	379872
<i>KRTAP 9-4L</i>	ABRR03001291.1	na	Exón 1	2	GGTAGGCTGGCAGCAGGTTTGTCCA CAGCTGCTGG[A/G]CTCACAGCAGG TGGGCTGGCAGCAGGTGGTCACAC	3527
					TGGGCTGGCAGCAGGTGGTCACACA AGTAGGTTTG[C/T]GGCAGGTGGTGC TGCAGCAGGTGGTCTGAGTAGCT	3574

na: no asignado.

**Tabla 11:** Ubicación de genes, polimorfismos de nucleótido simple (PNSs) y secuencias flanqueantes a los PNSs (35pb) el VicPac3.1 de *KRTAPs* que codifican proteínas con alto contenido en glicina y tirosina.

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP7-1</i>	ABRR03000004.1	1	Exón único	1	TATGGCGGCAGCTTCTGCAGGCCATG GGGCTCCAG[C/G]TCTGGCTTTGGCTA CAGCACCTTCTGATGGACCAA	12420218
<i>KRTAP8-1</i>	ABRR03000004.1	1	Exón único	2	AGCACCTACTCCCCAGTGGGCTATGG CTTCGGCTA[T/C]GGCTACAACGGCT GTGGGGCTTTCGGCTACCGAAG	12436062
					CAACGGCTGTGGGGCTTTCGGCTACC GAAGATCCT[G/A]CCCATTCTTTTCT AGTGATTTGCTGAAATCCCCA	12436103
<i>KRTAP19-1L</i>	ABRR03000004.1	1	UTR	1	TAATTCTCTGCTTCATGTCTCTTATTT GTGCTTTC[A/G]TAACTGTGGCATTTC CCTATTCTTTTGCAATAAAT	12804959
<i>KRTAP19-3L</i>	ABRR03000004.1	1	Exón único	1	CTGGGCTATGGCTACGGCGCTGGCTA TGGTGGCTA[T/C]GGATACGGCTGCT ACCGCCCGTGTTACTACGGAAG	12746893
<i>KRTAP20-2L</i> (LOC116286138)	ABRR03000004.1	1	Exón único	1	ACCATAGCCACAGCCATAACCATAG CCCAGGCCAC[C/T]GTAACCATAACC CAGACCTCCATAGTAGTTGCCGT	12616128
<i>KRTAP20-2L</i> (LOC116286182)	ABRR03000004.1	1	Exón único	1	GCAGCATGGGCGGTAGCAGGCATAG CCATAGCCAC[C/T]GTAGCCACATCC ATAGCCACATCCCAGGCCACCAT	12651734
<i>KRTAP20-2L</i> (LOC116276588)	ABRR03000004.1	1	Exón único	1	GCAGCATGGGCGGTAGCAGGCATAG CCATAGCCAC[C/T]GTAGCCACATCC ATAGCCACATCCCAGGCCACCAT	12660986

Continuación ...

Gen	Scaffold	Cromosoma	Región	# PNS	Secuencia de 71 pb	Posición del PNS en el scaffold
<i>KRTAP20-2L</i> (LOC116276577)	ABRR03000004.1	1	Exón único	1	ATGGGCGGTAGCAGGCATAGCCATA GCCACCGTAG[C/T]CACATCCATAGC CACATCCCAGGCCACCATAGCCA	12678801
<i>KRTAP21-1L</i>	ABRR03000004.1	1	UTR	2	AACCTCCTCTTTCAACCCACCTTAAT TCACATCCT[T/C]TGACAACATGTGTT GCAACTACTATGGCAACTCCT GAAATTACATGCCTGAAAGAAGGTT TGACCTCCAA[T/C]TACCCGGTTTTAG ATTGGCATCTTTGTAGCCTGAG	12501902 12502328

### 4.3. DISCUSIÓN

De los 34 genes anotados en la base de datos del National Center for Biotechnology Information (NCBI) (Tabla 3), 6 genes (*KRTAP3-3L*, *KRTAP4-7L*, *KRTAP12-1L*, *KRTAP12-2L*, *KRTAP13-1L*, *KRTAP20-2L*) poseen una o más copias a lo largo del genoma. Según el genoma de referencia de la alpaca (VicPac3.1), 19 de estos genes se ubican en el cromosoma 1, 10 en el cromosoma 16, 1 en el cromosoma 18, el resto de genes aún no han sido ubicados en cromosomas. Dentro de los 19 genes ubicados en el cromosoma 1 se encuentran ubicadas una de las copias del gen *KRTAP12-1L* (LOC116283958) y del gen *KRTAP12-2L* (LOC116280071), las otras copias de cada gen LOC102528691 y LOC102530622 respectivamente, no han sido asignadas a cromosomas. La diferencia entre las copias de los 6 genes mencionados anteriormente, es principalmente el tamaño de la secuencia (pb) y el porcentaje de identidad entre las copias, que va entre 70 y 90%.

El análisis comparativo de las secuencias de los genes *KRTAP* entre oveja y alpaca, nos permite reportar que los genes:

- a. *KRTAP1-1*, *KRTAP3-2* y *KRTAP4-12L* (Anexo 4), que se encuentran ubicadas en el cromosoma 16 de alpaca (VPA16), estas mismas secuencias se encuentran localizadas en el cromosoma 11 de oveja (OAR11) siendo esto consistente ya que VPA16 tiene una región homóloga con el cromosoma humano HSA17q (Richardson *et al.*, 2019), que a su vez es homóloga a BTA19 (Zimin *et al.*, 2009) y OAR11 (Goldammer *et al.*, 2009). Al analizar estas secuencias en el VicPac3.1, el gen SHEEP-*KRTAP1-1* es equivalente al segundo exón del gen anotado como VICPA-*KRTAP4-3*; el gen SHEEP-*KRTAP3-2* a VICPA-*KRTAP3-1* y el gen SHEEP-*KRTAP4-12L* a VICPA-*KAP4-7L* (LOC10703281).
- b. *KRTAP5-4* (Anexo 5) se encuentra ubicado en un scaffold de alpaca (ABRR03077255.1) que aún no ha sido asignado a algún cromosoma mientras que el gen de oveja ha sido asignado al cromosoma 21 (OAR21), Sin embargo, OAR21 (Goldammer *et al.*, 2009) tiene una región homóloga al cromosoma bovino BTA25 (Zimin *et al.*, 2009), que tiene regiones homólogas a los cromosomas humanos HSA16p y HSA7, por lo tanto, también al cromosoma 18 de alpaca (VPA18) (Richardson *et al.*, 2019), con lo cual se podría predecir que esta secuencia se ubica en dicho cromosoma. Al analizar esta secuencia en el VicPac3.1, SHEEP-*KRTAP5-4* es equivalente a VICPA-*KRTAP5-5L*.

- c. *KRTAP9-7L* tiene tres copias (LOC114116843, LOC114116849 y LOC114116851) ubicadas en el cromosoma 11 de oveja (OAR11), en alpacas sólo se encontraría la copia LOC114116849 (Anexo 6) que se encuentra ubicada en el cromosoma 16 de alpaca (VPA16), lo que sería consistente con las homologías anteriormente mencionadas. Con los datos que se tienen actualmente del secuenciamiento del genoma de la alpaca, no se encuentra la copia LOC114116843, el bajo porcentaje de similitud se debe a que en la secuencia hay bases nucleotídicas que no se han podido identificar (280 pb). En el caso de las copias LOC114116849 y LOC114116851, ambas tienen la misma posición de inicio, sin embargo, sólo en LOC114116849 se puede identificar un exón pequeño de 33 aminoácidos, mientras que en el otro no se ha identificado el codón de inicio de transcripción, por lo que es probable que sea un pseudogen. Basados en la distancia evolutiva entre alpaca y oveja, el gen de alpaca, podría interpretarse como de un locus ancestral de los tres loci que presenta la oveja. Al analizar esta secuencia en el VicPac3.1, SHEEP-*KRTAP9-7L* (LOC114116849) es equivalente a VICPA-*KRTAP9-2L*.
- d. *KRTAP13-2L* y las dos copias del gen *KRTAP10-7L* (LOC101104456 y LOC114110483) (Anexo 7) se encuentran ubicadas en el cromosoma 1 de alpaca (VPA1); estas mismas secuencias se encuentran localizadas en el cromosoma 1 de oveja (OAR1q) siendo esto consistente con las homologías de VPA1 con las regiones de los cromosomas humanos HSA3q y HSA21q (Richardson *et al.*, 2019), que a su vez tienen regiones homólogas a BTA1 (Zimin *et al.*, 2009) y OAR1q (Goldammer *et al.*, 2009). Al analizar estas secuencias en el VicPac3.1, SHEEP-*KRTAP13-2L* es equivalente a VICPA-*KRTAP13-2p* SHEEP-*KRTAP10-7* (LOC101104456) al segundo exón de VICPA-*KRTAP10-8L\*3* y SHEEP-*KRTAP10-7L* (LOC114110483) a VICPA-*KRTAP10-12L*.

Además, el análisis comparativo de las secuencias entre oveja y alpaca nos indica que no existiría una secuencia equivalente al gen SHEEP-*KRTAP13-4L* en alpaca, el bajo porcentaje de similitud se debe a que existen muchos nucleótidos no identificados en la secuencia de la alpaca (365 pb).

La comparación de las secuencias entre cabras y alpacas nos permite reportar equivalencias entre las secuencias de los genes:

- a. *KRTAP1-4*, *KRTAP4-11L* y *KRTAP9-3L* (Anexo 8), se encuentran ubicados en el cromosoma 19 de cabra (CHI19) y en el cromosoma 16 de alpaca (VPA16), siendo esto

consistente ya que CHI19 (Schibler *et al.*, 2009) tiene una región homóloga al cromosoma humano HSA17, y este a su vez a VPA16 (Richardson *et al.*, 2019). Al analizar estas secuencias en el VicPac3.1, CAPHI-*KRTAP1-4* es equivalente al primer exón de VICPA-*KRTAP4-3* y CAPHI-*KRTAP4-11L* codifica para el segundo y tercer exón del gen VICPA-*KRTAP4-7L\*3* (LOC102529530). En el caso de la secuencia equivalente al gen CAPHI-*KRTAP9-3L* que no posee un codón de inicio de transcripción, por lo que es probable que sea un pseudogen.

- b. *KRTAP9-8L* (Anexo 9) se encuentra ubicado en el cromosoma 16 de alpaca (VPA16), sin embargo, en cabra no se ha asignado el scaffold donde se encuentra ubicado (ASM170441v1) a algún cromosoma, pero por homologías cromosómicas se podría predecir que este gen se encuentra en el cromosoma 19 en cabra (CHI19) ya que VPA16 tiene regiones homólogas al cromosoma humano HSA17 (Richardson *et al.*, 2019) y este a su vez a CHI19 (Schibler *et al.*, 2009). Al analizar esta secuencia en el VicPac3.1, CAPHI-*KAP9-8L* es equivalente al gen VICPA-*KRTAP4-7L\*4* (LOC102529530).
- c. *KRTAP10-9L* y *KRTAP16-4* (Anexo 10), que se encuentran ubicadas en el cromosoma 1 de alpaca (VPA1); estas mismas secuencias se encuentran localizadas en el cromosoma 1 de cabra (CHI1), siendo esto consistente con las homologías de VPA1 con las regiones de los cromosomas humanos HSA3 y HSA21 (Richardson *et al.*, 2019), y estos a su vez con CHI1 (Schibler *et al.*, 2009). Al analizar estas secuencias en el VicPac3.1, CAPHI-*KRTAP10-9L* es equivalente a VICPA-*KRTAP10-12L* y CAPHI-*KRTAP16-4* a VICPA-*KRTAP21-1L*.

La comparación de las secuencias entre las tres especies (Anexo 11) indica que los genes *KRTAP1-3* y *KRTAP9-9L* se encuentran ubicados en el cromosoma 11, 19 y 16 de oveja, cabra y alpaca respectivamente y el gen *KRTAP6-1L* se encuentra ubicado en el cromosoma 1 de las tres especies, siendo esto consistente con las homologías mencionadas anteriormente. Esta comparación también indica que los genes *KRTAP4-8L* y *KRTAP4-9L* (Anexo 12) se encuentran en la misma posición (455442-455951) en el scaffold de alpaca, lo que significaría que la alpaca posee sólo uno de estos dos genes, al analizar el alineamiento se encuentra una mayor similitud con el gen CAPHI-*KRTAP4-8L*, ya que poseen un *gap* en posiciones similares, sin embargo, el *gap* en alpaca es de 69 pb siendo menor que el *gap* en cabra (144 pb). Las otras alternativas son: que este locus represente el locus ancestral del cual evolucionaron los genes *KRTAP4-8L* y *KRTAP4-9L* en ovejas y cabras y/o que este



locus represente un locus específico de alpaca de una proteína asociada a la queratina que no está presente en el genoma de oveja y cabra.

En el caso de las secuencias equivalentes a los genes *KRTAP1-3* (cabra y oveja) y *CAPHI-KRTAP1-4*, se encuentran en posiciones superpuestas en el mismo scaffold, al realizar el análisis de las secuencias se identificó que ambas secuencias codifican para el mismo exón, el cual corresponde al primer exón del gen *VICPA-KRTAP4-3L* mientras que el gen *SHEEP-KRTAP1-1* codifica para el segundo exón de *VICPA-KRTAP4-3L*, por lo cual se presume que el gen *KRTAP4-3L* represente el locus ancestral del cual evolucionaron los genes mencionados en oveja y cabra. Asimismo, en las secuencias equivalentes a *CAPHI-KAP10-9L* y el gen *SHEEP-KAP10-7L* (LOC114110483), también existe una superposición y ambas codifican para el mismo gen, que se encuentra anotado en el VicPac3.1 como *VICPA-KAP10-12L*.

Los genes *KRTAP10-11L* y *KRTAP17-1*, tienen un porcentaje de identidad muy bajo, lo que indicaría que no es un gen conservado entre estas especies o que son genes distintos.

El análisis comparativo entre las secuencias de los 34 genes de proteínas asociadas a queratinas de alpaca anotados en el NCBI (Tabla 3) y los genomas de referencia de oveja y cabra nos permite reportar que 21 de estos genes tienen un porcentaje de identidad mayor al 70% con los genes de cabra y oveja, 13 tienen un porcentaje de identidad menor al 70% con genes de cabra y 8 con los genes de oveja. Mientras que, entre oveja y cabra, sólo 3 tienen un porcentaje de identidad menor al 70%. Estos resultados sustentan la búsqueda de otros *KRTAP* mediante el análisis comparativo con las especies mencionadas. En cuanto a los análisis sobre homologías cromosómicas, son consistentes con lo mencionado anteriormente, las secuencias de genes que se encuentran ubicadas en el cromosoma 1 de alpaca (VPA1); tienen sus secuencias equivalentes localizadas en el cromosoma 1 de cabra (CHI1) y en el cromosoma 1 de oveja (OAR1), siendo esto consistente con las homologías de regiones cromosómicas de VPA1 con las cromosomas humanos HSA3 y HSA21 (Richardson *et al.*, 2019), y estos a su vez con CHI1 (Schibler *et al.*, 2009) y OAR1 (Dalrymple *et al.*, 2007). Las secuencias de genes que se encuentran ubicadas en el cromosoma 16 de alpaca (VPA16), tienen sus secuencias equivalentes localizadas en el cromosoma 19 en cabra (CHI19) y en el cromosoma 11 de oveja (OAR11), siendo consistente ya que VPA16 tiene regiones cromosómicas homólogas al cromosoma humano HSA17 (Richardson *et al.*, 2019) y este a su vez a CHI19 (Schibler *et al.*, 2009) y OAR11

(Dalrymple *et al.*, 2007). Además, la secuencia equivalente al gen *KRTAP10-10* que se encuentra ubicado en el cromosoma 18 de alpaca (VPA18), se encuentra localizada en el cromosoma 24 de oveja (OAR24), siendo consistente ya que VPA18 tiene regiones cromosómicas homólogas al cromosoma humano HSA7 y HSA16p (Richardson *et al.*, 2019) y estas a su vez con OAR24 (Dalrymple *et al.*, 2007).

Mediante este estudio bioinformático se logró identificar 67 marcadores de tipo PNS en 32 de los 34 genes *KRTAPs* anotados en el NCBI. Esta técnica es eficiente cuando se posee la base de datos de por lo menos dos genomas. Además, al realizar el filtrado con los parámetros que se tomaron en cuenta para este estudio, se puede considerar que los PNS son informativos y confiables.

En la identificación de polimorfismos de nucleótidos simples en los genes de proteínas asociadas a queratinas, según su clasificación se reporta lo siguiente:

**a. Genes *KRTAPs* que codifican proteínas con alto contenido de sulfuro**

Los genes LOC102528468 y LOC102528209, se encuentran anotados en el NCBI como genes modelo de VICPA-*KRTAP3-3L*. Sin embargo, al realizar un análisis de las secuencias de las proteínas para las que codifican estos genes son idénticas, lo que significaría que el primer exón de la versión del gen con dos exones (LOC102528209) podría tener una función reguladora.

Se identificó un PNS (g.12376109T>G) en el gen *KRTAP11-1*, este resultado concuerda con lo hallado por Foppiano (2016), quien identificó este mismo PNS al cual señala como SNP330 y con el que realiza un análisis de asociación, sin obtener resultados significativos entre la asociación de este PNS con el diámetro de fibra en alpacas. Sin embargo, Foppiano (2016) identificó otros cinco PNSs en el gen *KRTAP11-1*, ello se puede deber a que en el presente estudio los PNSs deben cumplir parámetros de calidad como frecuencia del alelo menor, tamaño de la secuencia flanqueante, entre otros, mientras que Foppiano (2016) sólo consideró el parámetro de frecuencia alélica (5%). En ovejas, Gong *et al.* (2011) identificaron cinco posiciones polimórficas en la región codificante, de las cuales una produce un cambio en la secuencia de la proteína presentándose los aminoácidos Serina o Glicina, por el contrario, en este estudio el único PNS identificado, produce un cambio presentándose los aminoácidos Serina o Alanina en la secuencia de la proteína.

El gen *KRTAP13-1* posee 3 copias (LOC102533350, LOC102532866, LOC102532037), el PNS (g.12846534G>C) identificado en *KRTAP13-1* (LOC102532866) concuerda con el identificado por Foppiano (2016) menciona que produce un cambio de Triptófano a Cisteína en la secuencia de la proteína, lo que es consistente con el resultado del presente estudio. Cabe mencionar que Foppiano (2016) realizó un alineamiento entre el genoma de alpaca y la secuencia del gen *KRTAP13-1* en humanos para obtener la secuencia del gen en alpacas. Además, identificó otros tres PNSs en la misma secuencia, ello se puede deber a que en el presente estudio los PNSs deben cumplir parámetros de calidad como frecuencia del alelo menor, tamaño de la secuencia flanqueante, entre otros, mientras que el autor mencionado sólo consideró el parámetro de frecuencia alélica (5%). La secuencia en la que Foppiano (2016) identificó dichos PNSs, corresponde a la copia del gen *KRTAP13-1* (LOC102532866), sin embargo, en la presente investigación se identificaron PNSs en todas las copias del gen anotadas en el NCBI, cuatro en LOC102533350 de los cuales dos son no sinónimos, produciéndose cambios en la secuencia de la proteína presentándose los aminoácidos Treonina o Serina (g.12923905C>G) y Valina o Fenilalanina (g.12924045G>T); y dos en LOC102532037 que producen cambios en la secuencia de la proteína presentándose los aminoácidos Alanina o Serina (g.12859532G>T) y Serina o Cisteína (g.12859761C>G). En cabras, Fang *et al.* (2010) identificaron un PNS en la región codificante para la proteína, dicho PNS está significativamente asociado con el diámetro de la fibra (cashmere) en cabras de la raza Xinjiang, por lo que sería importante considerar los PNSs identificados en este estudio para un análisis de asociación con características de la fibra en alpacas.

