

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

FACULTAD DE CIENCIAS BIOLÓGICAS

E.A.P. DE GENÉTICA Y BIOTECNOLOGÍA

**“Búsqueda de genes relacionados a la síntesis de la fibra y
marcadores SSR en los ESTs de piel de alpaca *Vicugna*
pacos.”**

TESIS

Tesis para optar el Título Profesional de Biólogo, Genetista, Biotecnólogo

AUTOR

Daniel Antonio Fernández Ugaz

Lima-Perú

2015



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

(Universidad del Perú, DECANA DE AMÉRICA)

FACULTAD DE CIENCIAS BIOLÓGICAS

ESCUELA ACADÉMICO PROFESIONAL DE GENÉTICA Y BIOTECNOLOGÍA

**Búsqueda de genes relacionados a la síntesis de la fibra y
marcadores SSR en los ESTs de piel de alpaca *Vicugna pacos*.**

TESIS

Tesis para optar al Título Profesional de Biólogo, Genetista, Biotecnólogo.

AUTOR

Bach. Daniel Antonio Fernández Ugaz

ASESOR

Dr. Lenin Maturrano Hernandez

Lima - Perú

2015

A mi familia, amigos y profesores.

AGRADECIMIENTOS

Agradezco a mi padre Dionisio Isaac Fernández Matencio, más que por sus palabras, por la forma de como verlo trabajar muy duro y siempre siguiendo adelante a pesar de las dificultades, me transmitió el trabajo duro, de niño me dijo “si te dedicas a algo que en verdad te apasiona, nunca buscaras trabajo, el trabajo vendrá a ti”. Agradezco a mi madre Luz Irene Ugaz de Fernández por su increíble altruismo ante todas las personas, sus cualidades humanas que me transmitió, generosidad, bondad y humildad. Agradezco a mi hermano Ernesto Isaac Fernández Ugaz, al verlo siempre creciendo en su carrera, ingeniería de electrónica, a pesar de las dificultades, cuando muchos se darían por vencido, el continua y continua, solo con un fin ayudar a la familia haciendo lo que desde muy niño le gustaba. Sin duda la educación básica viene de casa, dedicarnos a algo que en verdad nos apasiona no hará la vida más fácil, igual hay retos, pero si será más divertida y con gusto aceptaremos los retos.

Al Dr. Lenin Maturrano por confiar en mí para desarrollar este estudio, junto con el Dr. Raúl Rosadio me han enseñado bastante sobre cómo enfocar un estudio y la importancia de su impacto para ayudar a la sociedad. Destaco junto con ellos a la Dra. Rina Ramírez, que obtuve un gran enfoque en el tiempo que estaba en su laboratorio; entre los tres doctores me enseñaron que la ciencia no solo es para llenarnos de grados académicos y publicaciones científicas sino esencialmente está para entender los diversos fenómenos, un entendimiento amplio y profundo del tema que abarquemos y ayudar al desarrollo de la sociedad y el medio ambiente en que vive.

Agradezco a mis buenos amigos del laboratorio Jorge Maximiliano, Nataly Allasi, Raquel Hurtado, Rocio Rimac, Ana Chero, Eduardo Justamayta, Geraldine Marcelo, Andrea Rodriguez, Gerardo Diaz, Aldo Ciprian y Juan Siuce y Juan Manual Aguilar por hacer esta etapa muy amena.

INDICE

DEDICATORIA	ii
AGRADECIMIENTOS	iii
INDICE	iv
RESUMEN	vi
ABSTRACT	vii
1. INTRODUCCIÓN	1
2. MARCO TEÓRICO	3
2.1. La alpaca.	3
2.2. La piel de alpaca.	3
2.2.1. La epidermis.	3
2.2.2. La dermis.	4
2.2.3. La hipodermis.	4
2.3. El folículo piloso y su fibra.	5
2.3.1. El folículo piloso.	5
2.3.2. El eje del pelo: La fibra del pelo.	7
2.4. Factores no genéticos que influyen en la fibra.	10
2.5. La fibra del pelo y sus principales componentes moleculares.	12
2.5.1. Los filamentos intermedios de queratinas (KIFs) y Las Proteínas Asociadas a Queratinas (KAPs).	12
2.6. Aspectos moleculares de la morfogénesis del folículo piloso.	15
2.7. Los microsatélites en las regiones codificantes	20
2.8. Etiquetas de secuencia expresadas (ETSs).	21
3. HIPÓTESIS	23
4. OBJETIVOS	22
5. MATERIALES Y MÉTODOS	24
5.1. Materiales.	24
5.1. Métodos.	25

5.2. Ensamblaje de los ESTs.	26
5.3. Distribución de los ESTs en los Unigenes.	26
5.4. Homología en la Base de Datos.	26
5.5. Anotación de Secuencias.	27
5.6. Búsqueda de microsatélites.	28
6. RESULTADOS.	29
6.1. Ensamblaje de los ESTs.	29
6.2. Distribución de los ESTs en los Unigenes.	31
6.3. Homología en las bases de datos.	32
6.4. Anotación de secuencias.	33
6.5. Búsqueda de microsatélites.	35
6.6. Genes altamente expresados.	36
6.7. Genes de interés en la síntesis de fibra.	37
7. DISCUSIÓN.	38
7.1. Identificación de la cantidad de tipos y copias de mRNA.	38
7.2. Identificación de las funciones de los mRNA.	39
7.3. Identificación de genes asociados al desarrollo y la calidad de la fibra.	40
7.4. Identificación de las Queratinas y KAPs.	40
7.5. Identificación de Factores de Necrosis Tumoral (TNF).	41
7.6. Identificación de factores asociados a la vía de señalización Wnt/ β -catenin.	42
7.7. Identificación de colágenos.	43
7.8. Identificación de señales proteínas morfogénicas oseas.	43
7.9. Identificación de microsatélites.	44
8. CONCLUSIÓN.	45
9. RECOMENDACIONES.	46
10. BIBLIOGRAFIA.	47