En el gen *KRTAP24-1*, Wang *et al.* (2019) identificaron tres PNSs no sinónimos en caprinos c.301AT/C, c.319G/A y c.656T/C, que producen cambios en las secuencias de las proteínas presentándose los aminoácidos Cisteína o Arginina, Glicina o Arginina y Valina o Alanina respectivamente, en dicho estudio se determinó que los cambios mencionados afectan el diámetro de la fibra (cashmere), mientras que en el presente estudio uno de los cuatro PNSs identificados también produce un cambio en la secuencia de la proteína presentándose los aminoácidos Alanina o Valina (g.12980054C>T), el cual sería importante considerar en un análisis de asociación con características de la fibra en alpacas. De igual manera, se deben analizar los otros PNSs identificados, un PNS sinónimo (g. 12979560T>G) y dos PNSs no sinónimos que producen cambios en las secuencias de las proteínas presentándose los

aminoácidos Glicina o Histidina (g.12979659A>T) y Asparagina o Serina (g.12979802A>G).

En el gen *KRTAP26-1*, Li *et al.* (2017) identificaron cuatro PNSs de los cuales el denominado C (c.277A/G), produce un cambio de Serina a Glicina, que al realizar estudios de asociación en ovejas Merino x Southdown, aquellas que poseen C tienen en promedio una calidad de lana más alta, ya que tienen un diámetro medio de lana, desviación estándar de diámetro de lana y factor de confort promedio más bajo, y un mayor rendimiento al lavado, lo que es consistente con las correlaciones encontradas entre estos rasgos. En el presente estudio también se identificaron cuatro PNSs, sin embargo, ninguno produce un cambio de Serina a Glicina. Los PNSs identificados son: un PNS sinónimo y tres no sinónimos que producen los siguientes cambios en las secuencias de las proteínas presentándose los aminoácidos Serina o Prolina (g.12947455T>C), Leucina o Glutamina (g.12947708T>A), Arginina a Cisteína (g.12947881C>T), los cuales sería importantes considerar para un análisis de asociación con características de la fibra.

#### **b. Genes *KRTAPs* que codifican proteínas con muy alto contenido de sulfuro**

En el gen *KRTAP9-2*, Yu *et al.* (2008) identificaron en cabras productoras de fibra cashmere una variación del tipo inserción/delección de 30 nucleótidos en el exón del gen. También, Wang *et al.* (2012) identificaron un PNS con variación C/T en este gen. En otro estudio, Wang *et al.* (2014) realizaron un perfil de expresión del gen en dos fases distintas del ciclo de la fibra y encontraron que existe una diferencia significativa en la expresión de este gen entre cabras con alta producción y baja producción de fibra cashmere. Por último, Foppiano (2016) identificó tres PNSs en el genoma de la alpaca, dos en el exón y uno en el 3'UTR, este último no cumplió con el parámetro de frecuencia alélica (5%), mientras que en el presente estudio se encontró un PNS en la región no traducida del gen *KRTAP9-2*, que no es el identificado por Foppiano (2016) y que sería importante seguir analizando ya que puede afectar la expresión del gen.

#### **c. Genes *KRTAPs* que codifican proteínas con alto contenido de glicina y tirosina**

En el gen *KRTAP7-1*, Jin *et al.* (2011) encontró en cabras Liaoning que la expresión de este gen en los folículos secundarios era 1.28 veces mayor que en los folículos primarios, lo que sugiere que el gen *KRTAP7-1* probablemente tiene un rol importante en la regulación de la finura. Liu *et al.* (2014) identificó un PNS en oveja Merino China que producía un cambio en la secuencia de la proteína presentándose los aminoácidos Asparagina o Lisina. Sin

embargo, en el presente estudio se identificó un PNS no sinónimo en el único exón de este gen, el cual produce un cambio en la secuencia de la proteína presentándose los aminoácidos Serina o Arginina (g.12420218C>G).

En el gen *KRTAP8-1*, Liu *et al.* (2011), encontró dos mutaciones sinónimas, de las cuales una de ellas T113G estaba significativamente asociada al peso del vellón de cashmere, pero no al diámetro de la fibra. Por el contrario, en el presente estudio se hallaron dos PNSs, uno es sinónimo (g.12436062T>C) y el otro es no sinónimo (g.12436103G>A), el cual produce un cambio en la secuencia de la proteína presentándose los aminoácidos Cisteína o Tirosina, que sería importante considerar para un análisis de asociación con características de la fibra.

En el gen *KRTAP20-2*, Bai *et al.* (2018) identificaron una sustitución nucleotídica (c.160A>T) que produce un codón terminal prematuro, esta se asoció al diámetro medio de lana, pero además este rasgo en estas ovejas (Merino x Southdown) está correlacionado al diámetro de curvatura media. Sin embargo, en el presente estudio el único PNS no sinónimo produce un cambio en la secuencia de la proteína presentándose los aminoácidos Alanina o Valina (g.12678801C>T), mientras que los otros cuatro PNSs identificados son sinónimos.

En el gen *KRTAP21-1*, Li *et al.* (2019) identificaron en corderos Merino x Southdown cuatro variantes ubicadas en el único exón del gen, de los cuales una se asoció con una variación positiva en el rendimiento al lavado. Por el contrario, en el presente estudio no se identificaron variantes en el exón, se identificaron dos PNSs en el UTR del gen *KRTAP21-1* en alpacas y sería importante seguir analizando ya que pueden afectar la expresión del gen.

## V. CONCLUSIONES

- Se obtuvo la localización de 34 genes de proteínas asociadas a las queratinas en alpacas.
- Se identificó y se extrajo la secuencia de 71 pb de 67 PNSs en genes de proteínas asociadas a las queratinas, con valores mayores o iguales a 0.05 para el parámetro de frecuencia del alelo menor, mayores o iguales a 0.6 para Illumina Score y una tasa mayor al 45% de tasa de genotipaje por PNS basado en secuencias *reads* en nueve genomas (almacenados en NCBI) y *reads* de 300 bibliotecas de representación reducida de ADN de alpacas.
- Por homologías de regiones cromosómicas, se puede predecir que el scaffold ABRR03077255.1 de alpaca se ubicaría una región del cromosoma 16 (VPA16) y el scaffold ASM170441v1 de cabra en una región del cromosoma 19 (CHI19).

## VI. RECOMENDACIONES

- Validar los PNSs identificados haciendo uso de técnicas como el desarrollo de pruebas de PCR para el genotipado de PNS, genotipado de PNSs con sondas TaqMan, entre otros.
- Evaluar la asociación entre los PNS identificados en los genes de proteínas asociadas a las queratinas y las características de valor comercial de la fibra de alpaca.
- Realizar la búsqueda de PNSs en las regiones cercanas al gen *KRTAP2-4L* y la copia del gen *KRTAP4-7L* (LOC102529530).
- Hacer una evaluación bioquímica de los *KRTAPs* que se expresan en el folículo de la alpaca y alinear la nomenclatura dependiendo de los resultados.

## VII. BIBLIOGRAFÍA

- Antonini, M.; González, M. & Valbonesi, A. (2004). Relationship between age and postnatal skin follicular development in three types of South American domestic camelids. *Livestock Production Science*, 90: 241–246.
- Bai, L.; Gong, H.; Zhou, H.; Tao, J. & Hickford, J. (2018). A nucleotide substitution in the ovine KAP20-2 gene leads to a premature stop codon that affects wool fibre curvature. *Anim. Genet.*, 49: 357-358. doi: 10.1111/age.12668
- Balding, D.J. (2006). A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*, 7(10): 781-791.
- Bonavia, D. (1996). Los Camélidos Sudamericanos (Una introducción a su estudio). Instituto Francés de Estudios Andinos. Universidad Peruana Cayetano Heredia. Conservation International.
- Brookes, A. (1999). The essence of SNPs. *Gene*, 234(2): 177-186. doi: 10.1016/S0378-1119(99)00219-X.
- Calderon, M., More, M.J., Gutierrez, G.A., & Ponce de León, F.A. (2021). Development of a 76k Alpaca (*Vicugna pacos*) Single Nucleotide Polymorphisms (SNPs) Microarray. *Genes*, 12(2): 291. <https://doi.org/10.3390/genes12020291>
- Cargill, M.; Altshuler, D.; Ireland, J.; Sklar, P.; Ardlie, K.; Patil, N.; Shaw, N.; Lane, C.; Lim, E.; Kalyanaraman, N.; Nemesh, J.; Ziaugra, L.; Friedland, L.; Rolfe, A.; Warrington, J.; Lipshutz, R.; Daley, G. & Lander, E. (1999). Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet*; 22(3):231-8. doi: 10.1038/10290.
- Contreras, A. (2009). Estructura cuticular y características físicas de la fibra de alpaca Huacaya (*Vicugna pacos*) de color blanco en la región de Huancavelica. (Tesis de Ingeniero Zootecnista). Universidad Nacional de Huancavelica, Huancavelica, Perú.
- Czaplicki, Z. (2012). Properties and Structure of Polish Alpaca Wool. *Institute of Natural fibres & Medicinal plant*. 1(90), 8–12.



- Dalrymple, B.; Kirkness, E.; Nefedov, M.; McWilliam, S.; Ratnakumar, A.; Barris, W.; Zhao, S.; Shetty, J.; Maddox, J.; O'Grady, M.; Nicholas, F.; Crawford, A.; Smith, T.; De Jong, P.; McEwan, J.; Hutton Oddy, V. & Cockett, N.; the International Sheep Genomics Consortium. (2007). Using comparative genomics to reorder the human genome sequence into a virtual sheep genome. *Genome Biology*, 8: R152. doi: 10.1186/gb-2007-8-7-r152
- Delgado De La Flor, I. (2014). Caracterización de nuevos marcadores genéticos microsatélites e identificación de SNP en el gen de Tricohialina en alpacas (*Vicugna pacos*). Tesis Magíster en Bioquímica y Biología Molecular. Universidad Peruana Cayetano Heredia, Lima, Perú.
- Dunner, S. & Cañon, J. (2001). Aplicaciones de Genómica en Laboratorios de Producción Animal. Laboratorio de Genética. Facultad de Veterinaria de la UCM. Madrid - España.
- FAO. (2005). Situación Actual de los Camélidos Sudamericanos en el Perú. Organización de las Naciones Unidas para la Agricultura y la Alimentación. Proyecto de Cooperación Técnica en apoyo a la crianza y aprovechamiento de los Camélidos Sudamericanos en la Región Andina.
- FAO. (2007). Marker Assisted Selection: current status and perspectives in crops, livestock, forestry and fish. Organización de las naciones unidas para la agricultura y la alimentación. Roma.
- FAO. (2010). La situación de los recursos zoogenéticos mundiales para la alimentación y la agricultura. Comisión de recursos genéticos para la alimentación y la agricultura. Organización de las naciones unidas para la agricultura y la alimentación. Roma.
- Fang, Y.; Liu, W.; Zhang, F.; Shao, Y. & Yu, S. (2010). The polymorphism of a novel mutation of KAP13.1 gene and its associations with cashmere traits on Xinjiang local goat breed in China. *Asian Journal of Animal and Veterinary Advances*, 5(1): 34-42. doi: 10.3923/ajava.2010.34.42
- Fernández, A.; Gutiérrez, G. & Ponce de León, F. (2019). Identificación bioinformática de Polimorfismos de Nucleótido Simple (PNSs) en genes candidatos para las características de la fibra en alpacas (*Vicugna pacos*). *Revista Peruana de Biología*, 26(1): 87-94. doi: 10.15381/rpb.v26i1.15911

- Foppiano, F. (2016). Caracterización de marcadores genéticos en genes que codifican a proteínas asociadas a queratina y evaluación de la asociación del gen *KRTAP11-1* al diámetro de fibra en alpaca (*Vicugna pacos*) siguiendo una aproximación de gen candidato. Tesis Magister en Bioquímica y Biología Molecular. Universidad Peruana Cayetano Heredia, Lima, Perú.
- Fujimoto, S.; Takase, T.; Kadono, N.; Maekubo, K. & Hirai, Y. (2014). *Krtap11-1*, a hair keratin-associated protein, as a possible crucial element for the physical properties of hair shafts. *Journal of Dermatological Science*, 74(1): 39-47.
- Gentry, A.; Clutton-brock, J. & Groves, C. (2004). The naming of wild animal species and their domestic derivatives. *Journal of Archaeological Science*, 31: 645-651.
- Gillespie, J.M. (1972). Proteins rich in glycine and tyrosine from keratins. Comparative biochemistry and physiology. *Comparative Biochemistry*, 41(4):723–34.
- Gillespie, J.M. & Broad, A. (1972). Ultra-high-sulphur proteins in the hairs of the artiodactyla. *Australian journal of biological sciences*, 25(1): 138–45.
- Goldammer, T.; Meo, G.; Lühken, G.; Drögemüller, C.; Wu, C.; Kijas, J.; Dalrymple, B.; Nicholas, F.; Maddox, J.; Iannuzzi, L. & Cockett, N. (2009). Molecular Cytogenetics and Gene Mapping in Sheep (*Ovis aries*, 2n=54). *Cytogenetic and genome research*, 126: 63-76. doi: 10.1159/000245907
- Gong, H.; Zhou, H.; Mckenzie, G.; Yu, Z.; Clerens, S.; Dyer, J.; Plowman, J.; ... & Hickford, J. (2012). An updated nomenclature for keratin-associated proteins (KAPs). *International Journal of Biological Sciences*, 8 (2): 258-264.
- Gong, H.; Zhou, H.; Dyer, J. & Hickford, J. (2011). Identification of the ovine KAP11-1 gene (*KRTAP11-1*) and genetic variation in its coding sequence. *Molecular biology reports*, 38(8): 5429–5433. doi: 10.1007/s11033-011-0697-2
- Gong, H.; Zhou, H.; Forrest, R.; Li, S.; Wang, J.; Dyer, J. & Hickford, J. (2016). Wool Keratin-Associated Protein Genes in Sheep-A Review. *Genes*, 7(6): 24.
- Gutiérrez, G. (2008). Revisión de la estimación de los parámetros genéticos en alpacas. En: Quispe, E.; Mueller, J.; Ruíz, J.; Alfonso, L.; Gutiérrez, G. Actualidades sobre adaptación, producción, reproducción y mejora genética en Camélidos. (p.124). Huancavelica, Perú: Universidad Nacional de Huancavelica.

- Hoffman, E. & Fowler, M. (1995). The alpaca book. Clay Press Inc., Herald, California.
- INEI. (2012). Resultados Definitivos. IV Censo Nacional Agropecuario 2012.
- Jin, M.; Wang, L.; Li, S.; Xing, M. & Zhang, X. (2011). Characterization and expression analysis of KAP7.1, KAP8.2 gene in Liaoning new-breeding Cashmere goat hair follicle. *Mol. Biol. Rep.* 38: 3023-3028. doi: 10.1007/s11033-010-9968-6
- Höcker, H. (2002). Fibre morphology. En: W. Simpson & G. Crawshaw (Eds), Wool: Science and technology. (p. 60-79). Cambridge, England: Woodhead Publishing Limited.
- Kadwell, M.; Fernandez, M.; Stanley, H.; Baldi, R.; Wheeler, J.; Rosadio, R. & Bruford, M. (2001). Genetic analysis reveals the wild ancestors of llama and alpaca. *Proceeding of the Royal Society London B.*, 268: 2575- 2584.
- Khan, I.; Maldonado, E.; Vasconcelos, V.; O'brien, S.; Johnson, W. & Antunes, A. (2014). Mammalian keratin associated proteins (*KRTAPs*) subgenomes: disentangling hair diversity and adaptation to terrestrial and aquatic environments. *BMC genomics*, 15(1): 1.
- Langbein, L. & Schweizer, J. (2005). Keratins of the Human Hair. *International Review of Cytology*, 243: 1-78.
- Li, S.; Zhou, H.; Gong, H.; Zhao, F.; Hu, J.; Luo, Y. & Hickford, J. (2017). Identification of the Ovine Keratin-Associated Protein 26-1 Gene and Its Association with Variation in Wool Traits. *Genes*, 8(9), 225. doi: 10.3390/genes8090225
- Li, S.; Zhou, H.; Gong, H.; Zhao, F.; Wang, J.; Liu, X.; Hu, J.; Luo, Y. & Hickford, J. (2019). Identification of the Ovine Keratin-Associated Protein 21-1 Gene and Its Association with Variation in Wool Traits. *Animals: an open access journal from MDPI*, 9(7):450. doi: 10.3390/ani9070450
- Liu, H.; Yue, C.; Zhang, W.; Zhu, X.; Yang, G. & Jia, Z. (2011). Association of the KAP 8.1 Gene Polymorphisms with Fibre Traits in Inner Mongolian Cashmere Goats. *Asian-Australas J. Anim. Sci.*;24(10):1341-1347. doi: 10.5713/ajas.2011.11120
- Liu, Y.; Shi, G.; Wang, H.; Wan, P.; Tang, H.; Yang, H. & Guan, F. (2014). Polymorphisms of KAP6, KAP7 and KAP8 genes in four Chinese sheep breeds. *Genet. Mol. Res.* 13: 3438–45. doi: 10.4238/2014.April.30.5

- Marín, J.; Zapata, B.; González, B.; Bonacic, C.; Wheeler, J.; Casey, C.; Bruford, M.; Palma, E.; Poulin, E.; Alliende, A. & Spotorno, E. (2007). Sistemática, taxonomía y domesticación de alpacas y llamas: nueva evidencia cromosómica y molecular. *Revista Chilena Historia Natural*, 80: 121-140.
- Mengoni, G. (2008). Camelids in ancient Andean societies: a review of the zooarchaeological evidence. *Quaternary International*, 185, 59-68.
- Morin, P.; Luikart, G. & Wayne, R. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, 19 (4): 208-216.
- NCBI Resource Coordinators (2013). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 41, Issue D1: D8–D20. doi: 10.1093/nar/gks1189
- Orwin, D. (1979). The cytology and cytochemistry of the wool follicle [Includes fiber development and properties]. *International review of cytology*, 60: 331-374.
- Perez-Cabal, M.; Cervantes, I.; Morante, R.; Burgos, A.; Goyache, F. & Gutiérrez, J.P. (2010). Analysis of the existence of major genes affecting alpaca fiber traits. *Journal of Animal Science*, 88: 3783-3788.
- Pierce, B. (2006). *Genética: Un enfoque conceptual*. 915 pp. Editorial Panamericana.
- Plowman, J.E. (2003). Proteomic database of wool components. *Journal of Chromatography*, 787: 63–76.
- Ponzoni, R. (1996). *Manual de prácticas de manejo de alpacas y llamas. Estudio Producción y Sanidad Animal*. Vol. 130. Rome, Italy: FAO.
- Powell, B.C. & Rogers, G.E. (1997). The role of keratin proteins and their genes in the growth, structure and properties of hair. *Exs*, 78: 59-148.
- Quispe, E.; Paucar, R.; Poma, A.; Sachero, D. & Mueller, J. (2008). Perfil de diámetro de fibra en alpacas, Seminario Internacional de Biotecnología Aplicada en Camélidos Sudamericanos. Huancavelica, Perú: Universidad Nacional de Huancavelica.
- Quispe, E.; Rodríguez, T.; Ñequez, L. & Muller, J.P. (2009). Producción de fibra de alpaca, llama, vicuña y guanaco en Sudamérica. *Animal Genetic Resources Information* (FAO), 45: 1-14.

- Richardson, M.F.; Munyard, K.; Croft, L.J.; Allnutt, T.R.; Jackling, F.; Alshanbari, F.; Jevit, M.; Wright, G.; Cransberg, R.; Tibary, A.; Perelman, P.; Appleton, B. & Raudsepp, T. (2019). Chromosome-Level Alpaca Reference Genome VicPac3.1 Improves Genomic Insight into the Biology of New World Camelids. *Front. Genet.* 10:586. doi: 10.3389/fgene.2019.00586
- Rogers, G. (2004). Hair follicle differentiation and regulation. *The International Journal of Developmental Biology*; 48: 163-170.
- Rogers, M.; Langbein, L.; Praetzel-wunder, S.; Winter, H. & Schweizer, J. (2006). Human hair keratin-associated proteins (KAPs). *International review of cytology*, 251(6): 209–63.
- Salas, W. (2019). Caracterización de marcadores moleculares de genes involucrados en la estructura y desarrollo de la fibra de alpaca y su potencial asociación con el diámetro de la fibra. Tesis para optar el Grado Académico de Doctor en Ciencias Biológicas. Universidad Nacional Mayor de San Marcos, Lima, Perú.
- Schibler, L.; Meo, G.; Cribiu, E. & Iannuzzi, L. (2009). Molecular Cytogenetics and Comparative Mapping in Goats (*Capra hircus*, 2n=60). *Cytogenetic and genome research*, 126: 77-85. doi: 10.1159/000245908
- Schweizer, J.; Bowden, P.; Coulombe, P.; Langbein, L.; Lane, B.; Magin, T.; Maltais, L.; Omary, M.; Parry, D.; Rogers, M. & Wright, M. (2006). New consensus nomenclature for mammalian keratins. *The Journal of cell Biology*, 174(2): 169 – 174.
- Stanley, H.; Kadwell, M. & Wheeler, J. (1994). Molecular evolution of the family Camelidae: a mitochondrial DNA study. *Proceedings of the Royal Society*, 256: 1-6.
- Wang, X.; Huarong, X.; Li, T.; Qu, L.; Zhao, Z. & Zhang, Z. (2014). Expression analysis of KAP9.2 and Hoxc13 genes during different cashmere growth stages by qRT-PCR method. *Molecular biology reports*, 41(9). doi: 10.1007/s11033-014-3435-8
- Wang, B.; Yang, W.; Mckittrick, J. & Meyers, M.A. (2016). Keratin: Structure, mechanical properties, occurrence in biological organisms, and efforts at bioinspiration. *Progress in Materials Science*, 76, 229-318.

- Wang, J.; Zhou, H.; Luo, Y.; Zhao, M.; Gong, H.; Hao, Z.; Hu, J. & Hickford, J. (2019). Variation in the Caprine KAP24-1 Gene Affects Cashmere Fibre Diameter. *Animals: an open access journal from MDPI*, 9(1): 15. doi:10.3390/ani9010015
- Wang, X., Wang, L. & Liu, X. (2003). The Quality and Processing Performance of Alpaca Fibres: A report for the Rural Industries Research and Development Corporation. RIRDC Publication N° 03/128. Australia. 132 p.
- Wang, X.; Zhao, Z.; Xu, H.; Qu, L.; Zhao, H.; Li, T. & Zhang, Z.Y. (2012). Variation and expression of KAP9.2 gene affecting cashmere trait in goats. *Molecular biology reports*, 39(12), 10525-10529.
- Wu, D.; Irwin, D. & Zhang, Y. (2008). Molecular evolution of the keratin associated protein gene family in mammals, role in the evolution of mammalian hair. *BMC evolutionary biology*, 8: 241
- Wuliji, T.; Davis, G.; Dodds, K.; Turner, P.; Andrews, R. & Bruce, G. (2000). Production performance, repeatability and heritability estimates for live weight, fleece weight and fiber characteristics of alpacas in New Zealand. *Small Ruminant Research* 37, 189–201. doi: 10.1186/1471-2148-8-241
- Yu, H.; Wang, X.; Chen, H.; Wang, M.; Zhao, M.; Lan, X. & Wang, X.L. (2008). The polymorphism of a novel 30bp-deletion mutation at KAP9.2 locus in the Cashmere goat. *Small Ruminant Research*, 80(1): 111-115. doi: 10.1016/j.smallrumres.2008.09.010
- Yu, Z.; Gordon, S.; Nixon, A.; Bawden, C.; Rogers, M.; Wildermoth, J.; Maqbool, N. & Pearson, A. (2009). Expression patterns of keratin intermediate filament and keratin associated protein genes in wool follicles. *Differentiation*, 77(3):307-316. doi: 10.1016/j.diff.2008.10.009.
- Zimin, A.; Delcher, A.; Florea, L.; Kelley, D.; Schatz, M.; Puiu, D.; Hanrahan, F.; Pertea, G.; Tassell, C.; Sonstegard, T.; Marcais, G.; Roberts, M.; Subramanian, P.; Yorke, J. & Salzberg, S. (2009). A Whole-Genome Assembly of the Domestic Cow, *Bos Taurus*. *Genome biology*. 10: R42. doi: 10.1186/gb-2009-10-4-r42

## **VIII. ANEXOS**

**Anexo 1:** Valores de los parámetros considerados para cada PNS de los genes *KRTAPs* que codifican proteínas con alto contenido de sulfuro.

Gen	Posición en el scaffold	Illumina Score	Frecuencia del alelo menor				Genome Studio
			UNALM	UPCH	PstI_MspI	ApeKI	
<i>KRTAP3-1</i>	514886	0.737	0.08333				
<i>KRTAP3-3L</i> (LOC102528468)	526383	0.834	0.5	0.1667			0.387
<i>KRTAP3-3L</i> (LOC102528468)	526561	0.915	0.1667				
<i>KRTAP3-3L</i> (LOC102528209)	532094	1	0.08333				0.007
<i>KRTAP3-3L</i> (LOC102528209)	532200	0.685	0.25				0.062
<i>KRTAP10-1L</i>	12853114	0.988	0.25				0.098
<i>KRTAP10-1L</i>	12853197	0.949	0.1667			0.02459	0.02
<i>KRTAP10-1L</i>	12888037	1	0.08333				
<i>KRTAP10-1L</i>	12888077	0.949	0.08333			0.0303	
<i>KRTAP10-1L</i>	12888117	0.787	0.08333			0.05782	
<i>KRTAP10-1L</i>	12888170	0.737	0.25	0.1667		0.1681	0.168



Continuación...

Gen	Posición en el scaffold	Illumina Score	Frecuencia del alelo menor				Genome Studio
			UNALM	UPCH	PstI_MspI	ApeKI	
<i>KRTAP10-1L</i>	12888236	0.834	0.1667	0.5	0.3533	0.4727	
<i>KRTAP10-1L</i>	12888275	0.915	0.1667	0.5		0.4815	
<i>KRTAP10-1L</i>	12888315	0.877	0.25	0.1667		0.2632	
<i>KRTAP10-1L</i>	12888357	0.949	0.3333	0.1667		0.2632	
<i>KRTAP10-10</i>	8124973	0.915	0.2			0	
<i>KRTAP 11-1</i>	12376109	0.996	0.1667	0.3333		0.445	
<i>KRTAP12-1L</i> (LOC102528691)	764	0.633	0.08333				
<i>KRTAP12-1L</i> (LOC116283958)	3777628	0.877	0.5				
<i>KRTAP12-2L</i> (LOC102530622)	3775478	0.877	0.25	0.1667			
<i>KRTAP12-2L</i> (LOC116280071)	807	0.996	0.25				0.31
<i>KRTAP12-2L</i> (LOC116280071)	902	0.877	0.25	0.1667			0.314
<i>KRTAP12-2L</i> (LOC116280071)	1519	0.633	0.25				0.331

Continuación...

Gen	Posición en el scaffold	Illumina Score	Frecuencia del alelo menor				Genome Studio
			UNALM	UPCH	PstI_MspI	ApeKI	
<i>KRTAP13-IL</i> (LOC102533350)	12923762	0.787	0.08333				0.044
<i>KRTAP13-IL</i> (LOC102533350)	12923825	0.877	0.5	0.1667			0.375
<i>KRTAP13-IL</i> (LOC102533350)	12923905	0.616	0.1667	0.5			
<i>KRTAP13-IL</i> (LOC102533350)	12924045	0.949	0.08333				0.024
<i>KRTAP13-IL</i> (LOC102532866)	12846534	0.854	0.5	0.1667			
<i>KRTAP13-IL</i> (LOC102532037)	12859532	0.877	0.25				
<i>KRTAP13-IL</i> (LOC102532037)	12859761	1	0.08333				0.054
<i>KRTAP16-1</i>	353702	0.915			0.2199	0.1235	
<i>KRTAP16-1</i>	354957	0.787	0.08333				0.007
<i>KRTAP16-1</i>	355358	0.996	0.3333	0.1667	0.1081	0.1233	0.115
<i>KRTAP16-1</i>	355869	0.949	0.5	0.1667	0.2028		0.262
<i>KRTAP24-1</i>	12979560	0.949	0.25	0.3333			0.186

Continuación...

Gen	Posición en el scaffold	Illumina Score	Frecuencia del alelo menor				Genome Studio
			UNALM	UPCH	PstI_MspI	ApeKI	
<i>KRTAP24-1</i>	12979659	1		0.1667			0.027
<i>KRTAP24-1</i>	12979802	0.915	0.25	0.1667		0.2133	0.213
<i>KRTAP24-1</i>	12980054	1		0.3333		0.1039	0.104
<i>KRTAP26-1</i>	12947455	0.877		0.1667	0.01942		
<i>KRTAP26-1</i>	12947628	0.834	0.3333	0.1667	0.3514	0.2746	0.275
<i>KRTAP26-1</i>	12947708	0.737	0.08333			0.03125	
<i>KRTAP26-1</i>	12947881	0.915	0.25	0.1667		0.2542	
<i>KRTAP27-1</i>	12931105	0.996	0.5	0.1667		0.3933	
<i>KRTAP27-1</i>	12931214	0.996		0.1667			
<i>KRTAP27-1</i>	12931576	1	0.08333	0.1667			
<i>KRTAP29-1</i>	361421	0.965	0.1667				0.142
<i>KRTAP29-1</i>	361604	0.976	0.1667				0.142
<i>KRTAP29-1</i>	361748	0.915	0.3333	0.3333			0.429

UNALM: Datos de secuenciamiento de 6 genomas completos de la Universidad Nacional Agraria La Molina.

UPCH: Datos de secuenciamiento de 3 genomas completos de la Universidad Peruana Cayetano Heredia.

PstI\_MspI: Datos de secuenciamiento de 150 bibliotecas de representación reducidas obtenidas a partir de la digestión doble con enzimas PstI\_MspI.

ApeKI: Datos de secuenciamiento de 150 bibliotecas de representación reducidas obtenidas a partir de la digestión con la enzima ApeKI.

**Anexo 2:** Valores de los parámetros considerados para cada PNS de los genes *KRTAPs* que codifican proteínas con muy alto contenido de sulfuro.

Gen	Posición en el scaffold	Illumina Score	Frecuencia del alelo menor				Genome Studio
			UNALM	UPCH	PstI_MspI	ApeKI	
<i>KRTAP4-1</i>	400197	0.976	0.3333	0.3333			0.405
<i>KRTAP4-1</i>	403985	0.834	0.3333	0.3333			0.49
<i>KRTAP4-1</i>	404135	0.662		0.1667			0.061
<i>KRTAP4-3L</i>	500673	0.787	0.08333				
<i>KRTAP4-7L</i> (LOC107032861)	411461	0.737	0.08333		0.06015		0.06
<i>KRTAP9-2L</i>	379872	0.949	0.5				0
<i>KRTAP9-4L</i>	3527	0.834	0.5	0.5			
<i>KRTAP9-4L</i>	3574	0.877	0.1667				

UNALM: Datos de secuenciamiento de 6 genomas completos de la Universidad Nacional Agraria La Molina.

UPCH: Datos de secuenciamiento de 3 genomas completos de la Universidad Peruana Cayetano Heredia.

PstI\_MspI: Datos de secuenciamiento de 150 bibliotecas de representación reducidas obtenidas a partir de la digestión doble con enzimas PstI\_MspI.

ApeKI: Datos de secuenciamiento de 150 bibliotecas de representación reducidas obtenidas a partir de la digestión con la enzima ApeKI.

**Anexo 3:** Valores de los parámetros considerados para cada PNS de los genes *KRTAPs* que codifican proteínas con alto contenido en glicina y tirosina.

Gen	Posición en el scaffold	Illumina Score	Frecuencia del alelo menor por genoma				Genome Studio
			UNALM	UPCH	PstI_MspI	ApeKI	
<i>KRTAP7-1</i>	12420218	0.787		0.1667			
<i>KRTAP8-1</i>	12436062	0.834	0.08333	0.1667		0.02381	
<i>KRTAP8-1</i>	12436103	1	0.5	0.1667		0.4211	0.421
<i>KRTAP19-1L</i>	12804959	0.858		0.1667			0.007
<i>KRTAP19-3L</i>	12746893	0.685		0.1667			0.03
<i>KRTAP20-2L</i> (LOC116286138)	12616128	1	0.08333				
<i>KRTAP20-2L</i> (LOC116286182)	12651734	0.787	0.5	0.5			
<i>KRTAP20-2L</i> (LOC116276588)	12660986	0.787	0.5	0.5			0
<i>KRTAP20-2L</i> (LOC116276577)	12678801	0.787		0.1667			
<i>KRTAP21-1L</i>	12501902	1	0.08333				
<i>KRTAP21-1L</i>	12502328	1	0.08333				

UNALM: Datos de secuenciamiento de 6 genomas completos de la Universidad Nacional Agraria La Molina.

UPCH: Datos de secuenciamiento de 3 genomas completos de la Universidad Peruana Cayetano Heredia.

PstI\_MspI: Datos de secuenciamiento de 150 bibliotecas de representación reducidas obtenidas a partir de la digestión doble con enzimas PstI\_MspI.

ApeKI: Datos de secuenciamiento de 150 bibliotecas de representación reducidas obtenidas a partir de la digestión con la enzima ApeKI.

**Anexo 4:** Alineamiento de las secuencias de los genes *KRTAP1-1*, *KRTAP3-2* y *KRTAP4-12L*, con la herramienta Clustal Omega.

**-KRTAP1-1**

O.Aries/KAP1-1	ATGGCCTGCTGTTCCACCAGCTTCTGTGGATTTCCCATCTGTTCCACTGGTGGGACCTGT	60
V.Pacos/KAP1-1	ATGGCCTGCTGTTCCACCAGCTTCAACGGATTTCCCATCTGTTCTACTGGTGGGAACTGT	60
	*****	
O.Aries/KAP1-1	GGCTCCAGTCCCTGCCAGCAGACCTGCTGCCAGACCAGCTGCTGCCAGCCAACCTCCATC	120
V.Pacos/KAP1-1	GGCTCCAGCTGCTGCCAGCCAACCTGCTCCAGACCAGCTGCAGCCAGCCAACCTGCTCC	120
	*****	
O.Aries/KAP1-1	CAGACCAGCTGCTGCCAGCCAACCTCCATCCAGACCAGCTGCTGCCAACCGATCTCCATC	180
V.Pacos/KAP1-1	CAGACCAGCTGCTGCCAGCCAACCT-----	145
	*****	
O.Aries/KAP1-1	CAGACCAGCTGCTGCCAGCCAACCTCCATCCAGACCAGCTGCTGCCAGCCAACCTGCCCTC	240
V.Pacos/KAP1-1	-----GCTGCCAGCCAACCTGCTGCCAGACCAGCTGCTGCCAGCCAACCTGCCGC	195
	*****	
O.Aries/KAP1-1	CAGACCAGTGGCTGTGAGACGGGCTGTGGCATTGGTGGCAGCATTGGCTATGGCCAGGTG	300
V.Pacos/KAP1-1	CAGACCAGCAGCTGTGAGACTGGCTGTGGCATTGGTGGTAGCATTGGCTGTGGCCAGGAG	255
	*****	
O.Aries/KAP1-1	GGTAGCAGCGGAGCTGTGAGCAGCCGACAGGTGGTGCCGCCCTGACTGCCGCGTGGAG	360
V.Pacos/KAP1-1	GGTGGCAGCGGAGCTGTGAGCTGCCGACAGGTGGTGCCACCCTGACTGCCGCGTGGAG	315
	***	
O.Aries/KAP1-1	GGCACCAGCCTGCCTCCCTGCTGCGTGGTGGAGTGCACACCCCGTCTGCTGCCAGCTG	420
V.Pacos/KAP1-1	GGCACCTGCCTGCCTCCCTGCTGCGTGGTGGAGTGCACCCACCAACCTGCTGCCAGCTG	375
	*****	
O.Aries/KAP1-1	TACTATGCCAGGCCCTCTGCTGCCGCCATCCTACTGTGGACAGTCTGCTGCCGCCA	480
V.Pacos/KAP1-1	TACCTGCCCCAGGCCCTCTGCTGCCGCCATCCTACTGTGGACAGTCTGCTGCCGTCCA	435
	***	
O.Aries/KAP1-1	GCCTGCTGCTGCCAGCCACCTGCACCTGAGCCCGTCTGTGAGCCAGCTGCTGTGAGCCC	540
V.Pacos/KAP1-1	GCCTGCTGCTC-----TGCTGTGAGCCACCTGCTGCGAGCCC	474
	*****	
O.Aries/KAP1-1	ACCTGCTGA	549
V.Pacos/KAP1-1	ACCTGCTAA	483
	*****	