RESUMEN

La comercialización de la fibra de alpaca es una de las principales fuentes de sustento para más de un millón de personas en el Perú, siendo de mucha importancia para su economía por ser el principal productor de fibra. Teniendo en cuenta que dos de los factores importantes en la calidad de la fibra son su diámetro y su longitud, se analizaron los ESTs de la alpaca depositados en la base de datos del NCBI para identificar los genes relacionados a la síntesis de fibra y microsatélites que tendrían una función importante relacionada a su calidad. Se analizaron 7286 ESTs de piel de una alpaca del tipo huacaya, resultando en 2428 secuencias ensambladas o unigenes, posteriormente se realizó la búsqueda de homologías y la anotación de secuencias. Se identificaron 22 genes relacionados a la síntesis de la fibra (*KRT10*, *KRT25*, *KRT31*, *KRT14*, *KRT15*, *KRT5*, *KRT78*, *KRT81*, *KRT85*, *KRTAP3-2*, *KRTAP10*, *KRTAP11-1*, *TNFSF12*, *sFRP4*, *RACK1*, *COL1A1*, *COL1A2*, *DAB2IP*, *GLI1*, *BMP1*, *BMP4* y *CTNNB1*) donde se observó 4 de estos con una alta expresión (*TNFSF12*, *sFRP4*, *COL1A1* y *COL1A2*) y 10 genes poseen microsatélites (*COL1A1*, *TNFSF12*, *DAB2IP*, *GLI1*, *BMP4*, *KRT81*, *KRT5*, *KRT15*, *KRT85* y *KRTAP3-2*). El presente estudio aporta la identificación de genes relacionados a la síntesis de fibra y marcadores EST-SSR con los cuales se pueden probar en la alpaca para establecer de forma experimental si estos genes tienen una relación con el diámetro

Palabras clave: transcriptoma, huacaya, folículo piloso.

ABSTRACT

The marketing of alpaca fiber alpaca is one of the main sources of livelihood for more than one million people in Peru, being the country's leading producer of its fiber giving much importance to their economy. Two important factors in the quality fiber is its diameter and length, to know which genes are expressed and which of these are the ones that would have an important function related to its quality alpaca, the ESTs of alpaca fiber was studied. 7286 ESTs was analyzed in the present study, these belong a skin from alpaca huacaya type deposited in the NCBI database, resulting in 2428 sequences unigenes, later the homology search and the functional annotation. 22 genes were identified as related to the synthesis of fiber (*KRT10*, *KRT25*, *KRT31*, *KRT14*, *KRT15*, *KRT5*, *KRT78*, *KRT81*, *KRT85*, *KRTAP3-2*, *KRTAP10*, *KRTAP11-1*, *TNFSF12*, *sFRP4*, *RACK1*, *COL1A1*, *COL1A2*, *DAB2IP*, *GLI1*, *BMP1*, *BMP4* and *CTNNB1*) where 4 of these have high expression (*TNFSF12*, *sFRP4*, *COL1A1* and *COL1A2*) and 10 genes were found to have microsatellites (*COL1A1*, *TNFSF12*, *DAB2IP*, *GLI1*, *BMP4*, *KRT81*, *KRT5*, *KRT15*, *KRT85* and *KRTAP3-2*). This study provides genes and microsatellites markers which can be tested in the alpaca to establish experimentally the ratio of these genes the diameter.

Keywords: transcriptome, huacaya, hair follicle.

1. INTRODUCCIÓN.

Las alpacas son un recurso pecuario muy importante para la comunidad andina ya que son fuente de carne y fibra para los criadores que habitan en zonas mayores a 4 000 msnm, en este ámbito el Perú posee el 87% de la población mundial de alpacas 3'685,516 cabezas y se convierte en el primer productor de fibra. El 79% de la crianza alpaquera se da en los departamentos: Puno, Cuzco, Arequipa, Tacna y Moquegua donde el 99% se encuentra en poder de pequeños criadores y el 1% en diversas formas organizativas (medianas y grandes empresas, cooperativas, asociaciones, comunidades campesinas, etc) (CENAGRO 2012) y se estima que al menos un millón y medio de personas se dedican a su crianza en las zonas altoandinas (De los Ríos, 2006).

A la fibra extraída de los camélidos se le denomina fibra grasienta, mal llamada lana (De los Ríos, 2006). Por sus características de suavidad, calidez, variedad de colores y buena resistencia estas fibras son considerados por la industria textil como fibras especiales y sus derivados son clasificados como artículos de lujo (Wang *et al.*, 2003). La característica más importante en la fibra es su diámetro, la misma que está relacionado con el confort, por lo que un menor diámetro otorgará un mayor confort.

Su participación en el mercado de fibras finas de origen animal es del 10% mientras que la fibra de cabra de Angora, mohair, es el de mayor participación con un 56%. El Perú posee el 80% de la producción mundial de fibra de alpaca, seguido de Bolivia con 15%. También es el principal exportador de fibra de alpaca donde el Perú aporta el 88% del volumen, seguido de Bolivia que aporta el 10% (Brenes *et al.*, 2001).

La disminución de la calidad de la fibra de alpaca es causada, probablemente, por la hibridización que existe entre la alpaca y la llama, originada después de la conquista española, la cual ha modificado la composición genética de las actuales poblaciones. Cerca del 90% de la producción de la fibra de alpaca en Perú tiene un diámetro mayor a 25 μm alcanzando precios bajos. Antes del contacto europeo la fibra de alpaca tenía un diámetro entre 17-22 μm (Wheeler, 1995) similar al cashmere (15-17 μm). Se conocen dos tipos de alpaca, el tipo Suri que presenta fibras de gran longitud y lustrosas y el tipo Huacaya que es en general una fibra un poco más fina y menos larga, de vellón más compacto y esponjoso, con mayor rizado similar al de ovinos, el cual representa el 85% de la población de alpacas del Perú (Quispe *et al.*, 2009). Las características de la calidad de fibra, como el diámetro y la longitud así como la coloración de esta son de los factores más importantes para su comercio.

La fibra de alpaca también es muy apreciada por su gran diversidad de colores, normalmente se determina 22 colores, que incluyen el blanco puro, cremas, marrones, plata, grises y negras, sin embargo en la industria textil y las asociaciones de alpaqueros dentro y fuera del Perú se suele usar un número menor de colores (Oria *et al.*, 2009). Dentro de la gran gama de colores, el color blanco es el más apreciado por que permite su tinción, esta preferencia ha hecho que las poblaciones de alpacas blancas sean las de mayor número (Bustinza, 2001). Pero la problemática de la fibra de alpaca abarca muchos factores graves como la disminución de la calidad de fibra, la disminución de las poblaciones de colores que podría tender a su desaparición en algunas zonas, como el color gris en Huancavelica que es demasiado escaso, la hibridización, la frecuencia menor de machos lo que significa disponer de un menor tamaño efectivo de las poblaciones, la gran incidencia de castración que existe en algunas zonas, los defectos congénitos y enfermedades que se presentan (Oria *et al.*, 2009), además abarca también la producción, aunque en las poblaciones de alpacas se muestra un crecimiento, la calidad de su fibra no ha mejorado sustancialmente, así también su producción, como referencia se tiene que hasta el 2003 se obtenía, en promedio, 2.3 kilos por alpaca cuando se podría llegar a 5.4 kilos con buen programa de crianza (De los Ríos, 2006).

Uno de los enfoques para estudiar los genes que se están expresando en determinado momento es la transcriptómica. Uno de los métodos para estudiar transcriptómica es a través de la generación de marcadores de secuencias expresadas (ESTs), actualmente Fan *et al.* (2011) generaron ESTs de piel de alpaca huacaya, su estudio fue enfocado a identificar los genes involucrados en la coloración, el presente estudio fue enfocado a identificar genes, su expresión y marcadores EST-SSR relacionados al desarrollo de fibra, teniendo en cuenta que las bases de datos han aumentado, los genes y alelos involucrados al desarrollo de la fibra en otros mamíferos se encuentran en aumento, todo esto ayuda a la identificación de nuevas secuencias que en un inicio se desconocía su función.

Habiendo tantos factores que afectan tanto la calidad y la producción de la fibra, este estudio pretende ayudar a entender la biología del folículo de la fibra de alpaca Huacaya identificando genes asociados a la síntesis de fibra y la identificación de marcadores microsatélites dentro de estos genes y proponer los genes que podrían tener un mayor efecto en la calidad dado el complejo proceso de su morfogénesis

2. MARCO TEORICO.

2.1. La alpaca.

Los camélidos son animales pertenecientes al orden Artiodactyla, suborden Tylopoda y pertenecientes a la familia Camelidae. El antecesor de esta familia se originó en Norte América durante el Eoceno, con la división de la familia Camelidae en 2 tribus, la tribu Camelini que son los actuales camellos, el dromedario (*Camelus dromedarius*), el bactraciano (*Camelus bactrianus*) y el bactriano salvaje (*Camelus ferus*), estos se distribuyen en Eurasia y África y la tribu Lamini donde encontramos a la llama (*Lama glama*), el guanaco (*Lama guanicoe*), la vicuña (*Vicugna vicugna*) y la alpaca (*Vicugna pacos*), se distribuyen en el sur de América, (Wheeler, 1995; Wheeler *et al.*, 2004).

La alpaca es un animal domesticado descendiente de la vicuña, su domesticación se produjo en la puna húmeda de los Andes centrales del Perú hace 6-7 mil años atrás (Wheeler, 1995; Wheeler, 2004).

2.2. La piel de alpaca.

La piel de alpaca, al igual que en otros mamíferos, está formada por 3 capas, la epidermis, la capa más delgada y externa; la dermis, una capa gruesa e interna y la hipodermis, una capa de grasa (Bustinza, 2001).

2.2.1. La epidermis.

La epidermis es un tejido formado por células epiteliales estratificadas planas y ligeramente queratinizadas, dividido en 4 estratos, su espesor puede cambiar según la zona cutánea y no siempre se pueden evidenciar los 4 estratos nombrados (Lacolla *et al.*, 2010).

Estrato basal: Conformado por queratinocitos cúbicos o cilíndricos (Bustinza, 2001) de citoplasma claro, con presencia de melanocitos en proporción de uno de cada 14 a 20 queratinocitos basales (Torres de Jasauí *et al.*, 2007). En esta capa el tipo Suri posee un menor contenido de grasa comparado con el tipo Huacaya, también a diferencia que otras especies, la alpaca tiene este estrato más delgado (Bustinza, 2001).

Estrato espinoso: Conformado por células poliédricas (Bustinza, 2001) dispuestas en 1 a 2 capas y con apreciación de desmosomas muy finas (Torres de Jasauí *et al.*, 2007).

Estrato granuloso: Conformado por una capa muy fina de queratinocitos en forma de uso con granulaciones basófilas (queratohialina, sustancia precursora de la queratina) y núcleos pictónicos (Torres de Jasauí *et al.*, 2007).

Estrato córneo: Capa delgada de tejido con escamas córneas y queratinizado (Bustinza, 2001), donde los queratinocitos han perdido su núcleo (Torres de Jasauí *et al.*, 2007).

2.2.2. La dermis.

De manera general la dermis no presenta características diferenciales apreciables (Lacolla *et al.*, 2010). Esta está formada por el tejido conectivo, encontrando en ella fibras de colágeno, en esta capa encontramos el folículo piloso, las glándulas sebáceas sudoríparas y el músculo erector del pelo.

Esta capa también se divide en 2 capas: la Dermis Superficial Delgada, conformado por tejido conectivo laxo, células conjuntivas o fibrocitos; a esta capa también se le llama "lámina propia", mientras más profunda, se hace más densa, en esta parte densa encontramos los "**nidos foliculares**". La segunda capa es la Dermis Profunda formada por tejido conectivo denso, en esta zona encontramos los bulbos pilosos (Bustinza, 2001).

2.2.3. La hipodermis.

La hipodermis está formada por el tejido conectivo laxo, esta capa ejerce la función de fijar la dermis a los huesos o músculos y cuya característica principal es la presencia de un alto contenido de células adiposas. En ella se encuentran los capilares sanguíneos los cuales rodean los folículos, en los camélidos sudamericanos, estos superan en cantidad a la de los ovinos y cerdos. Una diferencia importante es que en el tipo Huacaya encontramos capilares que no llegan a acercarse a los grupos foliculares sino que terminan a una cierta distancia, lo cual sugiere que la difusión de sustancias necesarias para éstos sería a través del tejido conectivo, mientras que en el tipo Suri son más abundantes y se acercan más a los grupos foliculares, estas características sugieren que la piel del tipo Suri es semejante a la de los animales de climas calurosos, lo que es reforzado por su característica de tener un vellón abierto y su menor resistencia a las condiciones de las zonas altas (Bustinza, 2001).

2.3. El folículo piloso y su fibra.

El pelo es una característica primaria de los mamíferos y ejerce un amplio rango de funciones incluyendo la termorregulación, protección física, dispersión del sudor, actividad sensorial entre otros. Su estructura consiste en dos partes: el folículo piloso y el eje del pelo o fibra.

2.3.1. El folículo piloso.

El folículo piloso en general en los mamíferos es un complejo miniórgano de la piel el cual junto con la glándula sebácea y el musculo erector del pelo forman la unidad pilosebácea. El folículo en los mamíferos posee 4 regiones: la zona superior o **infundibulo** entre el nivel de la epidermis y la desembocadura de la glándula sebácea, el **itsmo** entre la desembocadura de la glándula sebácea y la inserción del musculo erector, es en esta sección donde se halla la protuberancia o *bulge* que alberga células madres, la región **suprabulbar** entre el musculo erector y el bulbo y la región del **bulbo** por debajo de la inserción del musculo erector (Schneider *et al.*, 2009), en esta capa se encuentra la matriz pilosa que también contiene células madres.

El folículo piloso en la alpaca, como otros mamíferos, posee la **vaina radicular externa (ORS, outer root sheath)**, cuyo extremo profundo dilatado constituye el bulbo, y la **vaina radicular interna (IRS, inner root sheath)** (Figura 1). La ORS se divide en la capa basal y la suprabasal mientras que la IRS se divide en la capa de Henle, la capa de Huxley y la cutícula. El bulbo contienen la matriz pilosa y melanocitos, de su región central superior surge el eje del pelo o fibra (Torres de Jasauí *et al.*, 2007).