**-KRTAP3-2**

O.Aries/KAP3-2	ATGGCTTGCTGCGTCCCGCTGCTGCAGCGTCCGCACTGTTCTGCCACCACCATCTGC	60
V.Pacos/KAP3-2	ATGGCTTGCTATGTACCTGCTCCTGCAGCGTCCGCACTGTTCTGCCACCACCATCTGC	60
	*****	
O.Aries/KAP3-2	TCCTCTGACAAATCTGTGCGTGGAGTCTGCTGCCAGCACCTGCCACACGACATC	120
V.Pacos/KAP3-2	TCCTATGACAAATCTGTGAGTGGAGTCTGCTGCCAGCACCTGCCAGCACGATC	120
	***	
O.Aries/KAP3-2	AGCCTCCTCAGCCCACTTGTGTGACAACTCCCCGTGCCCTGCTATGTGCTGACACC	180
V.Pacos/KAP3-2	AGTCTCCTCAGCCTACCTGCTGCGACTGCTGCCACCCTGCTGTGAGCCTGACACC	180
	**	
O.Aries/KAP3-2	TATGTGCCAACTTGTCTTCTGCTCAACTCTCCACCCCACTCCTGGACTGAGCGGGATC	240
V.Pacos/KAP3-2	TACGTGCCAACTTGTGGTGTCTCAACTCTCCCAACCGACTCCCGGACTGAGCGGGATC	240
	**	
O.Aries/KAP3-2	AACCTGACGACCTTCATTACGCTGGCTGTGAAAATGTCTGCGAGCCCCGCTGTAA	297
V.Pacos/KAP3-2	AACCTGATGACCTACGTTACGCTGGCTGTGAAAATGCTGCGAGCCCCGCTGTAA	297
	*****	

**-KRTAP4-12L**

0.Aries/KAP4-12	CTCCACCCCTGACACCATGGTCAGCTCCTGTTGTGGCTCCGTC TGCTCTGACCAGAGCT	60
V.Pacos/KAP4-12	CTCCACCC TCTGACACCATGGTCAGCTCCTGCTGTGGCTCCGTC TGCTCTGACCAGGGCT	60
	*****	
0.Aries/KAP4-12	GCGGCCGAAGTCTCTGCCAGGAGACC TGCTGCCGCCAGC TGCTGCCAGAC CACCTGCT	120
V.Pacos/KAP4-12	GTGGTCAAGGCCTCTGTCAGGAGACC TGCTGCCGCCAGC TGCTGCCAGAC CACCTGCT	120
	*****	
0.Aries/KAP4-12	GCAGGACCACC TGCTACCGCCC CAGCTGTGGTGTGTC CAGCTGC TGCCGCCCGTCTGCT	180
V.Pacos/KAP4-12	GTAGGACCACC TGCTACCGCCC CAGATGTTGTGAGTCCATCTGC -----	164
	*****	
0.Aries/KAP4-12	GCCAGCCCACC TGCCCTCGCCC CACC TGCTGCATCTCTAGC TGC TACCGCCCCTCTGCT	240
V.Pacos/KAP4-12	-----	164
0.Aries/KAP4-12	GTGT TACC CGC TGTGGCTCCAGCTGCTACAGGC CTACCGCTGCATCTCCAGCTGCAGGC	300
V.Pacos/KAP4-12	-----TG CAGAC	171
	*****	
0.Aries/KAP4-12	CCCAGTGC TGC CAGCTGTGTGCTGCCAGC CCACCTGCCCTCCCCTACCTGCTGCATCT	360
V.Pacos/KAP4-12	CCCAGTGC TGC CAGTCTGTGTGCTGCCAGC CCACCTGCTGCCCGCCAGATACTGCATTT	231
	*****	
0.Aries/KAP4-12	CTAGTGC TACCGCCCTTCAGCTGTGGTCCAGCTGTGGC TCCAGCTGCTGCAGGCCATA	420
V.Pacos/KAP4-12	CCAGTTGCTGCCGCCCTTTCCACTGTGGTCCAGGTGTGGT TCCAGCTGCTGCAGGCCCA	291
	*****	
0.Aries/KAP4-12	CATGCTGCATCTCCAGCTGCTGTAGGCCCAAGT GCTGCCAGCCTGTGTGCTGCCAGCCCA	480
V.Pacos/KAP4-12	CCTGCTGCATCTCTAGCTGCTGTCCAGCTGCTACCAGACCACGTGCTGC-----	344
	*****	
0.Aries/KAP4-12	CCTGCCCTCGC CCCACCTGTTGCATCTCTAGCTGCTACCGCCC TCCAGCTGTGGGTCCA	540
V.Pacos/KAP4-12	-----CGC CCCAGATGCTGTGTGCCAGCTCTGCGGCC-----	379
	*****	
0.Aries/KAP4-12	GCTGTGGCTCCAGCTGCTGCAGGCCTACCTGCTGCATCTCCA--GCTGCAGGCCCGGT	597
V.Pacos/KAP4-12	-----CAGATGCTGCCAGACCACCTGCTGCCGTACCACCGCTGTCGCCCATCT	429
	*****	
0.Aries/KAP4-12	GTTGCCAGTCTGTGTGCTGCCAGCCAGCTGCCCGGCATCTCCAGCTGCTGCCGCCCT	657
V.Pacos/KAP4-12	GCTGCCGCCATCTGCTGCCGC-----	452
	*****	
0.Aries/KAP4-12	CTTGCTATAGCTCCAGCTGCTGCCGCCAGCTGCTGCCCTGCCCGCCAGTGTGTGGCCGGG	717
V.Pacos/KAP4-12	-----CCCATCTGCTGCCGCCCATCTGCTGAGGATCCCTCT---TTTGTGAA	498
	*****	
0.Aries/KAP4-12	TCTCCTGCCCACACCACTTGCTATCGCCCCACCTGTGTATCTCCACCTGCCCGGCCCGG	777
V.Pacos/KAP4-12	CTTCATTCTCACCACCAGCCC TGAGCCAA TCACCATCATCCAG--TGTATCAGTC--T	554
	*****	
0.Aries/KAP4-12	TGTGCTGTCCC TCCTCTTGCTGCTGAGCCC GCTGCC TGTGACCACTGCTCCGTTTACC	837
V.Pacos/KAP4-12	--TGTCTCATT--TCCTTTGTC CATAGCACCTGGCT-----TGT TTTCA	596
	*****	
0.Aries/KAP4-12	CCTGTCCC CATAGATGCAGACCCTCTTTGGGGTTGATGTTGCTCAGA-----TAAACA	891
V.Pacos/KAP4-12	TCTGTACATATGTTTATTCA--TCCCATTGGCT---ATGCCGCCAAGCAACTGTTAAGA	651
	*****	
0.Aries/KAP4-12	GGGCTTCATAATTGAAGCAGATGGATAAGGCATGGAGCATTCTCCAGTAACAGATTT--	949
V.Pacos/KAP4-12	---AGCGCAATTTAAACT---AAATACCACACAGAAC--CTC-----CAAATTCTG	695
	*****	
0.Aries/KAP4-12	TCCCCTTT CACGTGCAAT TGGTAACTCA---TTCTCTAA-----	988
V.Pacos/KAP4-12	TGCC TATTACC GTGGCATT CGGTCAGATGACCAAACGCTACCC TAGTTCATAGGGACAG	755
	*****	
0.Aries/KAP4-12	-CTTCTTTCC TAAGACTTTCTA-----TAACTGATCTATT-----	1026
V.Pacos/KAP4-12	GCTCCTAATCCATGATTGCTGCACATGGCTTGGCCTTCATGATGACTTACGTGTGTTAG	815
	*****	
0.Aries/KAP4-12	-----CTGTACCAAAAATAAGTGATGCTTTC TGTATTAATAAAA	1063
V.Pacos/KAP4-12	GATAAATATGTCTC--AATAAA TACTT-----GGTATCAGGAA	851
	*****	

**Anexo 5:** Alineamiento de la secuencia del gen *KRTAP5-4* con la herramienta Clustal Omega.

O. Aries/KAP5-4	CACTGAGTCCC-----TTCTCTCTCCACCTGCTCTGACC-TACTC-CACCCCTC	51
V. Pacos/KAP5-4	CACCCAGAACCCCGAGGATGGTCTCTACCGC--AGGACTGAATTTAATCACACCTGC	57
	*** ** * * * ***** * * **** * * * * * * * * *	
O. Aries/KAP5-4	AA----CCCACCAGAACCATGGGCT--GCTCTGGTGTTCGGAGGCTGCGGCTCCAGC	104
V. Pacos/KAP5-4	AAAATTCCAGTAGATCCGAGGTAAGGGAACCTGGATG---CAAGGCTGTCACAGG	113
	** **** *	
O. Aries/KAP5-4	TGTGGGGGCTGCGGCTCGTGTGGGGGCTGCAGCTCCAGCTGCTGTGTGCCTGTCTGCTGC	164
V. Pacos/KAP5-4	-----TACAGTGCAGTCCACACC--AAG--GCTTGGGACCAAGT-----	147
	* *	
O. Aries/KAP5-4	TGCAAGCCGTGTGCTGTGTGCCAGCCTGCTCCTGCTCCAGCTGTGGCAAAGGGGGC	224
V. Pacos/KAP5-4	TGACAGG--GACTGTGGGAAGGACAGACTGAA---ATCTAGCA-TGG-----A	190
	** *	
O. Aries/KAP5-4	TGCGGCTCCTGTGGGGCTCCAAGGGGGCTGCAGCTCCTGTGGGGGCTAAGGGGAGC	284
V. Pacos/KAP5-4	TGA-----GGAAGA----CGTCAAATCTGGAACCTAAG-AAAGT	224
	** *	
O. Aries/KAP5-4	TGT-----GGCTCATGTGGAGG-----CTG---T-	305
V. Pacos/KAP5-4	TGTAGTGACACAGGCACTAGGAAGTCTGAAAAAGAAACAAGCAAAAACCTGATCCTG	284
	*** **** *	
O. Aries/KAP5-4	----GGCTCCAGCTGCTGCAAGCCGCTGT-----GCTGCTG	337
V. Pacos/KAP5-4	GCAACAGTCCAG---ACACAAGGAGGTGAGTGGTGCCAGCACACACCTGGCAGGTG	341
	***** *	
O. Aries/KAP5-4	TGTGCCGCTGTG---CTCCTGCTCCAGCTGTGGCAAAGGGGGCTGTGGCTCCAGCTGTG	393
V. Pacos/KAP5-4	TGTAGTTTCTGGCCCGGAGGCTCCACACTGA-ACAACCCCTCTCTCCATCTGC-	399
	*** *	
O. Aries/KAP5-4	GGGGCTCCAAGGGGGCTGCGGCTCCTGTGGGGGCTAAGGGGGGCTGTGGCTCGTGTG	453
V. Pacos/KAP5-4	----TCCTTGG---ACTGCCCCCCACC---CCCACCCAGCAGAACCCTG-----G	440
	* *	
O. Aries/KAP5-4	GGGGCTGTGGCTCTGGCTGCGGCTCCAGCTGCTGTGTGCCCGTGTGCTGTGTGCCAG	513
V. Pacos/KAP5-4	GCCG-----CTGTGGTCCAGCAGT-----GGCGGCTG	468
	* *	
O. Aries/KAP5-4	CCTGCTCCTGCTCCAGCTGTGGCAAAGGGGGCTGCGGCTCCTGTGGCTCCTCCAGTCCA	573
V. Pacos/KAP5-4	TCTGCTCCTGCTCCAGCTGTGGCAAAGGGGGCTGTGGCAGTGTGGCTCCTCCAG---	524
	***** *	
O. Aries/KAP5-4	GCTGCTGCAGACCCCTGCTGCTCCAGTCCAGTGTGTGTGCCCGTGTGCTGCTGCTGCGCA	633
V. Pacos/KAP5-4	-----TCCAGTGTGTGTGCCCATTTGCTGCGCAGGACA	558
	***** *	
O. Aries/KAP5-4	AGATCTGATGTAACGGTCACTCCAGCCTTCCAGCTAACCCTTCCAGCCGCGCATCT-GT	692
V. Pacos/KAP5-4	AGCCAGATGTAGTGCCAGTCCGGTATTACAGTGTCTTCCACTTGCAGCAGCTCAT	618
	** *	
O. Aries/KAP5-4	GAGGCTGTGGCTGTGCTTCTCCTGGACACATCTTGTCTGTGGGCTCTTCCCTGC-	751
V. Pacos/KAP5-4	GAGGCCATGGTCTTGTCACTCCCTGAACACTTCTTGCCTGCACACTCTTCTTCA	678
	***** *	
O. Aries/KAP5-4	ATGGCCTGACTCCAGC--AGGAAGCACCATTACATGAGGAGCCACGGCGTATCCCCAT	809
V. Pacos/KAP5-4	CTGTCTGTGTCCAGCACAGGAAACACCATTACATGAGTGGAATGGCTCTGTCTTGT	738
	** *	
O. Aries/KAP5-4	CCCTCTAGGATGTCATTATGAGAGGAGGGCCGTGCTAATAGTCTCCAAACACCAAG	869
V. Pacos/KAP5-4	CCCTCCAGATACAATACCCAGAAGAAGGCAATGTTCTAACAGTC-TCTAAAAGCAAAG	797
	***** *	
O. Aries/KAP5-4	TGGCCCCACCTGTG---CTGACTTCTAAGGGCCATCATGTGCCACAACACAGGCGCTG	926
V. Pacos/KAP5-4	TGGCCACTGCATGTCTTTGACCTTCTAAGGGCCACTAGCTGNNNNNNNNNNNNN--	855
	***** *	
O. Aries/KAP5-4	ACCCTCACGTGAAGGGTCTGGCCAAGGACAGCCAGCTCCTGGCTTCTAATAAGAAAAGA	986
V. Pacos/KAP5-4	-----NNNNNTGGCCGGTCGAGAATCC--CTCCGGCTCCTCATGAGGAAAAGA	905
	* *	
O. Aries/KAP5-4	TGCTGCTTTGATCACATCCAAGTTTCCTTCTCCAGTGAAGCCAGTCCAGCAAGACTCCC	1046
V. Pacos/KAP5-4	TGCTGCTCGGATCACATCCACGCTCCTCCCCGACAAGGCTAGCCCCGAGCGCCCC	965
	***** *	
O. Aries/KAP5-4	TTCTCACAGCAGGTTTCTATACTTGTGGTAAATAACAATAAAATGCCCTGTCAATAA	1106
V. Pacos/KAP5-4	TGCCTGATAGCAGGGTCTACACTGTCTGCTAAAGTACGACAAAATGCTTTTCAATAA	1025
	* *	
O. Aries/KAP5-4	AA 1108	
V. Pacos/KAP5-4	AA 1027	
	**	



**Anexo 6:** Alineamiento de la secuencia del gen *KRTAP9-7L* (LOC114116849) con la herramienta Clustal Omega.

O.Aries/KAP9-7	ATGACTCACTCCTGCTGCTCCCCGTGCTGCCAGCCACCTGCTGTGAGTCCAGCTGCTGC	60
V.Pacos/KAP9-7	NNNNCCCCTTGTGCCAGCCACCTGCTGCCAGCCACCTGCTGTGAGCCAGCTGCTGC * * * **** * * * *****	60
O.Aries/KAP9-7	CAGCCTTGCTGTGCCTATCAGCT-----GTCAGACCACCTGCTGCAGGACCACC	111
V.Pacos/KAP9-7	CAGCCTTGCTGCCCCAAACTTGCTATGAAGTACTCAGACCACCTGCTGCAGCACCACC ***** * * * * * *****	120
O.Aries/KAP9-7	TGCCTCAAGCCTGTTTGTGCGACCACCTGCTGCCAGCCACCTGCTGTGAGTCCAGCTGC	171
V.Pacos/KAP9-7	TGCCGCAAACCTACTTGTGTGACCACCTGCTGCCAGCCACCTGCTGTGAGCCAGCAGC **** * * * * * *****	180
O.Aries/KAP9-7	TCTCAGCCCTCCTGCCCCAAACTTGCTGTCAA--ATCACTGAAACCACCTGCTGCAAA	228
V.Pacos/KAP9-7	TGTGGAC-----AAACCTGCTGCCAGCCTACCTGTGTAACCCAGCTGTTGCAGC * * * * * *****	228
O.Aries/KAP9-7	CCTACTTGTGTGACCAGCTGCTGTGAGCCACCTGCTGTGCATCCAGCAGCTGTGGACAA	288
V.Pacos/KAP9-7	CCTTC-----CTGCTGCCAGTCCACTTGCTGTGAGTCTAGCAGCTGCGGACAA * * * * * *****	276
O.Aries/KAP9-7	CCCTGCGGTGGTTCTAACTGCTGTGAGCCAGCTTCTGTGCACCCGTGTACTGCCACAGA	348
V.Pacos/KAP9-7	AACTGCGGTGGGTCCAGCTGCTGCCAGCCAGCTAGCTGTGCACCCGTGTACTGCCACCGA ***** * * * * * *****	336
O.Aries/KAP9-7	ACCTGTACCACCCACATGCTGCTGCCTGCCTGGCTGCCAAGCCAGAGCTGTGGATCC	408
V.Pacos/KAP9-7	ACCTGTACCACCCACATGCTGCTGCCTGCCTGGGTGCCTGGCCAGAGCTGTGGATCC ***** * * * * * *****	396
O.Aries/KAP9-7	AGCTGTGCCAGCCTTGACAGTGCCTGTCTGCTGTGAGCCACCTGCTGTGCGGACCACC	468
V.Pacos/KAP9-7	AGCTGTGCCAGCCTCGCTGCCGCCCTGTCTGCTGTGAGCCACCTG----- ***** * * * * * *****	443
O.Aries/KAP9-7	CGCTGCGCCCGAGCTGTGTGTCCAGCTGCTGCCTGCCTTCTGCTGTGATCACCTTGC	528
V.Pacos/KAP9-7	--CTGCCGCCCTAGCTGTGTGTCCAGCTGCTGCCAGCCCTCCT---GCTGATCACCTCAC ***** * * * * * *****	498
O.Aries/KAP9-7	CAAGCATCCACTGTCCCTATACAGCACCTTCTATTAATTGACTTGCTACCTG-GGGACAA	587
V.Pacos/KAP9-7	CAAGA-GCCATCCCCTGCATCCAACACAATCTGTCAACTGAGTTGCCGTTTTGGGGCAA **** * * * * * *****	557
O.Aries/KAP9-7	AACAACATCACTGCCTAGTCTTTGCTGACAACCTGACGTGCTCGACCAGAATTCAGTTAC	647
V.Pacos/KAP9-7	AT----TCACTTC-TTGGGTTTGCAATTACCAGCATGCTCTCCAGAACTTCTGTTAC * * * * * *****	611
O.Aries/KAP9-7	TCATCTACTTGCTTAAGAGCTTGTGAAGTACCTTGACGGAGTGCAGGGGGCTTCCCTCAA	707
V.Pacos/KAP9-7	TCATCTGCTTGTAA-AGCTTGTGAATCAGCTTGTGGAGTGCAGAGGACTTCCCCCAA ***** * * * * * *****	670
O.Aries/KAP9-7	TTGTCTTCTCTGCTGTAGGAGGACCATGTGCCACTTTCCTGCATCTGTGATAAGGAAGT	767
V.Pacos/KAP9-7	TTCTCTTCTCTCGTGTAGGTGGATGATGAGCCAGCTTGTGATCCTTGATATGG-AGT ** ***** * * * * * *****	729
O.Aries/KAP9-7	ATGTCTTGGTTTTGACTCTGAAAATAGGATTGACCATCTAGTCTTCTGAGTAACCTAAG	827
V.Pacos/KAP9-7	TTGTCTGGGCTTTGACTCTGAAGTCAGG---TCCACGCAGTCCCTCTGAGTGACTAAGG ***** * * * * * *****	785
O.Aries/KAP9-7	AAAACCAGATCTCAGAATGTTTCTATAGGTTTCTATAGCCTATCAT-AAACCTATAATC	886
V.Pacos/KAP9-7	AAAACATAATCTTGTAACTGTTTACAGGTTTCTGCAACTGATCAATATTATTATAATC ***** * * * * * *****	845
O.Aries/KAP9-7	ATGTTTTCTTTTTGCTACCCCTAGTCTTGTATCATGCTTCTTCTTTTACATTCAC	946
V.Pacos/KAP9-7	ATATTTCTTTATCAATGACTCTGTGGCTCTCACAGCTTCTTGTCTTTTATGATCAC ** ***** * * * * * *****	905
O.Aries/KAP9-7	TTTGAGTTTTA--CCTGTATG-----TTTCTCAATAAATCGTGTACCACAA	990
V.Pacos/KAP9-7	TTTGAGTGTCTTCTGGATACAGGAGTCTCCCTA-----TGTGTCTCCCAA ***** * * * * * *****	953

**Anexo 7:** Alineamiento de las secuencias del gen *KRTAP13-2L* y las dos copias del gen *KRTAP10-7L* (LOC114110483 y LOC101104456) , con la herramienta Clustal Omega.