El folículo piloso pasa por un ciclo de 3 fases, **Anágen** que es la fase de crecimiento del epitelio folicular, este periodo es el más largo, en oveja puede llegar hasta los 2 años produciendo ~15 cm de longitud anual y un diámetro de 20 μm , **Catágen** donde ocurre una apoptosis y regresión de la fase junto con la reducción de la papila dérmica y **Telógen** que es de una inactividad relativa, una fase de reposo (Torres de Jasauí *et al.*, 2007; Stenn y Paus, 2001).

Este ciclo se repite durante toda la vida y en cada ciclo se genera una fibra nueva dependiendo de la activación de las células madres multipotentes epiteliales pelo-específicas localizadas en la parte inferior del bulbo (Schneider *et al.*, 2009). Aunque el folículo piloso es altamente sensitivo a numerosos factores de crecimiento, citoquinas, neuropéptidos y hormonas que en parte son producidas por el mismo

folículo, su ciclo es un fenómeno autónomo ya que es capaz de continuar incluso en cultivos de folículos piloso aislados (Stenn y Paus, 2001).

El bulbo del folículo en la etapa de crecimiento, anágen, contiene la matriz de queratinocitos y la unidad de pigmentación, la matriz de queratinocitos contiene células las cuales migran fuera de la parte inferior del bulbo, son células que proliferan rápidamente y su número determina el tamaño del bulbo y el diámetro de la fibra (Stenn y Paus, 2001). Estas células de la matriz proliferan y luego se diferencian, dando lugar a varios tipos de células de la fibra del pelo y de la IRS, mientras que la ORS tiene un origen policlonal y muestra un crecimiento regional (Legué y Nicolas, 2005).

En los mamíferos no solo existe un tipo de folículo, se reconocen por lo menos 2 tipos, los **folículos primarios (FP)** y los **folículos secundarios (FS)**. Estos pueden encontrarse en forma simple o en agrupaciones, en forma simple se encuentran los FP simples o aislados mientras que en las agrupaciones se puede encontrar a un FP rodeado por varios FS, a estos últimos se les llama también nido folicular y son los FS los que mayormente forman el vellón. En ovejas se han identificado un tercer tipo, los folículos secundarios derivados (FSD) (Hardy *et al.*, 1956), mientras en alpaca se ha identificado solo dos tipos clásicos (Bustinza, 2001; Schneider *et al.*, 2009; Lacolla *et al.*, 2010).

En general todos los camélidos sudamericanos presentan FP simples y nidos foliculares. En la alpaca los FP están acompañados de una glándula sudorípara (Bustinza, 2001) mientras que todos los FS desembocan en un infundíbulo común conjuntamente y se relacionan con la glándula sebácea, la glándula sudorípara y el músculo erector (Torres de Jasauí *et al.*, 2007; Bustinza, 2001).

En los nidos foliculares de los camélidos sudamericanos el FP tiene una ubicación excéntrica con respecto a los FS que lo acompañan, muchas veces el FP de los nidos son más grandes que los FS y otras veces todos son de un tamaño similar, este nido tiene un número variable de folículos según la zona corporal y oscila entre 5 a 50, donde en la cabeza del guanaco y el costillar de la vicuña presentan las mayores agrupaciones (Lacolla *et al.*, 2010), esta agrupación de folículos se compactan y fusionan (Bustinza, 2001). En la alpaca el FP no está rodeado completamente de los FS si no que se encuentra a un lado de ellos (Bustinza, 2001). En todos los camélidos sudamericanos los FP presentan un diámetro mucho mayor que los FS y son los FS

los que determinan el volumen del vellón (Lacolla *et al.*, 2010; Torres de Jasauí *et al.*, 2007) al igual que en ovejas (Hardy *et al.*, 1956).

En las ovejas Merino tiene una población de folículos muy alta, de 10 a 100 millones comparado con los 5 millones de folículos en piel de humanos. Los FP son los primeros en formarse, esta es una invaginación de la epidermis que en ovejas aparecen entre los 50 a 75 días del feto, seguido de los FS que aparecen entre los 85 a 135 días del feto, por último se forman los FS ramificados o también llamados FS derivados (FSD) que se forman a los 105 días. Esta ramificación de los FSD determina la densidad de la población de los folículos, la mayoría de los folículos son FSD, un aproximado del 80% del total además que muchas fibras emergen del mismo orificio de la piel (Hardy *et al.*, 1956).

2.3.2. El eje del pelo: La fibra del pelo.

El eje del pelo o también llamada fibra de pelo se compone de queratinocitos muertos (tricocitos) los cuales están compactados en una fibra, esta fibra o eje del pelo es generada por el folículo piloso.

La fibra de alpaca, al igual que en otros mamíferos, presenta 3 capas: la cutícula que es la más extrema, la corteza y la médula que es la más interna.

La cutícula presenta células planas poligonales y se sobreponen unas a otras formando estructuras llamadas escamas. Existen diferencias entre las fibras del tipo Suri y Huacaya, en el exterior de la cutícula en el tipo Suri las fibras son más suaves mientras que en Huacaya son ásperas, sus diferencias en las propiedades de fricción se debe a la conformación de las células de la cutícula que forman las escamas de la fibra, a medida que el diámetro aumenta el tamaño de las escamas disminuye (Bustinza, 2001).

La proporción de la corteza varía según el diámetro de la fibra, las fibras más finas en alpacas no presentan médula por lo que la corteza forma más del 90% de la masa de la fibra, esta característica es similar a la lana fina de ovinos, mientras que en las fibras gruesas encontramos las 3 capas, la médula, la corteza y la cutícula donde la corteza forma menos del 50% de la masa de la fibra, entre estos dos extremos hay muchos casos intermedios (Bustinza, 2001).