**- KRTAP13-2L**

O.Aries/KAP13-2	GTGGCCTCCATGTAAGTCAAGTCTGATCTCACATCTCTGTCAACAACATGCTCTACAAATG	60
V.Pacos/KAP13-2	GTCTCCTCAGTGTAAGTCAAGTCTGAGCTCACACCTCCTGTCA--CCATGCTCTACAAGT	57
	** ** ** ** *	
O.Aries/KAP13-2	CTGTTCTGAAAACCTCTCCTCCCGCTCCCTCGGGTCCACCTGGGCTACCCAGGCTCCTC	120
V.Pacos/KAP13-2	CTGCTCTGAAAACCTCTCCTCCACTCCCTTGGGGGCCACCTGGCTGCCAGGCTCCTC	117
	** * ** ** *	
O.Aries/KAP13-2	CTGTGGCTCCTCCTCTGGCTCCTCTCCTCAGCAACCTGGTCTACAGGACTGACCTCTG	180
V.Pacos/KAP13-2	CTGTGGCTC-----TTCCTACCCAGCAACCTGGTCTACACCACGGACTCCG	165
	*****	
O.Aries/KAP13-2	TTGTCCCAGGATCTGCCGGCCACCCTGTGGTGTCAAGCCCTGTCAAGACTCTCGTA	240
V.Pacos/KAP13-2	CTCTCGCAGCACCTGCCAGCCCTCCTGTGTGGTGTCCAGCCCTGCCAGAGTCTCGTA	225
	* * * * *	
O.Aries/KAP13-2	CTGCCAAGGACCTCCACTCTGCAAGTCCCTCCAGATAACTTTTCTGGTTCTCCGGG	300
V.Pacos/KAP13-2	CCGCCTGAGGACCTCCAGCTCCAGTCCCTGCCAGCAACTTATGCTGGTCTCTGGG	285
	* ** * * * *	
O.Aries/KAP13-2	ATGTGGGTCTATTAGAGGCTGCTCCCTGGGTTATGGATCTAGAAGTCTACTCTTGGG	360
V.Pacos/KAP13-2	CTGCGGGTCTAGCAGAGGCTGCTCCCTGGGTTCTGGATCTAGAAGTCTGCTCCCTGGG	345
	** * * * * *	
O.Aries/KAP13-2	CTGGGGATCCAGA-----GTGGGCTGTGGATCCAGAGTCTTCAGACTGCTGGG	408
V.Pacos/KAP13-2	CTATTGATACAGAAGTTCCCTTATTGGGATGGGGATCCGGTGTCTCAGACCCCTGAG	405
	** * * * * *	
O.Aries/KAP13-2	TTATGGAATTCATGGCTTCCCTTCCCTGAGGTGTGGATCCAGGTTCTACCACCCAGCCT-	467
V.Pacos/KAP13-2	TTTTGGAGCCCATGCTTCCCTTCCCTGAGCTATGGATCCAGACTTTGTCTTCCAAACAAA	465
	** * * * *	
O.Aries/KAP13-2	-ACTTGGCTTCTAGGAGCTGCCAGTCTTCTTGTCTATAGGCAATCTGTAGATCAAACTTC	526
V.Pacos/KAP13-2	AGTTTTGGCTTCCAGGAGCTGCCAGCCTTCTTGTCTATGGCCAGTTTATAGATCAGCTTTC	525
	*****	
O.Aries/KAP13-2	TGTAGATCAACTTGTGAATTTCCAGAACTTCTGAGCAAAACGTTTCAGTCTCTACTTAC	586
V.Pacos/KAP13-2	TGTAATCAGCTTGTGAGTTTCTGGACTTCTGAGCAAAATGTTTCAGTCTCTTCTTAG	585
	*** * * * *	
O.Aries/KAP13-2	A---GCTACCATAGCCCCTTTCCAGAAATGTTAGCTAAGCCCTCTCACCATCAGCATCTCA	643
V.Pacos/KAP13-2	AACTGCTATCATAGCCTTTCCAGCAATGTTAGCTAATCCATCTCACCATCAGCTCTTCA	645
	* * * * *	
O.Aries/KAP13-2	TCCTTTCTCTAGTGTCAACAC---TGACTACTAATCCTTTCA---GTGAAATTA--	693
V.Pacos/KAP13-2	CCCTTTCTCAGTCAATCAAGCACCAGGCTGACAGACTAGTTCTTCAAATGGTAAATTAAT	705
	*****	
O.Aries/KAP13-2	--ATAACCAAAATATTGAGTTCAA-ATCTGTAATTTTGTGAAACAATTGAATCAGAATT	750
V.Pacos/KAP13-2	GAATGACCAAAAGTATTGAGTTCAATGCTGTAATTTGTGAAATAGTTAATTCAGAATT	765
	** * * * * *	
O.Aries/KAP13-2	ACATAAACTTCTTTACAAAATTCATGGAAATCCAGCTTTGCTTACCTAATAAAATTCAT	810
V.Pacos/KAP13-2	ACATAAATTTCTGTATACAATATGTAGAAATTCCTTTTGCTTACCTAATAAAATTCAT	825
	***** * * * * *	
O.Aries/KAP13-2	TCACT	815
V.Pacos/KAP13-2	TCACC	830
	***	

**- KRTAP10-7L (LOC114110483)**

O. Aries/KAP10-7 V. Pacos/KAP10-7	CACCCACACACGAAACCTCCTCCAGCCACCCGCTCCACCATGGCAGCCTCCACC CACTACTCACACATCACCTCCTACAGCCACCCGCCCCGCATGGCCGCGCCGCC *** **	60 60
O. Aries/KAP10-7 V. Pacos/KAP10-7	CTGTCCATCTGCTCCAGCGACTGAGCTATGC----- CTGTCCACTGCTCCGGCGACTGAGCTACGGCAGCCGGGCTGCCTGGCCGGTCCCTGT *****	92 120
O. Aries/KAP10-7 V. Pacos/KAP10-7	-----CTGTC CAGAGAGCTGCTGCGAGCC GACTCTGCACCGCTCCTCCGGCAGGTGAGACTGCC CAGAGAGCTGCTGTGAGCCC ***	117 180
O. Aries/KAP10-7 V. Pacos/KAP10-7	CCCTGCTGCGCTCCAGCTGTTGCGCCCGGCCCTCGCTGACCTCCTCTGTGCCCA CCGTGCTGTGAGCCCGCTGCTGCGCCCGGCCCTGCCTGAGCCTCCTCTGTGCTCA ** **	177 240
O. Aries/KAP10-7 V. Pacos/KAP10-7	GTGAGCTGCGAGTCCAGCCCTGCTGCCAGGCAGCTGCAGCAGCTCCTGCCAACCTCG GCGAGCTGTGAGCCCTGCCCTGC--CCATCTGCTGCACAGCTCCTGCACGGCTGTG * **	237 297
O. Aries/KAP10-7 V. Pacos/KAP10-7	TGTGTCAGCAGTCTAGCTGCCAGCCCTCCAGCTGCACCTCCTCCCCGCCAGCAGGCC TGTGCCCGCAGTCTAGCTGCCAGCCCTCCGTTGACTCCTCCCCGCTGCCAGCAGGAC *****	297 357
O. Aries/KAP10-7 V. Pacos/KAP10-7	TGTGTGAGCCTGCTGCTGCAGGCCATCTGCTGCACACTGTCTGCTGCAGGCCGTGC TGTGTGAGCCCGTGTGCTGCAGGCTGTGCTGCAGGCCCGTGTGCTGCACG----- *****	357 411
O. Aries/KAP10-7 V. Pacos/KAP10-7	TGTGTGAGGCTCCTCCCTGTTCTACCTCCTCATGTTGCCAGCAGTCCAGCTGCCAGTCC -----	417 411
O. Aries/KAP10-7 V. Pacos/KAP10-7	TCCTGCTGCACCTCCTCCCCGCCAGCAGGCTGCTGTGAGCCGCTGCTGTGAGCCC ----- ***	477 414
O. Aries/KAP10-7 V. Pacos/KAP10-7	GCCTGCTGCAGGCTGTCTGCTGCAGGCCGTCTGCTGTACACTGTCTGCTGCACACT GTGTGCTGCAGGCTGTCTGCTGCAGGCCGTGTG----- * *****	537 449
O. Aries/KAP10-7 V. Pacos/KAP10-7	GTCTGCTGTGAGGCCCTCCCTGCTCAACCTCCTCATGTTGCCAGCAGTCTAGCTGCAGG -----	597 449
O. Aries/KAP10-7 V. Pacos/KAP10-7	CCCTCCTGCTGCACCTCCTCCCCGCCAGCAGGCTGCTGTGAGCCATCTGCTGCAGG -----CTGCACG **** *	657 456
O. Aries/KAP10-7 V. Pacos/KAP10-7	CCTGTCTGCTGCAGGCCCGTGTGCTGTGAGGCTTCCCCCTGCTCAGCCCTCATCTGC CCTGTCTGCTGCAGGCCCGTGTGCTGTGAGCCACCCCTGC-----CCCTGCTCTGC *****	717 510
O. Aries/KAP10-7 V. Pacos/KAP10-7	TGCAGACACTCCTCCTGCTGCCCTCCTGCGCCCGTGTGCCGCCCGCCCTGCTGC TGCAGACCTCCTCCTCCGCTGCTCCTGCTGCGCCCGTGTGCAGACCCGATGCTGT *****	777 570
O. Aries/KAP10-7 V. Pacos/KAP10-7	GTGCCATTTCCTCCTGCCAGTCCAGCTGCTGCCGCCGGCCCTCCTGTGTCCCTGCTC GCCCCGCCCCCTCCTGCCAGCCAGCTGCTGCCGCCGGCCCTCCTGCTGCTCCTGCTC * **	837 630
O. Aries/KAP10-7 V. Pacos/KAP10-7	TGCCAGCTGCATGCTCCCGCCAGCCTGCTATGCTCCCATCTCGGCCCCAGAGTCTGCTGC TGCCGCCCTGCATGCCCTGCCCACTGCTGTGCTCCCAACTCGGCC TAGAAGCCCTGC ***	897 690
O. Aries/KAP10-7 V. Pacos/KAP10-7	TGCTGACTGGCATGTCACTTGGGGCCAGCCAGCCTCAGGTCCACC TG TGACTCGGTC TGCTGACCGGC-CGTCCCCCAGGGCCAGCCGGGCTCAGTTCCACCAGTGACTTGGCC *****	957 749
O. Aries/KAP10-7 V. Pacos/KAP10-7	CTCCAGCCT-CCCTCTGCTGACTGGGCTGGTGGCTGCCCTCAC TCAAGACAGGGT C--CCAGCTTCCCTCAGTCCGACTGAGTTATGACTGTCCAC TGAAGACATGTC * *****	1016 807
O. Aries/KAP10-7 V. Pacos/KAP10-7	CCCCGATGCTCCACTGCTCCTGAC-TGACTGACCTTCACTCCTGCTCCAGGAACCC-- CTGCTATGATCCATTGCTCCTGACTGACTGACTCCTGACTGCTCCAGGACCCCTC * * **	1073 867
O. Aries/KAP10-7 V. Pacos/KAP10-7	AACCTC-----ACGG-----CTCCAGGAATTGCC TCCAGCTGTCTTGGGTC GACTTGTGCTGCCCGAGGCTCTTGGCTCCA--GGTGACCCACCTGCTTGGGTC ****	1116 924
O. Aries/KAP10-7 V. Pacos/KAP10-7	ATTGCCCCCTGCCCTCCAGGTCTCTGTCTGGGCTCACTTGGCCTCGACTTCCAGCCTT ACCTGTCCC---CTCCCATTTCTGCTGGTCACTTGGCCTCTCTTCCAACCT- * ****	1176 980
O. Aries/KAP10-7 V. Pacos/KAP10-7	GGGTGCATCCCCTGGGCACCGGATAAATAAACAAGTGTGTGCTGCTGAA GTCAGCATTTCTCAGGCACCTGGAGGAATAAAC TCACTTACCACCTCAT * ****	1227 1031

**- KRTAP10-7L (LOC101104456)**

O. Aries/KAP10-7	TCCACTCCCT-----CACACTCACACCT-	25
V. Pacos/KAP10-7	CCCACCAC TCACC ACTCACAGCCCGTCCAGCTACC TCCGGGCAAAAGCCACCAT	60
	**** * **	
O. Aries/KAP10-7	-----CCTCCAGCCCCATCACCTCCACCATGGCAGCCTCCACCCTGTGTGTGTGCTC	77
V. Pacos/KAP10-7	GGCTGACGC CAGCTGCTCCAGGACTACGTATCTCTGCATCCACCCTGTGTGTGTCT	120
	** *	
O. Aries/KAP10-7	CAGCGACTGAGCTATGACT-----	97
V. Pacos/KAP10-7	CAGTGACCCGAAC TG TGGCAGCCGCTGCTCACC ACTGCCTGCAC TGGCTTCTCCG	180
	** * * * * * * * * * *	
O. Aries/KAP10-7	-----GTCCAGAGAGCTGCTGCGAGCCCCCTGCTGTGCCCCAGCTGCTG	143
V. Pacos/KAP10-7	GCAGGTGACGATTGCCAGGAGAGCTACTGCGAG-----CCCACTGCTG	225
	* *	
O. Aries/KAP10-7	CACCTCGGCCCCCGCTGACCTCCTCTGTGCCCCAGTGAGCTGCGAGTCCAGCCCCTG	203
V. Pacos/KAP10-7	CACCCCGGGCCCTGCTGAGCCTCCTCTGCGCCCGAGCGAGCTGTGAGCCCTGCCCCTG	285
	*** ** *	
O. Aries/KAP10-7	CTGCCAGCCAGCCTGACGAGCTCCTGTCCGGCTCTGTGCTGCCAGCAGTCTAGCTGCCA	263
V. Pacos/KAP10-7	---CCCATCAGCTGCACCAGCTCCTGCACGGCCTTGCTGCCCGCAGTCTAGCTGCCA	342
	** *	
O. Aries/KAP10-7	GCCCACCTGCTGCACCTCC TCCCTGCCAGCAGGCTGCTGTGAGCCCGTCTGCTGCAG	323
V. Pacos/KAP10-7	GCCCTCCTGCTGCACCTCC TCCCTGCCAGCAGGACTGCTGTGCACCCTCTGCTGCAG	402
	*** *	
O. Aries/KAP10-7	GCCCGTCTGCTGTAGGCTGTCTGCTGCACACCTATCTGTTGCACACC TG TGTGCTATGA	383
V. Pacos/KAP10-7	GCCGTCTGCTGCAGCCGCTGCTGCTGCAGGCCGCTGCTGCACCGCTGTCTGCT---	458
	** *	
O. Aries/KAP10-7	AGCCTCCCCCTGTCTGCC TCTCATGTTGCCAGCAGTCCAGCTGTACTCTCCCTCCCTG	443
V. Pacos/KAP10-7	-----	458
	-----	
O. Aries/KAP10-7	CCAGCAGCCACTGTGAGCCTGCTGCTGCAGGCTGCTGCTGCAGGCCGCTGCTGCTG	503
V. Pacos/KAP10-7	-----GCAGGCC TGTGCTGCAGGCCGCTGCTGCAGGCCGCTGCTGCTG	504
	* *	
O. Aries/KAP10-7	TGAGGCC TCCCTGCTGCAAACTCTCTCTCTCCGCTGCCCTCTGCTGCCGGCCCGT	563
V. Pacos/KAP10-7	TGAGCCACCCCTGCC-----	522
	*** * * * * * * * * * *	
O. Aries/KAP10-7	GTCCGCCCCACCTGCTGTGTGCCACTCTCTCTGCCAGCCAGCTGCTGCAGACCCTC	623
V. Pacos/KAP10-7	-----	522
	-----	
O. Aries/KAP10-7	CTCTCTCTGTGCTCCCTCTGTGCCAGCCGCTGCTGCCGCCCCACTGCTGTGTGCCAC	683
V. Pacos/KAP10-7	-----	522
	-----	
O. Aries/KAP10-7	CTCTCTGCCAGCCAGCTGCTGCAGACCCTCTCTCTCTCCGCTGCCCTTTGTGCCG	743
V. Pacos/KAP10-7	-----CTGCTCTGCTGCAGACCCTCC---TCTCCGCTGCCCTGCTCTGCCG	567
	* *	
O. Aries/KAP10-7	GCCCGTGTGCCGCCACCCTGCTGTGTGCCACTCTCTCTGCCAGCCAGCTGTTGCCG	803
V. Pacos/KAP10-7	CCCGTGTGCAGACCCTGCTGTGCCCTGCCCTCTGCCAGCCAGCTGTTGCCG	627
	***** *	
O. Aries/KAP10-7	CCCGCCTCCTCCGTGTCTGCTGCTGCCAGCCGCTGTGCCGCCCTCCACTGCTGCAT	863
V. Pacos/KAP10-7	CCCGCCTCCTGCTGTCTGCTGCTGCCAGCCAGGCTGCTCCGCTGCCCTGCTGTGT	687
	***** *	
O. Aries/KAP10-7	CCCTGCCTCAGCCAGAGCCCTGCTGCTGAC CAGGCACATCCCTCAGGGCCAGCCGGG	923
V. Pacos/KAP10-7	CCCGACTCGGCC TAGAAGTCTGCTGCTGAC CAGGCCGCTCCCTGAGGGCCAGCCGGG	747
	** *	
O. Aries/KAP10-7	CACCATAGGGCAACCCCCACCAGCCTCATGACCGCCCTCAGGGTTTCTCCAGGAGT	983
V. Pacos/KAP10-7	CTCCAGGGAT--CCCTCTCCGCGCAGTGCCTTGGAGCTCTGA-----CCTCTGCGCT	800
	* *	
O. Aries/KAP10-7	TCCAAGCTACATGGACACCCCTTGCCTCTCTGCCAGGAGTGGGGCCCTCAAATC	1043
V. Pacos/KAP10-7	CCTCAGC-----CCATCTACCAGGCTTTGCTCA--GCACAGAGGA	840
	* *	
O. Aries/KAP10-7	CAGATGACCATGCCCGTGTCTACTGCCACTCTCTCTGGGACCCAGTGTCTCAGGCC TGGG	1103
V. Pacos/KAP10-7	AA---GACCAGCTC-----CCAC--CTC-----CCCAGCCAGACCTGGG	876
	* *	
O. Aries/KAP10-7	GACCA---GGTTTCAATCCAGGCC--CCATGGTCTCACCTCGGA-----AGCTGCTTG	1153
V. Pacos/KAP10-7	GCCCAACAGGGTACAGAAACCCTTCTTGACCGTGGCCACAGGAGCCGGGCTGCTCC	936
	* *	
O. Aries/KAP10-7	AAAACGTAATAAATTGACTGCATCTACTGGT	1185
V. Pacos/KAP10-7	ACAC-----CTGCTTCTCGCTCTGGA	958
	* *	

**Anexo 8:** Alineamiento de las secuencias de los genes *KRTAP1-4*, *KRTAP4-11L* y *KRTAP9-3L* con la herramienta Clustal Omega.