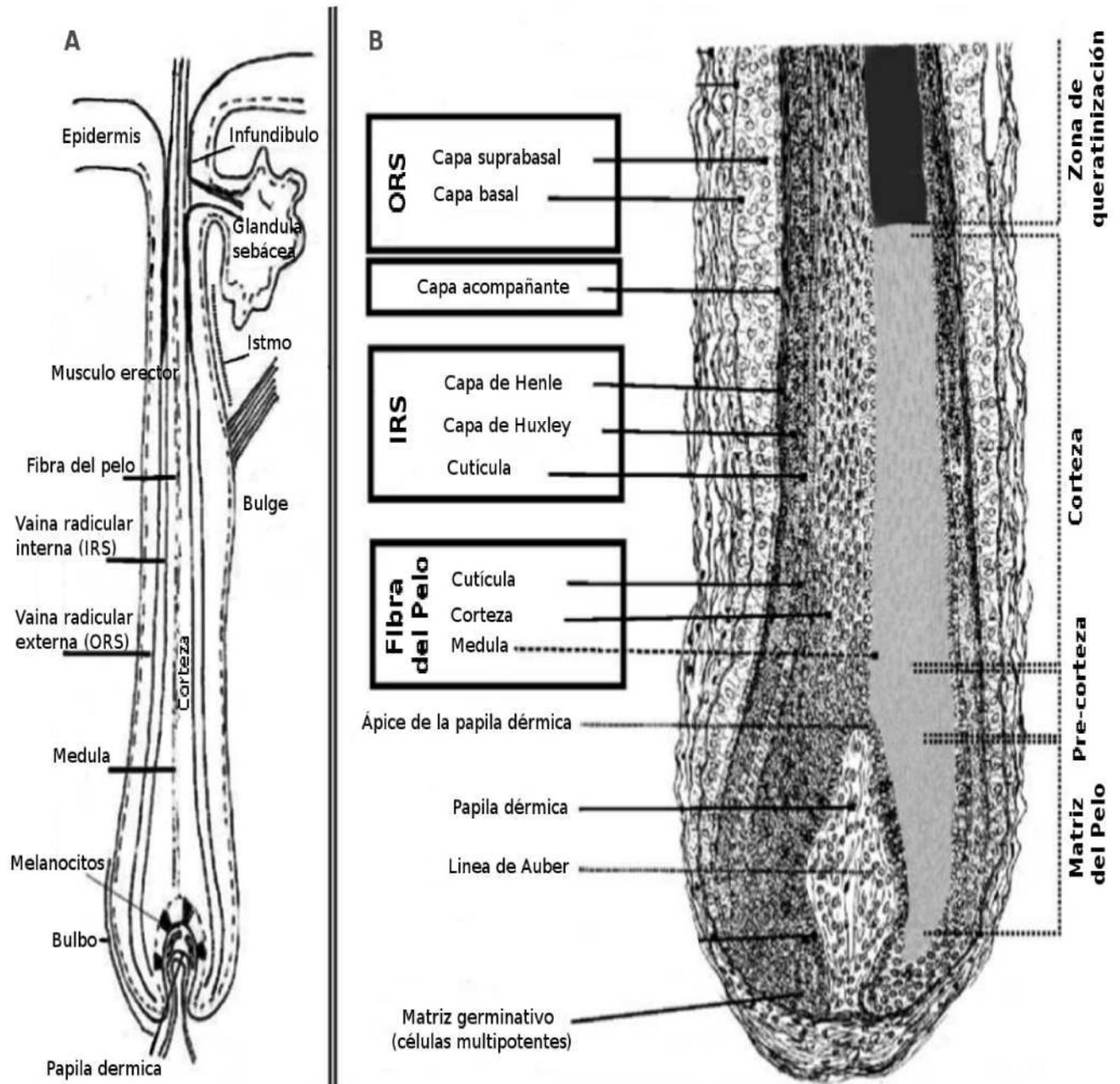


Figura 1. Morfología del folículo piloso. (A) Ilustración de los mayores componentes estructurales en la fase de anagen. (B) Ilustración detallada de las partes del folículo piloso (figura tomada y modificada de Langbein y Schweizer, 2005).

Una característica aparentemente envuelto en el rizado de la fibra es la estructura bilateral de la fibra, muchas veces bien definida, la ortocorteza y la paracorteza (Figura 2). La ortocorteza es la parte convexa o la región exterior que consiste en un tejido denso de células mientras que la paracorteza es la parte cóncava o región interior de la curvatura de la fibra que consiste en una densidad menor de células. En la oveja se ha identificado estas dos estructuras gracias a su tinción con azul de metileno donde la ortocorteza se tiñe más fuertemente que la paracorteza (Rogers, 1959).

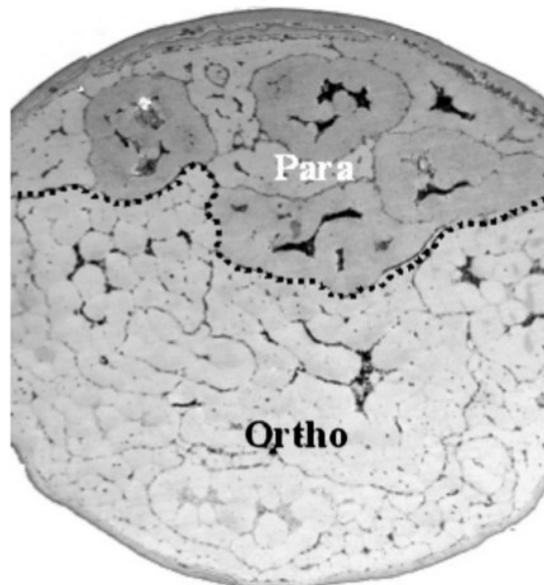


Figura 2. Fotografía de la estructura bilateral en la lana de oveja mostrando la ortocorteza y la paracorteza (figura tomada de Rogers *et al.*, 2006).

En alpacas si hay una estructura definida en la corteza, en la fibra del tipo Huacaya que se caracteriza por ser más rizada que en el tipo Suri, está bien definida la región de la ortocorteza y la paracorteza, viéndose una mayor diferenciación de estas regiones en las fibras de mayor rizado, esta característica es similar en ovinos, también en las fibras más finas y de mayor rizado se observa una mayor diferenciación, en las fibras finas (menores a 25 μm) con médula y sin médula, la estructura entre orto y para son similares, en las fibras medias (entre 25 a 35 μm) estas regiones se distribuyen de forma variable y menos nítida por la tinción, en las fibras gruesas (mayores a 40 μm) no se distingue bien la región fuertemente teñida, la ortocorteza (Bustinza, 2001) mientras que en la fibra del tipo Suri que se caracteriza por ser más lacias y lustrosas con una superficie suave, tiene una menor afinidad a los

tintes, en las fibras finas no se diferencia bien las regiones orto y para, así también en las fibras medias y gruesas no se distingue bien la región tipo orto por lo que se considera que las fibras en el tipo Suri son mayoritariamente formadas por la paracorteza (Bustinza, 2001).

La alpaca es uno de los animales con una gran diversidad en colores, tiene 22 coloraciones desde el blanco puro hasta cremas, marrones y negras. El proceso de coloración del pelo en los mamíferos es muy conservada, esta es controlada por el tipo y la cantidad de del pigmento melanina, éste es producido dentro de los melanocitos, en la organela melanosoma, estas células están localizados bajo la piel y están relacionados con los queratinocitos del pelo. Los melanosomas son transportados desde los melanocitos hasta los queratinocitos por medio de las protuberancias o dentritas que existen en los melanocitos. Se distinguen dos tipos de melaninas en los mamíferos: la eumelanina la cual está involucrada desde el color negro al marrón y la feomelanina involucrada desde el color amarillo al rojo.

Al parecer existe una relación entre el color de la fibra con el tamaño del diámetro. Las fibras de color claro en la alpaca presentan un diámetro menor que las fibras más oscuras, mientras más aumenta el tono oscuro en la fibra el diámetro tiende a aumentar y por ende el confort que está relacionado al diámetro disminuye así también va acompañado de una disminución del índice de curvatura (McGregor y Butler, 2004; Oria *et al.*, 2009).

2.4. Factores no genéticos que influyen en la fibra.

La edad influye mucho en el diámetro de la fibra de alpaca, cuando son menores de un año de edad, es cuando presentan la fibra más fina conocido como "baby alpaca" (de 20 a 23 μm), incluso algunas poblaciones tienen un pool genético que le permite tener una fibra de mejor calidad (menor a 20 μm) considerado como "royal alpaca" (Tabla 1). La primera esquila es generalmente al año ya que en esta etapa es la más cotizada por su mejor calidad, luego de la primera esquila y conforme va aumentando su edad, el diámetro de la fibra aumenta considerablemente hasta los 5 años, luego de ese tiempo aumenta a menor velocidad (Bustinza, 2001). Similares características se encuentran en ovinos y cabras, cuando recién nacen su tasa de crecimiento de la lana se incrementa hasta 3 o 4 años después decae y su diámetro aumenta (Summer y Bigham, 1993).

Tabla 1. Clasificación de la calidad de fibra de alpaca en el mercado. Según Norma Técnica Peruana (NTP 231.301:2004).

Clasificación	Diámetro	Longitud
Alpaca royal	Hasta 21,5 μm	65 mm
Alpaca baby	21,6 a 23,0 μm	65 mm
Alpaca fleece	23,1 a 26,5 μm	70 mm
Alpaca médium fleece	26,5 a 29,0 μm	70 mm
Alpaca huarizo	29,1 a 31,5 μm	70 mm
Alpaca gruesa	Más de 31,5 μm	70 mm
Alpaca corta	-----	20 a 50 mm

Respecto al sexo no se ha establecido una relación clara, algunos indican que ejerce una diferencia mínima en el diámetro de la fibra (McGregor y Butler, 2004) mientras otros no encuentran una relación (Badajoz, 2007). Una relación en cuanto al sexo y al clima se observa en puna húmeda, un estudio reporta que en la puna húmeda el diámetro entre machos y hembras presentan diferencia significativa (22,5 μm en machos y 22,8 μm en hembras) contrario en la puna seca (22,7 μm en machos y 22,8 μm en hembras) que no presenta una diferencia significativa, así también el diámetro se incrementa más en la puna seca que en la puna húmeda entre la primera y segunda esquila (Huanca et al., 2007), sin embargo este estudio no determina la causa directa de la variación del diámetro, lo que muy probablemente se deba a los diferentes nutrientes en estas dos zonas geográficas.

La nutrición también juega un rol importante, en ovinos una mejor nutrición da como resultado un aumento de la longitud y el diámetro de la lana (Summer y Bigham, 1993), el diámetro de la fibra está relacionado con el tamaño del bulbo folicular, por otro lado la longitud de la fibra está relacionada con el tamaño del bulbo folicular, con la longitud de las células corticales y la proporción de las células que ingresan a la propia fibra (Hynd y Masters, 2002), también existe una diferencia entre los ovinos de fibra fina y gruesa, los ovinos de fibra gruesa son más susceptibles a la nutrición comparados con los de fibra fina, esto solo en cuanto a su diámetro, en este caso estas características serían dictadas por el genotipo (Hynd y Masters, 2002). Por otro lado se han comparado alpacas en condiciones deficientes de nutrientes, debajo de los requerimientos de mantenimiento y con el doble de los requerimientos, se observó que la producción de fibra de alpaca es muy sensible a la manipulación nutricional y el

efecto de la nutrición en el peso de la fibra es más acentuada en el largo de la fibra que en el diámetro, siendo esta última no significativa (Russel y Redden, 1997).

2.5. La fibra del pelo y sus principales componentes moleculares.

El estudio de los genes involucrados en la fibra de la lana se vuelve complicado por la pleiotropía de los genes responsables del fenotipo (Adelson *et al.*, 2004). Numerosos estudios se han hecho en humanos, ratones y ovejas principalmente, más en alpacas u otro camélido existe muy poca información.

Las proteínas queratinas son los mayores componentes del folículo piloso de los mamíferos y son responsables de la mayoría de las propiedades estructurales. Estas se dividen en 2 tipos, los filamentos intermedios de queratinas (KIFs) y las proteínas asociadas a queratinas (KAPs) las cuales están ampliamente estudiadas (Rogers, 2004).

2.5.1. Los filamentos intermedios de queratinas (KIFs) y las proteínas asociadas a queratinas (KAPs).

Las queratinas están presentes en dos tipos de los filamentos intermedios (IFs), los IF de tipo I y tipo II, sin embargo dentro de la fibra del folículo, solo se expresan algunas las cuales en este estudio se las nombrará como KIFs relacionados al pelo o fibra, hKIF (hair keratin intermediate filament) (Tabla 2). En humanos está bien estudiado y se conocen 17 genes de queratinas del folículo, 11 del tipo I y 6 del tipo II, en ovejas, en los últimos años, se han identificado más queratinas, llegando a igualar en número al de los humanos, 17 genes de queratinas, 10 del tipo I y 7 del tipo II, el gen *KRT37* solo se ha identificado en humanos mientras que el *KRT87* solo en oveja lo que muestra genes especie-específicas (Yu *et al.*, 2011). Esta familia es muy conservada en los mamíferos.

Tabla 2. Clasificación de los genes de queratinas de los mamíferos. Las proteínas son designadas como una K seguida del número correspondiente, mientras que los genes como *KRT* (Schweizer *et al.*, 2006).

Ubicación	IF tipo I	IF tipo II
Epiteliales	<i>KRT9, KRT10</i>	<i>KRT1 – KRT5</i>
	<i>KRT12 – KRT20</i>	<i>KRT6a – KRT6c</i>
	<i>KRT23 – KRT25</i>	<i>KRT7, KRT8</i>
		<i>KRT71 – K80</i>
Fibra del folículo	<i>KRT31, KRT32</i>	<i>KRT81 – KRT85</i>

Por otro lado las proteínas asociadas a queratinas (KAPs) son mucho más numerosas (Tabla 3), dividiéndose en 3 grupos, las proteínas que contienen entre 35 a 60 mol% de glicina y tirosina nombradas como proteínas ricas en glicina y tirosina (HGT, high glycine-tyrosine proteins), las proteínas ricas en sulfuro (HS, high sulphur proteins) las cuales contienen menos de 30 mol% de cisteína y las ultra ricas en sulfuro que contienen valores mayores a 30 mol% de cisteína (UHS, ultra-high sulphur proteins). Estas proteínas han sido puestas en 27 familias y a su vez están subdivididas, por ejemplo la familia KAP18 de humanos se divide en 12 miembros (KAP18-1 hasta KAP18-12) (Gong *et al.*, 2012). Se han identificado más de 100 KAPs en los mamíferos, en humanos se han identificado 89 KAPs funcionales puestos en 25 familias mientras que en oveja se han identificado tan solo 13 KAPs en 7 familias.