**-KRTAP1-4**

C.Hircus/KAP1-4	TGTAGCCAAAACCTCAAAAACCTTCTCTAACAACCCCTCCTCTCAACCCAACTCTGTATACC	60
V.Pacos/KAP1-4	TGAAGCCAAAACCTCAGAACTTCTCTAACAACCAACTCTCAGCCCAACTCTTGATACC	60
	** ***** *	
C.Hircus/KAP1-4	ATGGCCTGCTGTCCACTAGCTTCTGCGGATTTCCATCTGTTCCTACTGGTGGAACTGT	120
V.Pacos/KAP1-4	ATGGCCTGCTGTCCACCAGCTTCAACGGATTTCCATCTGTTCCTACTGGTGGAACTGT	120
	***** *	
C.Hircus/KAP1-4	GGCTCCAGTCCCTGCCAGCCGACCTGTGCGCAGACAGCTGCTGCCAGCCAACCTCCATT	180
V.Pacos/KAP1-4	GGCTCCAGTCCCTGCCAGCCGACCTGTGCGCAGACAGCTGCTGCCAGCCAACCT----	175
	***** *	
C.Hircus/KAP1-4	CAGACCAGTCTGCGCAGCCAACTTCCATCCAGACCAGTCTGCCAGCCAACCTCCATC	240
V.Pacos/KAP1-4	-----GCGGCAGCCAACTTCT-----	193
	* *	
C.Hircus/KAP1-4	CAGACCAGTCTGCCAACCGATCTCCATCCAGACCAGTCTGCCAGCCAACCTGCCTC	300
V.Pacos/KAP1-4	-----CCAGACCAGTCTGCCAGCCAACCTGCCTC	225
	***** *	
C.Hircus/KAP1-4	CAGACCAGTGGTGTGAGACCGGCTGTGGCATTGGTGCAGCATGGCTATGCCAGGTG	360
V.Pacos/KAP1-4	CAGACCAGCAGCTGTGAGACTGGCTGTGGCATTGGTGTAGCATGGCTATGCCAGGTG	285
	***** *	
C.Hircus/KAP1-4	GGCAGCAGCGGAGCTGTGAGCAGCCGACCAAGTGGTCCCGCCCTGACTGCCGCTGGAG	420
V.Pacos/KAP1-4	GGTGGCAGCGGAGCTGTGAGCTGCCGACCAAGTGGTCCCGCCCTGACTGCCGCTGGAG	345
	** ***** *	
C.Hircus/KAP1-4	GGCACCAGCTGCCTCCCTGCTGTGTGGTGGAGTGCACACCCCTGCTGTGCCAGCTG	480
V.Pacos/KAP1-4	GGCACCAGCTGCCTCCCTGCTGTGTGGTGGAGTGCACCCCAACCTGCTGCCAGCTG	405
	***** *	
C.Hircus/KAP1-4	TACTATGCCAGGCCCTCTGCTGCCGCCATCTACTGTGGACAGTCTGCTGCCGCCA	540
V.Pacos/KAP1-4	TACCTCGCCAGGCCCTCTGCTGCCGCCATCTACTGTGGACAGTCTGCTGCCGCCA	465
	** ***** *	
C.Hircus/KAP1-4	GCCTGCTGCTGCCAGCTACCTGCACTGAGCCCATCTGTAGCCACCTGCTGTGAGCCC	600
V.Pacos/KAP1-4	GCCTGCTGCT-----CTGCTGTGAGCCACCTGCTGTGAGCCC	504
	***** *	
C.Hircus/KAP1-4	ACCTGCTGAAAGCAAGGCTGCTGATTACTTAAGAATGAAGGGGAAGCAAAACAACTG	660
V.Pacos/KAP1-4	ACCTGTTAAAGCAAGGTTGCTGATTACTTAAGAATGAAGGGGAAGCAGACTAAATTA	564
	***** *	
C.Hircus/KAP1-4	CGCTGATCTCTGAAGAACGTTCAATTTTTCTGGGTGCTAAGTATCATGAAGTTCTAA	720
V.Pacos/KAP1-4	TGCAGATCTCTGAAGGAGTGTCAATTTTACCTGGATGCTAAGTATCATGCTATTCTAA	624
	** ***** *	
C.Hircus/KAP1-4	GCTGCTCAGAAATGAAAGTCTTAGGACACTTTTGAAATGTTGAAATAGCACACTGAAA	780
V.Pacos/KAP1-4	GCCACTCAGAAATGAAAGCTTAGGACCCCTTGGATGTTTGAATTAGCA- ACTGAAA	683
	** ***** *	
C.Hircus/KAP1-4	CAGGCCTCAACCACGGTGTCTGGATTTCCCTGCCTGCTGGAGAGTACAGACTTGAG	840
V.Pacos/KAP1-4	CATGCATTGGCCACAGTATCTGGATATCCTGCCCTGCTGGAGAGGGCCCAAGAGTGGAG	743
	** *	
C.Hircus/KAP1-4	TAGCTCTTCATTTCAAGTGCATCTCCGATCTCTTCTCTCTTGTG-CCCATATATTT	899
V.Pacos/KAP1-4	TAATCTCTCATTTAAGTGGCATCTGGTCTCTTCTCTCTTGTG-CCCATATATTT	803
	** ***** *	
C.Hircus/KAP1-4	TCAGACCTGCTTTCTCCTCAATATCCTCTGTTAATTTCAAATAAAATATTTAACTTGC	959
V.Pacos/KAP1-4	TCAAACCTCCTTTCTCCTGATAGCCTCTGTAATGTCAAATAAAATGTTTAAATGTGC	863
	***** *	
C.Hircus/KAP1-4	AAAGCAAA 967	
V.Pacos/KAP1-4	AAAGCAAA 871	
	*****	



***-KRTAP4-11L***

C.Hircus/KAP4-11	ATGGTCAGCTCCTGTTGTGGCTCTGTCTGCTCTGACAGAGCTGTGGCCGAAGTCTCTGC	60
V.Pacos/KAP4-11	ATGGTCAGCTCCTGCTGTGGCTCTGTCTGTTCTGAGGAGGGCTGTGGCCAAGGCCTCTGC	60
	***** **	
C.Hircus/KAP4-11	CAGGAGACCTGCTGCCGCCCCAGCTGCTGCCAGACCACCTGCTGCCGGACCACTGCTAC	120
V.Pacos/KAP4-11	CAGGAGACCTGCTGCCGCCCCAGCTGCTGCCAGACCACCTGCTGCCAGAACCACTGCTAC	120
	***** * *****	
C.Hircus/KAP4-11	CGCCCCAGCTGTGGTGTGTCCAGCTGCTGCCGCCCGTCTGCTGCCAGCCCACTGCCCT	180
V.Pacos/KAP4-11	CACCCAGGTGCCATGAGTCCAGCTGCTGCCGCCCACTGCTGCCCTGTGTGTGCTGC	180
	* ***** ** * ***** * * *	
C.Hircus/KAP4-11	CGCCCCACCTGCTGCATCTCTAGTGTCC-----CGCCCTCTGCTGT	225
V.Pacos/KAP4-11	CAGCCCACTGCTGCCGCCCCAGTGTGAGTCCAGCTGCTGCAGACCCCACTGCTGC	240
	* ***** *	
C.Hircus/KAP4-11	GT--TTCCAGCTGTGGTTCAGCTGTACAGGCCGACCACTGCTCATCTCCAGCTGCTGC	282
V.Pacos/KAP4-11	CAGTCTGTGTGCTGCCAGCCCACTGCCGCCCGCCAGGTGCTGTGAGTCCAGCTGCTGC	300
	* *	
C.Hircus/KAP4-11	AGGCCCCAGTGTGCCAGCCTGTGTGCTGCCAGCCCACTGCCCTCGCCCCACTGCTGC	342
V.Pacos/KAP4-11	CGCCCCAGTGTGCCAGTCTGTGTGCTGCCAGCCCACTGCTGCCACCACTGAGTGTGCT	360
	* ***** *	
C.Hircus/KAP4-11	ATCTTAGCTGCTACCGCCCTCCAGCTGTGGGTCCAGC-----	381
V.Pacos/KAP4-11	GAGTCCAGCTGCTGCCGCCCCAGTGTGCCAGTCCGTGTGCTGCCAGCCCACTGCTGC	420
	* *	
C.Hircus/KAP4-11	-----TGTGGCTCCAGCTGCTGCAGGCCTACCTGCTGCATCTCCAGCTGCAGG	429
V.Pacos/KAP4-11	CGCCCCAGGTGTGAGTCCAGCTGCTGCCAGCCAGATGCTGCATCTCCAGCTGCTGC	480
	* *	
C.Hircus/KAP4-11	CCCCGGTGTGCCAGACTGTGTGCTGCCAGCCCGTGTGCTGCCAGCCCACTGCCCCGC	489
V.Pacos/KAP4-11	CGCCCGGATGTTGCATCTCCAGCTGCCAGCCCAAGCTTGCTGATGCTCCAGCAGCTCT	540
	* *	
C.Hircus/KAP4-11	ATCTCAGCTGCTGCCGCCCTCTTGTGTGGTCTAGTGTGCTGCCGCCAGCTGCTGC	549
V.Pacos/KAP4-11	AGTTCAGCTGTACCAACCTCTTGTGCTGCTGCTGCTGCCGCCAGCTGCTGCTGCTGC	600
	* *	
C.Hircus/KAP4-11	CTGCG-CCCAGTGTGCGCGGGTCTCTGCCACCACTTGCTATCGCCCACCTGTGT	608
V.Pacos/KAP4-11	CTGCGCCCAAGTCTGTGGCCAGTCTTCTGCCACCACTTGCTATCGCCCACCTGTGT	660
	***** *	
C.Hircus/KAP4-11	CATCTCCACCTGCCCGGCCCGTGTGAGTGTCCCTCCTTGTGCTGCTGA	657
V.Pacos/KAP4-11	CATCTCCACCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGA	709
	***** *	

***-KRTAP9-3L***

C.Hircus/KAP9-3	ATGACCCCTCGTGCTGCTCCCGCACTGTCAGCCCACTGCTGCAGGACCACTGCTGT	60
V.Pacos/KAP9-3	TGCTGCCCCAACTTCTATGAAGCTACTAGACCACTGCTGCAGCACTGCTGCCG	60
	***** *	
C.Hircus/KAP9-3	GAGTCCAGCTGCTGCAAGCCCTGCTGCCCCCTACATGCTGTCAAACCACTGCTGCAGG	120
V.Pacos/KAP9-3	AAACCTACTTGTGTGACCACTGCTGCCAGCCCACTGCTGTGAGCCCACTGCTGCA	120
	* *	
C.Hircus/KAP9-3	-CCACTGCTGCAAACTATCTGTGTGACCAGCTGTTGCCCGCCCACTGTTGCAGCAA	179
V.Pacos/KAP9-3	CAAACCTGCTGCCAGCTACTG-----TGTAACCACTGTTGCAGCCCT	165
	***** *	
C.Hircus/KAP9-3	CCCTGCTGTGAGCCCACTGCTGTGGTCCAGCAGCTGTGGACAAACCTTCACTGGGTC	239
V.Pacos/KAP9-3	TCCTGCTGCCAGTCCACTTGTGTGAGTCTAGCAGCTGGGACAAACTGCGGTGGGTCC	225
	***** *	
C.Hircus/KAP9-3	AGCTGCAGCCAGTCTTGTGTCAGCCAGCTTCTGTGCACCTGTGTAAGCCAGAAATC	299
V.Pacos/KAP9-3	AGCT-----GCTGCCAGCCAGCTAGCTGTGCACCCGTTACTGCCACCGAACCC	273
	***** *	
C.Hircus/KAP9-3	TGTAACCACTCACGTGCTGCTGCCTGGTGCCAGCCAGAAATGTTGGATCCAGC	359
V.Pacos/KAP9-3	TGTTACCACTCACGTGCTGCTGCCTGGTGCCAGCCAGAAATGTTGGATCCAGC	333
	* *	
C.Hircus/KAP9-3	TGCTGCCAGCCTTGCAGCAGCCCTGCTGCTCTAGACCACTGCTGTAGGACCACTCC	419
V.Pacos/KAP9-3	TGCTGCCAGCCTGCTGCCAGCCCTGCTGCTGCTCTAGACCACTGCTGCCAGCCCTGCTG	392
	***** *	
C.Hircus/KAP9-3	TGCCACCCAGCTTCTGTGCAAGCTGCTGCCAGCCTTCTGCTGCTGA	467
V.Pacos/KAP9-3	-----TGTTCCAGCTGCTGCCAGCCTTCTGCTGCTGA	426
	***** *	

**Anexo 9:** Alineamiento de la secuencia del gen *KRTAP9-8* con la herramienta Clustal Omega.

C. <i>Hircus</i> /KAP9-8	CTCCACCCCTGACACCATGGTCAGCTCCTGTTGGCTCCGTCGCTCTGACCAGAGCT	60
V. <i>Pacos</i> /KAP9-8	CTCCACCCCTGACACCATGGTCAGCTCCTGCTGGCTTCGTCGTTCTGAGGAGGCT	60
	*****	
C. <i>Hircus</i> /KAP9-8	GCGGCCGAAGTCTCTGCCAGGAGACCTGCTGCCGCCAGCTGCTGCCAGACCCTGCT	120
V. <i>Pacos</i> /KAP9-8	GTAGCCAAGGCTCTGCCAGGAGACCTGCTGCCGCCAGCTGCTGCCAGACCCTGCT	120
	* ** * * *	
C. <i>Hircus</i> /KAP9-8	GCAGGACCACCTGCTACCGCCCAGCTGTGGTGTGCCAGCTGCTGCCGCCCGTCTGCT	180
V. <i>Pacos</i> /KAP9-8	GCAGGACCACCTGCTACCGCCCAGCTGTGTGAGTCCAGCTGCTGCCGCCCGTCTGCT	180
	*****	
C. <i>Hircus</i> /KAP9-8	GCCAGCCCACCTGCCCTCGCC-----CCAG	205
V. <i>Pacos</i> /KAP9-8	GCCAGTCCTTGCTGCCAGCCACCCNNNNNNNNNNNNNNNNNNNTGCTGTCAACCTGC	240
	***** * * * * *	
C. <i>Hircus</i> /KAP9-8	CTGCTGCATCTTAGCTGTACCGCCCTCCTGCTGTGTTAC--CCGCTGTGGCTCCAG	262
V. <i>Pacos</i> /KAP9-8	TTGCTGTATCTCCAGCTGCTGCCGCCAGCTGCTGCCAGACCACCTGCTGCAGGACCAC	300
	**** * * * * *	
C. <i>Hircus</i> /KAP9-8	CTGCTACAGGCCTACCGGCTGCATCTCCAGCTGCAGGCC--CCAG--TGCTGCCAGCCTGTG	320
V. <i>Pacos</i> /KAP9-8	CTGCTGCCGCCAGGTGCTGTGAGTCCAGCTGCTGCCGCCAGATGCTGCAT--CTCC	357
	***** * * * * *	
C. <i>Hircus</i> /KAP9-8	TGCTGCCAGCCACCTGCCCTCCCTACCTGCTGCATCTTAGCTGTACCGCCCTTCC	380
V. <i>Pacos</i> /KAP9-8	AGCTGCCAGCCATCTGTGCTGCCAGCCAGATGCTGTGAGTCCAGCTGCTGCCGCCAGCA	417
	***** * * * * *	
C. <i>Hircus</i> /KAP9-8	AGCTGTGGTCCAGCTGTGGCTCCAGCTGCTGCAGGCCTACATGCTGCATCTCCAGCTGC	440
V. <i>Pacos</i> /KAP9-8	TGCTGCATCTCCAGCTGCCAGCCACCTGTTGCCACCCAGATGCTGTGAGTCCAGCTGC	477
	**** * * * * *	
C. <i>Hircus</i> /KAP9-8	TGTAGCCCCAGTGTGCCAGCTGTGTGCTGCCAGCCACCTGCCCTGCCGCCACCTGT	500
V. <i>Pacos</i> /KAP9-8	TGCAGACCCAGTGTGCCAGTCTGTGTGCTGCCAGCCACCTGCTGCCGCCAGCTG-	536
	** * * * * *	
C. <i>Hircus</i> /KAP9-8	TGCATCTTAGCTGTACCGCCCTCCAGCTGTGGTCCAGCTGTGGCTCCAGCTGCTGC	560
V. <i>Pacos</i> /KAP9-8	-----CTG-----TGAGTCCAGTTG----	551
	*** * * * * *	
C. <i>Hircus</i> /KAP9-8	AGGCCTACCTGCTGCATCTCCAGCTGCAGGCCCGGTGTTGCCAGTCTGTGTGCTGCCAG	620
V. <i>Pacos</i> /KAP9-8	-----CTGCAGACCCAGTGTGCCAGTCTGAGTGTGCCAG	588
	***** * * * * *	
C. <i>Hircus</i> /KAP9-8	CCCAGCTGCCCCGCATCTCCAGCTGCTGCCGCCCTCTTGCTATAGCTCCAGCTGCTGC	680
V. <i>Pacos</i> /KAP9-8	CCCAGTGTGCCGCCCTCTGCTGT--GGTTCAGCTCCTGTGGCTCCAGCTGCTGC	645
	**** * * * * *	
C. <i>Hircus</i> /KAP9-8	CGCCCGAGTGTGCTGCCGCCAGTGTGTGGCCGGTCTCCTGCCACACCACTTGCTAT	740
V. <i>Pacos</i> /KAP9-8	CGCACCTGCTTTGCCCTGCACCCAGTCTGTGCCAGTCTCCTGCCACACCACTTGCTAT	705
	*** * * * * *	
C. <i>Hircus</i> /KAP9-8	CGCCCCACCTGTGCATCTCCACCTGCCCCCGCCCGTGTGCTGTCCCTCCTCTTGCTGC	800
V. <i>Pacos</i> /KAP9-8	CGCCCCACCTGTGCATCTCCACCTGCCCCCGCCACAGTGTGCTGTCCCTCCTCTTGCTGC	765
	*****	
C. <i>Hircus</i> /KAP9-8	TGAGCCCGTGCCTGTGACCACTGTCTCCGTTTACCCCTGTCCCATAGATGCAGACCC	860
V. <i>Pacos</i> /KAP9-8	TGA-----TGCCCGTGACCACTCTCTGTTTGC--TCTGTTCTCACAGATGTAGACCC	818
	*** * * * * *	
C. <i>Hircus</i> /KAP9-8	TCCTTTGGGTTGATGTTGCTCAGATAAACAGGGCTT--CATAATTGAAGCAGATGGATAA	919
V. <i>Pacos</i> /KAP9-8	TCCTACTGTGCTGACCCCTT--AGGATACATGGAGCCTGGCTAACACATTAGT-----	869
	**** * * * * *	
C. <i>Hircus</i> /KAP9-8	GGCATGGAG--CATTTCTCCAGTAACAGATTTTCCCTTTACCTGACAATTGGTAAACTC	978
V. <i>Pacos</i> /KAP9-8	GGCTTGACCCATGATTCCAATGAGTCCATT--ACTATCCCACTGACTTAGG--AGCAC	925
	*** * * * * *	
C. <i>Hircus</i> /KAP9-8	ATTCT-----TCTAACTCTTTTCTTA--AGACTTTTCTATAACCTGATCTATTCTT	1028
V. <i>Pacos</i> /KAP9-8	ATTCTGGTTCACTTAAACTCTCCCTTGCTATCCTTTCC-----TTTCAATATCC	977
	***** * * * * *	
C. <i>Hircus</i> /KAP9-8	GTACCAAATAAAGTGATGCTTTCTGT--ATTAATAA	1063
V. <i>Pacos</i> /KAP9-8	GCACCAAATAGAAGTTAATT---TGCAATCCACCA	1009
	* * * * * *	