La expresión de las KAPs solo se han encontrado en células del folículo piloso envueltas en la matriz proliferativa del bulbo la cual da la formación a la fibra y también se encontró en las células de la corteza y la cutícula, además la mayoría de estas proteínas presentan polimorfismos (Rogers, 2004). Los genes se representan por KRTAP (Gong *et al.*, 2012).

Tabla 3. Clasificación de los genes codificantes de las proteínas asociadas a queratinas en mamíferos (Gong *et al.*, 2012).

Grupo	Familia
HS	KRTAP1 – KRTAP3, KRTAP10 – KRTAP16, KRTAP23 – KRTAP27
USH	KRTAP4, KRTAP5, KRTAP9, KRTAP17
HGT	KRTAP6 – KRTAP8, KRTAP19 – KRTAP22

En la lana de oveja se ha observado que tipo de queratinas y donde empiezan a expresarse en las diferentes capas de la fibra (Figura 3), habiendo una alta expresión de los genes *KRT34* y *KRT36* en la médula (Yu *et al.*, 2011), mientras que en los humanos los genes *KRT33A*, *KRT33B*, *KRT34*, *KRT36*, *KRT37* y *KRT81* son las que más se expresan en la médula (Langbein y Schweizer, 2005). En general el patrón de expresión de las queratinas es muy similar en ovejas y humanos, no solo en

los tipos de queratinas que se expresan en cada capa de su estructura si no en el lugar donde empieza a expresarse.

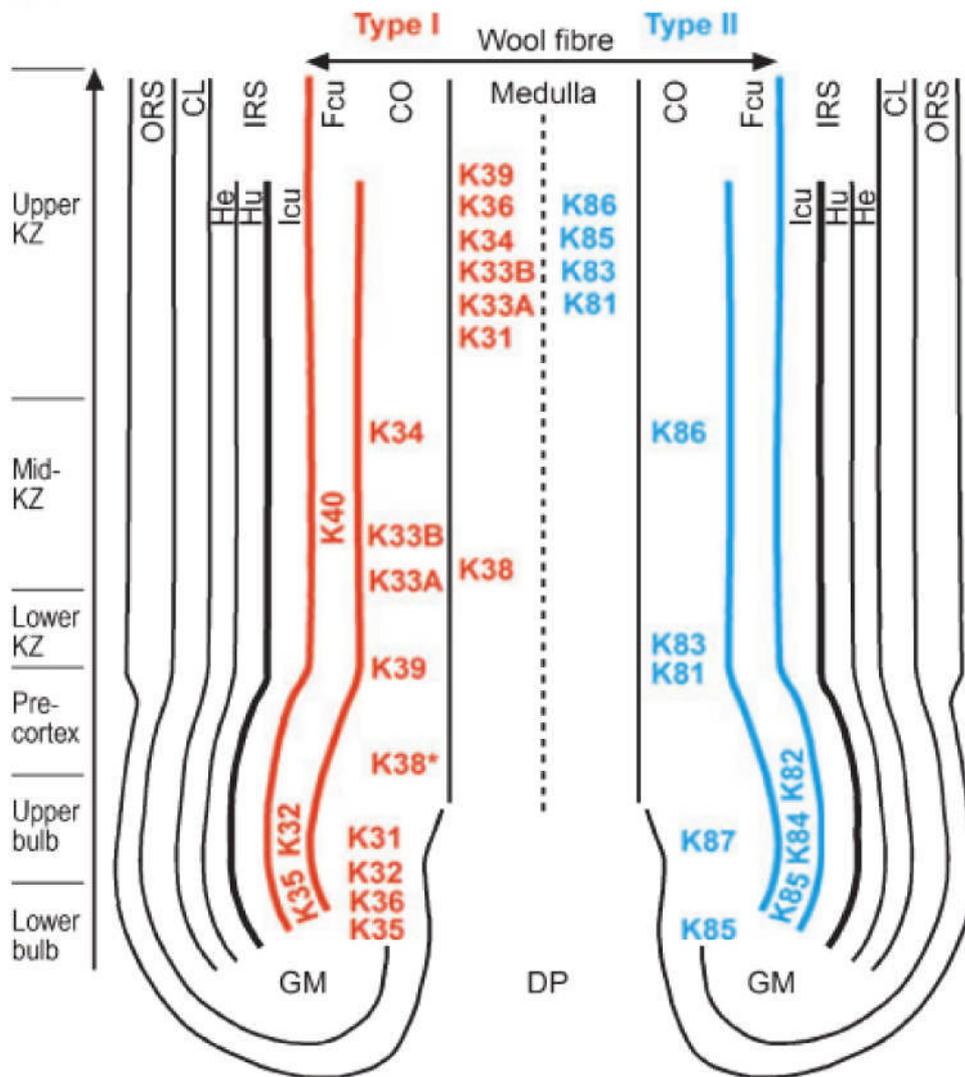


Figura 3. Diagrama de la expresión de las queratinas en el folículo de la lana de oveja. La posición de los nombres indica el lugar de inicio de su expresión. ORS, vaina radicular externa; CL, capa acompañante; IRS, vaina radicular interna; He; Capa de Henle; Hu, capa de Huxley; Icu, cutícula de la IRS; Fcu, cutícula de la fibra; CO, corteza; DP, papila dermica; GM, matriz germinativo (figura tomada de Yu *et al.*, 2011).

En la estructura bilateral orto-para, la organización de las queratinas es diferente en los dos tipos de células e indican que son el producto de dos vías diferentes, se conoce que la expresión de HGT son relativamente más altas en la ortocorteza que en la paracorteza, mientras que las HS y UHS se expresan un poco más en la paracorteza (Figura 4) sin embargo la función que ejerce esta estructura bilateral sobre la curvatura todavía no es clara ya que en ovejas no Merinos, que

producen una fibra más gruesa mayor a 20 μm , está ausente o menos desarrollada y en humanos está ausente (Rogers, 2004; Rogers, 2006).