**Anexo 10:** Alineamiento de las secuencias de los genes *KRTAP10-9L* y *KRTAP16-4* con la herramienta Clustal Omega.

**- KRTAP10-9L**

C.hircus/KAP10-9	CACACACAC-AGTCATACACCCACACACTCAAACCTCCTCCAGCCCCACCGACTCC	59
V.Pacos/KAP10-9	CACACTCCCTCGCACTCACACTCACTCACACACTCACCCTCTACAGCCCCACCGCCCC	60
	***** * * * * * ***** * ***** * ***** * *	
C.hircus/KAP10-9	ACCATGGCAGCCTCCACCTGTCCATCTGCTCCAACGACCTGAGCTAT-----	107
V.Pacos/KAP10-9	GCCATGGCCGCGCCCGCTGTCCACCTGCTCCGGCAGCTAGCTACGGCAGCCGGGTG	120
	***** ** * ***** ***** *****	
C.hircus/KAP10-9	-----GACTGTCCG	116
V.Pacos/KAP10-9	TGCCTGGCCGGTCCCTGTGACTCCTGCACCGGCTCCTCTGGCAGGTGGACGACTGCCCA	180
	***** **	
C.hircus/KAP10-9	GAGAGCTGCTGGAGCCCCATGCTGTGGCCAGCTGCTGCACCCCGCCCCCGCTG	176
V.Pacos/KAP10-9	GAGAGCTGCTGTGAGCCCCGCTGCTGTGAGCCCGCTGCTGCGCCCCGGCCCCCTGCTG	240
	***** ***** ***** * ***** ***** *****	
C.hircus/KAP10-9	ACCCTCTCTGTGCCAGTGTGAGCTGTGAGTCCAGCCCTGCTGCCAGGACCTGCAGC	236
V.Pacos/KAP10-9	AGCTCCTCTGTGCCCTCAGCGAGCTGTGAGCCCTGCCCTGC---CCATCTGCCAGCACC	297
	* ***** * ***** * ***** * * ***** *	
C.hircus/KAP10-9	AGCTCCTGCCAGGCTCATGCTGCCAGCAGTCTAGCTGCCAGCCTCCAGCTGCACCTCC	296
V.Pacos/KAP10-9	AGCTCCTGCCAGGCTGTGCTGCCCGCAGTCTAGCTGCCAGCCTCCTGTGACCTCC	357
	***** * ***** ***** ***** ***** * * ***** *	
C.hircus/KAP10-9	TCCCTGCCAGCAGGCTGCTGTGAGCCGCTGCTGCAGGCTGTCTGCTGCAGGCT	356
V.Pacos/KAP10-9	TCCCTGCCAGCAGGCTGCTGTGAGCCGCTGCTGCAGGCTGTGTGCTGCAGGCC	417
	***** ***** ***** ***** *****	
C.hircus/KAP10-9	GTGTGCTGTGAGGCTCCCTGCTCAACTCCTCATGTTGCCAGCAGTCCAGTGCCAG	416
V.Pacos/KAP10-9	GTGTGCT-----	424
	** *****	
C.hircus/KAP10-9	CCCTCAGCTACACTCCTCCCTGCCAGCAGGCTGCTGTGAGCCATCTGCTGCAGG	476
V.Pacos/KAP10-9	-----GCACG	429
	*** *	
C.hircus/KAP10-9	CCCGTGTGCTGTGAGCCTGTGCTGTGAGCCGCTGTGCTGCAGGCCGTGTGCTGCACA	536
V.Pacos/KAP10-9	CCCGTGTG-----CTGAGGCTGTGCTGTGAGGCCGCTGTGCTGCAGGCCGTGTGCTGCAGG	474
	***** * ***** ***** ***** *****	
C.hircus/KAP10-9	CCTGTTGCTGCAGGCCGTGCTGTGAGGCTCCGCTGCTCATCCCCCTCGTCTGTC	596
V.Pacos/KAP10-9	CCTGTTGCTGCAGGCCGTGCTGTGAGGCTCCGCTGCTCATCCCCCTCGTCTGTC	528
	***** ***** ***** ***** ***** *****	
C.hircus/KAP10-9	TGCAGACCTCCTCCTCGTGTCCCTCCTTGCGCCCGGTGTGCCGCCCGCTGCTGTC	656
V.Pacos/KAP10-9	TGCAGACCTCCTCCTCGTGTCCCTCCTTGCGCCCGGTGTGCCGCCCGCTGCTGTC	588
	***** ***** ***** ***** ***** ***** ***** *****	
C.hircus/KAP10-9	GTGCCGACTCCTCCTGCCAGCCAGTTGCTGCCGCCAGCCTCCTCTGTGCCCTGCTC	716
V.Pacos/KAP10-9	GCCCGGCCCTCCTGCCAGCCAGTGTGCCGCCCGCCTCCTGCTGTCCCTGCTC	648
	* ***** * ***** ***** ***** ***** ***** *****	
C.hircus/KAP10-9	TGCAGCCCGTATGCTCCCGCCTGTTGCTGCTTCCCAACCTTGGCCCGGAGCCCTGC	776
V.Pacos/KAP10-9	TGCCGCCCTGATGCCCTGCCACCTGCTGTGCCCAACTCGCCTAGAAGCCCTGC	708
	***** * ***** ***** ***** ***** ***** ***** *****	
C.hircus/KAP10-9	TGCTGACCTGGCATGTACCCAGAGCCAGCCAGCTCAGGTTCCACCTGTGACTTGGTC	836
V.Pacos/KAP10-9	TGCTGACCGGC-CGTCCCCAGGGCCAGCCGGCTCAGTTCACCACTGACTTGGCC	767
	***** * ***** ***** ***** ***** ***** ***** *****	
C.hircus/KAP10-9	-----CTCTATCCTGACCTGGGCTGGGTGATTGCCCTCACTCAGGACGGGTCC	885
V.Pacos/KAP10-9	CCCAGTTCCTCAGTCTGACCTGAGTTATGTACTGTCCCACTGAGGACATGTCT	827
	***** ***** * * ***** * * ***** * * *	
C.hircus/KAP10-9	CCGATGTCTCACTCCCCTGACC-GACTGACCTCTCATTCTGCCCC-AGGAAC-CCCGA	942
V.Pacos/KAP10-9	GCTATGTATCCATTGCTGACTGACTGACTCTGACCACTGCTCCAGGACCCTCGA	887
	* ***** * ***** ***** ***** * ***** * * ***** *	
C.hircus/KAP10-9	CCTCAGGCTCCCTGGGCTTCTGCCTCCCAGGAGTCGTCTCCAGTGTCTCTGGGCCAG	1002
V.Pacos/KAP10-9	CCTTGTGTGCCCCAGGCTCTTGCCTCCAGGTTGA--CCCACTGCCTCTGGGTCAC	944
	*** * * ***** ***** ***** * * ***** ***** *	
C.hircus/KAP10-9	CTGCCCCATCCGTCAGATCCTATCTGGGTGTTTGGTCTCGACTTCCAACCTGT	1062
V.Pacos/KAP10-9	CTGCCCCCTC---CCATTTCTGCTGGGTCACTTGGCTCTCTTTCCAACCTGT	1000
	*** ***** * * ***** ***** ***** ***** *****	
C.hircus/KAP10-9	CAGCGC-CCCTCAGGCACTGGGTGAATAAACTCACATCATC	1103
V.Pacos/KAP10-9	CAGCATTTCTCAGGCACTGGAGGAATAAACTCACTCACC	1042
	***** ***** ***** ***** ***** * * *	



***-KRTAP16-4***

C.Hircus/KAP16-4	GACACACACTTCAGAAACATCCTCTTGCAGTCCACCTGAATCTGCTCCCCTGGACAACAT	60
V.Pacos/KAP16-4	CACACCCATTGAGAAACTTCCTCTTGCAACCCACCTGAATTCACATTCTAGACAACAT	60
	*** ** * ***** ***** * ** *****	
C.Hircus/KAP16-4	GTGTTGTAACCTACCTATGGCAACCTCCTGTGGCTATGGCTGTGGAAACAGCTATAGCTG	120
V.Pacos/KAP16-4	GTGTTGCAACTT--CTATGGCAA-CTCCTGTGGCTATGGCTGTGGAAACAGCTATAGCTG	117
	***** ** * ***** ***** ***** *****	
C.Hircus/KAP16-4	TGGGTTACGCCCTATTATGGCTGTGGATATGGAAGTAGATACAGTTGTGGATATGGCTC	180
V.Pacos/KAP16-4	TGGATTTAGCCCTATTATGGCTGTGGTTATGGTAC-----TGGATATGGCTG	165
	*** ** ***** ***** ***** * *****	
C.Hircus/KAP16-4	AGGATACGGCTGTGGATACGGTTCAGGATACGGCTGTGGATATGGACAGGATACAGCTG	240
V.Pacos/KAP16-4	TGGATACGGCTGTGGATATGGCTCCCGCTATGGCTGTGGTTATGGAAGTGGATATGGCTG	225
	***** ***** ** * ** ***** ***** * ** *****	
C.Hircus/KAP16-4	TGGATTTAGGCCCTATTATGGCTGTGGCTATGGAACAGATATGGCTGTGGATACGGCTC	300
V.Pacos/KAP16-4	TGGAT-----AT-----GGATCTGGCTT	243
	**** * ** **** **	
C.Hircus/KAP16-4	AAGCTATAGCAGCCACTGGCCAGTTTGCTACAGGAGATGTTATTC--TTCCTGTAGAA	357
V.Pacos/KAP16-4	TGGCTATTGCAGCTACCGCCAGTTTGCTATAGGACGTGTTATTCCTCTTGCTGTAGAA	303
	***** ***** ** ***** ***** ***** * ** *****	
C.Hircus/KAP16-4	CATCACCGTCTTAGCCTAGCTGCTTCTGAAATGAAAGATCTAAAGATAATACTGCTTTG	417
V.Pacos/KAP16-4	CATCACTATCCAAGCCCATTTGTTTCTGTAATGATACATCTAAAGACAATGCTGAC-TC	362
	***** ** ***** * ** ***** ***** ***** ***** * **	
C.Hircus/KAP16-4	AGGATCAGTAACCCAGGATTTCTACTTGCAAGAAATTACATGCCGACAGAGACTTT-TT	476
V.Pacos/KAP16-4	AGGATCTATAACCCACGGTTGCTATGTCCAAGAAATTACATGCCGACAGAGGCTTTGAT	422
	***** ***** * ** * ** ***** ***** ***** ***** *	
C.Hircus/KAP16-4	CTCCAACCTACCCAGCTTTAGAACACAGTCTCT-GAAGTCTGAGGTCTATGGTGATAACAT	535
V.Pacos/KAP16-4	CTCCAACCTGCCCATTCTTTCTATTGGCATCTTTGTGGCTTGAGGCCCC--GTGGTAGCAT	479
	***** ***** ** * ** * ** * ** ***** * ** * ** *	
C.Hircus/KAP16-4	CAGAC--TCTTCCAAATGTTAGCCTTTACTCCCTTAAT-CTCTGAGTC--TCTTATGACT	590
V.Pacos/KAP16-4	CTGACCATTTCCAAAAGTTAGCCTTTACCCTCCTCATTTTTGACTCTCTGTTATGACT	539
	* ** * ***** ***** ***** * ** * ** * ** * ** *****	
C.Hircus/KAP16-4	GTGTTAATGATCCTAACATTCATAGCATGGAACCCCTTATTCTCAATAAAAACTACT	650
V.Pacos/KAP16-4	GTATGATGATCCCAACATTTGCAATTGAGGAAATCACTTATTCTCAATAAAAAATGGTT	599
	** * ***** ***** * * ** * ***** ***** ***** * **	
C.Hircus/KAP16-4	CATTCCTTCAATAAAAAAAAAAAAAA-----AA	681
V.Pacos/KAP16-4	CATTGTTCTAACAACTCTATGTATACTTGCTTGTA	637
	*** ***** * ** * ** **	

**Anexo 11:** Alineamiento de las secuencias de los genes *KRTAP6-1L*, *KRTAP1-3* y *KRTAP9-9L* con la herramienta Clustal Omega.

**-KRTAP6-1L**

V.Pacos/KAP6-1	ATGTGGCTACTACGGAACTACTATGGGGCCGTGGCTACGGCTGCTGGCTACGGA	60
O.Aries/KAP6-1	ATGTGTGGCTACTACGGAACTACTATGGCGCCCTCGGCTGTGGAAGCTATGGCTATGGA	60
C.Hircus/KAP6-1	ATGTGTGGCTACTACGGAACTACTACGGCGCCCTCGGCTGTGGGAGCTATGGCTATGGA	60
	*****	
V.Pacos/KAP6-1	GGCCTGGGCTGCGGCTATGGGTCTGCTATGGCTGTGGCTCCGAGACT-----	110
O.Aries/KAP6-1	GGCCTGGGCTATGGCTATGGCTGTGGCTACGGTCTGGCTCCGAGGCTGGGCTGTGGC	120
C.Hircus/KAP6-1	GGCCTGGGCTGTGGCTATGGCTGTGGCTACGGCTGGCTTCTGAGGCTGGGCTGTGGC	120
	*****	
V.Pacos/KAP6-1	-----GGGCTGTGGCTATGGCTGTGGCTAT	135
O.Aries/KAP6-1	TATGGCTCCTGCTACGGCTCTGGCTCCGTAGGCTGGGCTGTGGCTATGGCTGTGGCTAT	180
C.Hircus/KAP6-1	TATGGCTCCTGCTACGGCTCTGGCTCCACAGGCTGGGCTGTGGCTATGGCTGTGGCTAC	180
	*****	
V.Pacos/KAP6-1	GGCTTTGGCTCCGCTCTCTCTGTGGCTGTGGCTATGGGTGTGGCTGTGGCTGTGGCTCT	195
O.Aries/KAP6-1	GGCTATGGATGCGGCTATGGCTCTGGCTTTGGCTACTACTATTGA-----	225
C.Hircus/KAP6-1	GGCTATGGCTCCGCTCTCTCTGTGGCTGTGGCTATGGATGCGGCTATGGCTGTGGC---	237
	*** ** *	
V.Pacos/KAP6-1	GGCTTTGGCTACTACTATTGAG	217
O.Aries/KAP6-1	-----	225
C.Hircus/KAP6-1	---TTGGCTACTACTATTGAA	256

**-KRTAP1-3**

V.Pacos/KAP1-3	ATGGCTGCTGTTCCACCAGCTTCAACGGATTTCCATCTGTTCTACTGGTGGAACTGT	60
O.Aries/KAP1-3	ATGGCTGCTGTTCCACCAGCTTCTGTGGATT-----	32
C.Hircus/KAP1-3	ATGGCTGCTGTTCCACTAGCTTCTGCGGATTTCCATCTGTTCCACTGGTGGAACTGT	60
	*****	
V.Pacos/KAP1-3	GGCTCCAGCTGCTGCCAGCCAACCTGC-----	87
O.Aries/KAP1-3	-----	32
C.Hircus/KAP1-3	GGCTCCAGTCCCTGCCAGCCGACCTGCTGCCAGACCAGCTGCTGCCAGCCAACCTCCATT	120
V.Pacos/KAP1-3	-----TCCCAGACCAGCTGCTGC	105
O.Aries/KAP1-3	--TCCATCTGTTCCACTGCTGGGACCTGTGGCTCCAGCTGCTGCCAGTCAACCTGCAGT	90
C.Hircus/KAP1-3	CAGACCAGCTGCTGCCAGCCAACCTCCA TCCAGACCAGCTGCTGCCAGCCAACCTCCATC	180
	* * * * *	
V.Pacos/KAP1-3	CAGCCAACCTGCGGCCAGCCAACCTGCTCCCAGACCAGCTGCTGCCAGCCAACCTGCTGC	165
O.Aries/KAP1-3	CAGACCAGCTGCTGCCAGCCAACCTCCA TCCAGACCAGCTGCTGCCAGCCAACCTGCCTC	150
C.Hircus/KAP1-3	CAGACCAGCTGCTGCCAGCCAACCTCCA TCCAGACCAGCTGCTGCCAGCCAACCTGCCTC	240
	*** *	
V.Pacos/KAP1-3	CAGACCAGCTGCTGAGACTGGCTGTGGCATTGGTGGTAGCATTGGCTGTGGCCAGGAG	225
O.Aries/KAP1-3	CAGACCAGTGGCTGTGAGACCAGCTGTGGCATTGGTGGCAGCATTGGCTATGGCCAGGTG	210
C.Hircus/KAP1-3	CAGACCAGTGGCTGTGAGACCAGCTGTGGCATTGGTGGCAGCATTGGCTATGGCCAGGTG	300
	*****	
V.Pacos/KAP1-3	GGTGGCAGCGGAGCTCTGAGCTGCCGACCAAGGTGGTGGCCCTGACTGCCGCTGGAG	285
O.Aries/KAP1-3	GGTAGCAGCGGAGCTGTGAGCAGCCGACCAAGGTGGTGGCCCTGACTGCCGCTGGAG	270
C.Hircus/KAP1-3	GGCAGCAGCGGAGCTGTGAGCAGCCGACCAAGGTGGTGGCCCTGACTGCCGCTGGAG	360
	** *****	
V.Pacos/KAP1-3	GGCACCTGCCTGCCTCCCTGCTGTGTGGTAGCTGCACCCACCAACCTGCTGCCAGCTG	345
O.Aries/KAP1-3	GGCACAGCCTGCCTCCCTGCTGTGTGGTAGCTGCACATCCCGTCTGCTGCCAGCTG	330
C.Hircus/KAP1-3	GGCACAGCCTGCCTCCCTGCTGTGTGGTAGCTGCACACCCCGTCTGCTGCCAGCTG	420
	*****	
V.Pacos/KAP1-3	TACCTGCCCCAGGCTCCTGCTGCCGCCATCCTACTGTGGACAGTCTGCTGCCGCCCA	405
O.Aries/KAP1-3	TACTATGCCAGGCTCCTGCTGCCGCCATCCTACTGTGGACAGTCTGCTGCCGCCCA	390
C.Hircus/KAP1-3	TACTATGCCAGGCTCCTGCTGCCGCCATCCTACTGTGGACAGTCTGCTGCCGCCCA	480
	** *****	
V.Pacos/KAP1-3	GCCTGCTGCCTCTGCTGTGAGCCACC-----TGCTGTGAACCCACCTGTTAAAAGCAA	459
O.Aries/KAP1-3	GCCTGCTGCTGCCAGCCACCTGCACTGAGCCCGTCTGTGAGCCACCTGCTGCCAACCC	450
C.Hircus/KAP1-3	GCCTGCTGCTGCCAGCTACCTGCACTGAGCCATCTGTGAGCCACCTGCTGTGAGCC	540
	***** *	
V.Pacos/KAP1-3	GGTTGCTGA	468
O.Aries/KAP1-3	ATCTGTTAA	459
C.Hircus/KAP1-3	ACCTGCTGA	549
	** * *	

**-KRTAP9-9L**

V. Pacos/KAP9-9	NNNNNNNNNNNNNNNNNNNNCCCTTGCTGCCAGCCACCTGCTGCCAGCCACCTG----	56
O. Aries/KAP9-9	ATGACCACCTCTGTGCTGCTCCCCACGCTGTCAGCCCCCTGCTGCAGGACCACCTGCTGT	60
C. Hircus/KAP9-9	ATGACCACCTCTGTGCTGCTTCCCGCGTGTGTCAGCCACCTGCTGCAGGACCACCTGCTGT *** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *	60
V. Pacos/KAP9-9	-----	56
O. Aries/KAP9-9	GAGTCCAGCGGCTGCCAGCCCTGCTGCCCCCGCAGTGTGTCAAACCACCTGCTGCAGG	120
C. Hircus/KAP9-9	GAGTCCAGCTGTGCAAGCCCTGCTGCCCCCCACTTGTGTCAAACCACCTGCTGCAGG	120
V. Pacos/KAP9-9	-----	56
O. Aries/KAP9-9	ACCACCTGTACAAACCTACCTGTGTGACCAGCTGCTGCCAGCCACCTGTTGCAGCAAA	180
C. Hircus/KAP9-9	ACCACCTGTGCAAACTACCTGTGTAATCAGCTGCTGCCCGTCCACCTGTTGCAGCAAA	180
V. Pacos/KAP9-9	-----CTGTGAGCCAGCTGCTGCCAGCCTTGCTGCCCCCAAAC	96
O. Aries/KAP9-9	CCCTGTGTGAGCCACCTGCTGTGAGTCCATCTGCTGCCAGCCCTCATGCCCTCCAAC	240
C. Hircus/KAP9-9	CCCTGTGTGAGCCACCTGCTGTGAGTCCATCTGCTGCCAGCCTCATGCCCTCCAAC ***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *	240
V. Pacos/KAP9-9	TGCTATGAAGCTACTCAGACCACCTGCTGCAGCACCACCTGCCGCAAACTACTTGTGTG	156
O. Aries/KAP9-9	TGCTATCAAAC TAGCGAAACCACCTGCTGCAGGACCACCTGCCACAAGCCTACTTGTGTG	300
C. Hircus/KAP9-9	TGCTATCAAAC TAGTGAACCACCTGCTGTAGAACCACCTGCCACAAGCCTACTTGTGTG ***** ** * ** * * ** * ** * ** * ** * ** * ** * ** * ** * ** *	300
V. Pacos/KAP9-9	ACCACCTGCTGCCAGCCACCTGCTGTGAGCCAGCAGCTGTGGACAAACCTGCTGCCAG	216
O. Aries/KAP9-9	AGCACC-----	306
C. Hircus/KAP9-9	ACCACC----- * ** *	306
V. Pacos/KAP9-9	CCTACCTGTGTAACCAGCTGTTGCAGCCCTTCTGCTGCCAGTCCACTTGTGTGAGTCT	276
O. Aries/KAP9-9	-----TGCTGTGAGCCACCTGCTGTGGGTCC	333
C. Hircus/KAP9-9	-----TGCTGTCAACCCACCTGCTGTGGGTCC ***** ** * ** * ** * ** * ** * ** *	333
V. Pacos/KAP9-9	AGCAGCTGCGGACAAAACCTGCGGTGGGTCCA-----GCTGCTGCCAGCCAGCT	324
O. Aries/KAP9-9	AGCAGCTGTGGACAAAACCTTCAAGTGGGTCCAGCTGCAGCCAGCCTTGCTGCCAGCCAGCT	393
C. Hircus/KAP9-9	AGCAGCTGTGGACAAAACCTTCAAGTGGGTCCAGCTGCAGCCAGCCTTGCTGCCAGTCAAGT ***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *	393
V. Pacos/KAP9-9	AGCTGTGCACCCGCTGTACTGCCACCGAACCTGGTACCACCCACGTGCTGCTGCCTGCCT	384
O. Aries/KAP9-9	TCCTGTGCACCTGTGTACTGCCACAGAACTGCTACCACCCACGTGCTGCTGCCTGCCT	453
C. Hircus/KAP9-9	TACTGTGCCCCGTGTACTGCAACAGAACTGCTACCACCCACGTGCTGCTGCCTGCCT ***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *	453
V. Pacos/KAP9-9	GGGTGCCTGGCCAGAGCTGTGGATCCAGCTGCTGCCAGCCTCGCTGCCGCCCTGTCTGC	444
O. Aries/KAP9-9	GGGTGCCAAGCCAGGAATGTGGGTCCAGCTGCTGCCAGCCTTGC AACGCCCTGTCTGC	513
C. Hircus/KAP9-9	GGGTGCCAAGCCAGGAATGTGGGTCCAGCTGCTGCCAGCCTTGCAGCCGCCCTGTGTGC ***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *	513
V. Pacos/KAP9-9	TGTCAGACCACGTGCTGCCGCCCTAGCTGTGTGTC-----CAGCTGCTGC	489
O. Aries/KAP9-9	TGTCAGACCACCTGCTGTAGGACCACCTGCTGCCGCCACAGCTGTGTGTCACCTGCTGC	573
C. Hircus/KAP9-9	TGTCAGACCACCTGCTGCCGCCACAGCTGTGTGTCACCTGCTGTGTCAGTCTTCTGTTGC ***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *	573
V. Pacos/KAP9-9	CAGCCCTCCTGCTGA	504
O. Aries/KAP9-9	CCTTCCTTCTGCTGA	588
C. Hircus/KAP9-9	TGA-----	576

**Anexo 12:** Alineamiento de las secuencias de los genes *KRTAP4-8L* y *KRTAP4-9L* con la herramienta Clustal Omega.

**-KRTAP4-8L**

V.Pacos/KAP4-8	ATGGTCAGCTCCTGTGTGGCTCCATCTGTTCTGAGGAGGACTGTGGCCAAGGACTCTGC	60
O.Aries/KAP4-8	ATGGTCAGCTCCTGTGTGGCTCCGCTGCTCTGACCAGAGCTGCGGCCAAAGTCTCTGC	60
C.Hircus/KAP4-8	ATGGTCAGCTCCTGTGTGGCTCCGCTGCTCTGACCAGAGCTGTGGCCGAAGTCTCTGC	60
	***** **	
V.Pacos/KAP4-8	CAGGAGACCTGCTGCCGCCAGCTGTGCCAAACAACCTGCTGCAGGACCACCTGCTAC	120
O.Aries/KAP4-8	CAGGAGACCTGCTGTCGCCCCAGCTGTGCCAGACCACCTGCTGCAGGACCACCTGCTAC	120
C.Hircus/KAP4-8	CAGGAGACCTGCTACCGGCCAGCTACTGCCAGACCACCTGCTGCAGGACCACCTGCTAC	120
	***** **	
V.Pacos/KAP4-8	CGCCCAAGGTGCTGTGAGTCCAGCTGTGCCGCCCCAGATGCTGCATCTCCAGTGCCAG	180
O.Aries/KAP4-8	CGCCCCAGCTGCGGTGTGTCCAGCTGTGCCGCCCATCTGCTGCCAGCCCACCTGCCCCT	180
C.Hircus/KAP4-8	CGCCCCAGCTGTGGTGTGTCCAGTTGCTGTGGCCCTGTGTGCTGCCAGCCCACCTGCCCCT	180
	***** **	
V.Pacos/KAP4-8	CCCACC-----	186
O.Aries/KAP4-8	CGCCCCACCTGCTGCATCTCTAGCTGTCCCGCCCTCCTGCTGTGTTCCAGCTGTGGT	240
C.Hircus/KAP4-8	CACCCCACCTGCTGCATCTCTAGCTGTACCACCCCTCCTGCTGTGTTCCAGCTGTGGT	240
	* * **	
V.Pacos/KAP4-8	-----TGTTGCCGCCCCAGATGCTGTGAGTCCAGCTGCTGCAGACCCAGTGTGCCAG	240
O.Aries/KAP4-8	TCCAGCTGTACAGACCAACCAGCTGCATCTCCAGCTGTGCAGGCCCCAGTGTGCCAG	300
C.Hircus/KAP4-8	TCCAGCTGTACAGGCCGACCAGCTGCATCTCCAGAGCTGCAGGCCCCAGTGTGCCAG	300
	* * * * * **	
V.Pacos/KAP4-8	TCTGTGTGCTGCCAGCCACCTGCTGCCGCCGAGCTGTGTGAGTCCAGTT-----	292
O.Aries/KAP4-8	CCTGTGTGCTGCCAGCCACCTGCCCTGCCCCACCTGTTGCATCTCTAGCTGTACCGC	360
C.Hircus/KAP4-8	CCTGTGTGCTGCCAGCCACCTGCCCCGCATCTCC-----	336
	***** ** * *	
V.Pacos/KAP4-8	-----	292
O.Aries/KAP4-8	CCCTCCAGCTGTGGCTCCAGCTGTGGCTCCAGCTGTGCAGGCCTACCTGCTGCATCTCC	420
C.Hircus/KAP4-8	-----	336
V.Pacos/KAP4-8	-GCTGCAGACCCAGTGTGCCAGTCTGAGTGTGCCAGCCACCTGCTGCCGCCCT--	349
O.Aries/KAP4-8	AGCTGCAGGCCCGGTGCTGCCAGTCTGTGTGCTGCCAGCCAGCTGCCCCGCATCTCC	480
C.Hircus/KAP4-8	-----	336
V.Pacos/KAP4-8	-CCTGCTGTGGTTCCAGCTCCTGTGGCTCCAGCTGTGCCGCACCTGCTGTTGCCTGCGC	408
O.Aries/KAP4-8	AGCTGCTGCCGCCCTCTTGCTGTGGCTCCAGCTGTGCCGCCGAGCTGTGCCTGCGC	540
C.Hircus/KAP4-8	AGCTGCTGCCGCCCTCTTGCTGTGGCTCCAGCTGTGCCGCCGAGCTGTGCCTGCGC	396
	***** * ** * ***** * **** *****	
V.Pacos/KAP4-8	CCAGTCTGTGGCCAGGTCTCCTGCCACACCACTTGCTATCGCCACCTGTGTCATCTCC	468
O.Aries/KAP4-8	CCAGTGTGTGGCCGGTCTCCTGCCACACCACTTGCTATCGCCACCTGTGTCATCTCC	600
C.Hircus/KAP4-8	CCAGTGTGTGGCCGGTCTCCTGCCACACCACTTGCTACCGCCACCTGTGTCATCTCC	456
	***** ***** ***** ***** *****	
V.Pacos/KAP4-8	ACCTGCCCCGCCCCAGTGTGCTGTGCCTCCTTGTCTGA 510	
O.Aries/KAP4-8	ACCTGCCCCGCCCCGTGTGCTGTCCCTCCTTGTGCTGA 642	
C.Hircus/KAP4-8	ACCTGCCCCGCCCCGTGTGCTGTCCCTCCTTGTGCTGA 498	
	***** ***** ***** *****	

**-KRTAP4-9L**

V. Pacos/KAP4-9	ATGGTCAGCTCCTGCTGTGGCTCCATCTGTTCTGAGGAGGACTGTGGCCAAGGACTCTGC	60
O. Aries/KAP4-9	ATGGTCAGCTCCTGTTGTGGCTCCGTCTGCTCTGACCAGAGCTGGCCGGAAGTCTCTGC	60
C. Hircus/KAP4-9	ATGGTCAGCTCCTGTTGTGGCTCTGTCTGCTCTGACCAGAGCTGCAGCCGAAGTCTCTGC	60
	***** **	
V. Pacos/KAP4-9	CAGGAGACCTGCTGCCGCCACGCTGCTGCCAAACAACCTGCTGCAGGACCACCTGCTAC	120
O. Aries/KAP4-9	CAGGAGACCTGCTGCCGCCACGCT-----GCTGCCAGACCACCTGCTGCCAGACCACCTGCCCT	85
C. Hircus/KAP4-9	CAAGAGACCTGCTACCGTCCCAGCTGCTGCCAGACCACCTGCTGCAGGACCACCTGCTAC	120
	** ***** **	
V. Pacos/KAP4-9	CGCCCAAGGTGCTGTGAGTCCAGCTGCTGCCGCCCCAGATGCTGCATCTCCAGCTGC---	177
O. Aries/KAP4-9	-----GCTGCCAGACCACCTGCTGCCAGACCACCTGCCCT	120
C. Hircus/KAP4-9	CGCCCAAGGTGCTGTGAGTCCAGCTGCTGCCGCCCCGTCGCTGCCAGCCACCTGCCCT	180
	***** ** ***** **	
V. Pacos/KAP4-9	-----CAG	180
O. Aries/KAP4-9	CGCCCAAGGTGCTGTGAGTCCAGCTGCTGCCGCCCCCTCCGCTGTGTTCCAGCTGTGGT	180
C. Hircus/KAP4-9	CGCCCAAGGTGCTGTGAGTCCAGCTGCTGCCGCCCCCTCCGCTGTGTTCCAGCTGTGGT	240
V. Pacos/KAP4-9	CCCACCTGTTGCCGCCACGATGCTGTGAGTCCAGCTGCTGCAGACCCAGTGTGCCAG	240
O. Aries/KAP4-9	TCCAGCTGCTACAGGCCTACCGGTGCTCTCCAGCTGCTGCAGGCCACCTGCTGCCAG	240
C. Hircus/KAP4-9	CCCAGCTGCTACAGGCCTACCGGTGCTCTCCAGCTGCTGCAGGCCACCTGCTGCCAG	300
	*** ** * * * * * *****	
V. Pacos/KAP4-9	TCTGTGTGCTGCCAGCCACCTGCTGCCGCCAGCTGCTGTGAGTCCAGTTGC-----	294
O. Aries/KAP4-9	CCTGTGTGCTGCCAGCCACCTGCCCTGCCCCACCTGCTGCATCTCTAGCTGTACC	300
C. Hircus/KAP4-9	CCTGTGTGCTGCCAGCCACCTGCCCTGCCCCACCTGCTGCATCTCTAGCTGTACC	360
	***** ** ***** ** * * *	
V. Pacos/KAP4-9	-----	294
O. Aries/KAP4-9	CCCTCCAGCTGTGGGTCCAGCTGTGGCTCCAGCTGCTGCAGGCCTACTTGTGCATCTCC	360
C. Hircus/KAP4-9	CCCTCCAGCTGTGGGTCCAGCTGTGGCTCCAGCTGCTGCAGGCCTACTTGTGCATCTCC	420
V. Pacos/KAP4-9	---TGCAGACCCAGTGTGCCAGTCTGAGTGTGCCAGCCACCTGCTGCCGCCCT---	349
O. Aries/KAP4-9	AGCTGCAGGCCCGGTGTTGCCAGTCTGTGTGCTGCCAGCCAGTGCACCCGATCTCC	420
C. Hircus/KAP4-9	AGCTGCAGGCCCGGTGTTGCCAGTCTGTGTGCTGCCAGCCAGTGCACCCGATCTCC	480
	**** * * * * * ***** ** * * * *	
V. Pacos/KAP4-9	-CCTGCTGTGGTTCCAGCTCCTGTGGCTCCAGCTGCTGCCGCACCTGCTGTTGCCTGCGC	408
O. Aries/KAP4-9	AGCTGCTGCCGCCCTCTTGCTATAGCTCCAGCTGCTGCCGCCAGCTGCTGCCCTGCGC	480
C. Hircus/KAP4-9	AGCTGCTGCCGCCCTCTTGCTATAGCTCCAGCTGCTGCCGCCAGCTGCTGCCCTGCGC	540
	***** * * * * * ***** * * * * *	
V. Pacos/KAP4-9	CCAGTCTGTGGCAGGTCTCCTGCCACACCACTTGCTATCGCCCACTGTGCATCTCC	468
O. Aries/KAP4-9	CCAGTGTGCCGCCGGTCTCCTGCCACACCACTTGCTATCGCCCACTGTGCATCTCC	540
C. Hircus/KAP4-9	CCAGTGTGTGGCAGGTCTCCTGCCACACCACTTGCTATCGCCCACTGTGCATCTCC	600
	***** ** *****	
V. Pacos/KAP4-9	ACCTGCCCCCGCCCGTGTGCTGCTCCTCTTGCTTCTGA	510
O. Aries/KAP4-9	ACCTGCCCCCGCCCGTGTGCTGCTCCTCTTGCTTCTGA	582
C. Hircus/KAP4-9	ACCTGCCCCCGCCCGTGTGCTGCTCCTCTTGCTTCTGA	642
	***** * * * * * *****	