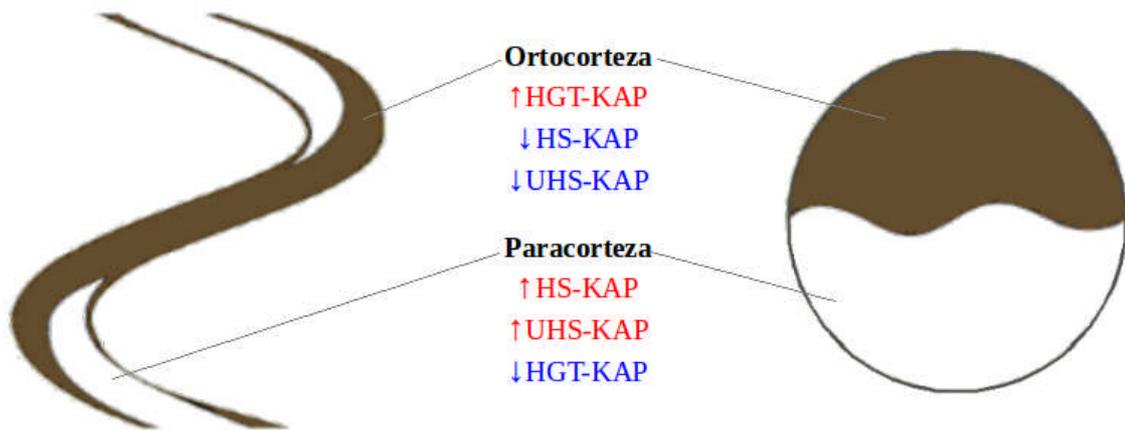


Figura 4. Diagrama transversal de la fibra de lana de la oveja, mostrando la distribución de las KAPs en la orto y paracorteza. En la ortocorteza hay una mayor expresión de las KAPs ricas en glicina y tirosina (HGT) mientras que en la paracorteza hay una mayor expresión de las KAPs ricas en sulfuro (HS y UHS).

2.6. Aspectos moleculares de la morfogénesis del folículo piloso.

El folículo piloso tiene dos tipos de células, las células de la papila dérmica, las cuales derivan del mesénquima y las células epiteliales de la vaina dérmica y de la fibra del pelo, derivadas de la superficie del epitelio. El prerequisite para el desarrollo del folículo piloso es la comunicación molecular epidermis-dermis. El desarrollo del folículo empieza durante la embriogénesis y requiere de señales entre las células dérmicas mesenquimales y las células epiteliales superficiales (Figura 5), los cuales causan cambios en los dos tipos de células, resultando en la diferenciación de la fibra del pelo, las vainas de la raíz y la papila dérmica (Hardy, 1992).

La primera señal surge de la dermis causando la formación de la placoda en las células epiteliales, por lo que se le llama "**Primera Señal Dérmica**" y genera el mensaje "**hacer una placoda**" o apéndice. Posteriormente se genera una "**Señal Epitelial**" de la placoda la cual forma una agrupación de células subyacentes en las células mesenquimales formando un condensado dérmico lo que será el brote del folículo, por lo que el mensaje es "**hacer una papila dérmica**"; la "**Segunda Señal Dérmica**" es transmitida desde la papila dérmica hacia las células epiteliales adyacentes de la matriz germinativa, causando la proliferación de las células

epiteliales de la placoda e invaginándose hacia la dermis, por lo que el mensaje será **"hacer un folículo piloso"**, las células hijas de la matriz posteriormente se mueven hacia arriba del folículo diferenciándose en lo que será la vaina radicular interna (IRS, inner root sheath), luego hay una queratinización interna que emerge con 3 capas, la cutícula, la corteza y la medula (Hardy, 1992).

Una serie de interacciones moleculares ocurren durante estas señales. Esas señales que lo controlan incluyen proteínas secretadas por la vía WNT/Wingless, la vía de Sonic Hedgehog (SHH) (Figura 6) y miembros de la familia BMP (proteína morfogénica ósea), TGF- β (factor de crecimiento transformante β), FGF (factor de crecimiento de fibroblasto) y TNF (factor de necrosis tumoral) (Schneider *et al.*, 2009). Algunas señales promueven el desarrollo folicular (señales Wnt, FGF) mientras otras lo inhiben (señales BMP) (Millar, 2002).

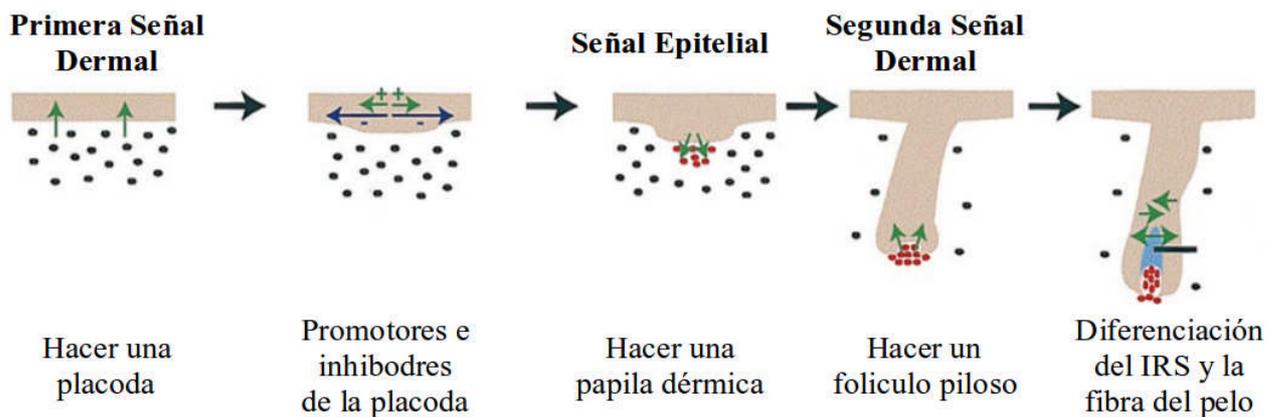


Figura 5. Morfogénesis del folículo piloso, mostrando las señales dermal-epiteliales. La Primera Señal Dermal inicia la formación de la placoda, donde van a interactuar genes que van a promover e inhibir su formación. En respuesta hay una Señal Epitelial la cual inicia la formación de la papila dérmica. En la Segunda Señal Dermal se inicia la formación del folículo piloso, posteriormente hay una migración y diferenciación celular en la cual se van a formar la vaina radicular interna (IRS) y la fibra del pelo (figura tomada y modificada de Reddy *et al.*, 2001).

En la Primera Señal Dérmica existe una competencia entre los promotores e inhibidores. Esta envuelve la activación de la señal EDA/EDAR (molécula relacionada al factor de necrosis tumoral y su receptor) en el epitelio seguido de la señal epitelial WNT y la posterior activación de la señal BMP, donde la activación de EDA/EDAR y WNT promueve la formación de la placoda, mientras que la señal BMP la reprime, la señal FGF y TGF- β 2 también la promueven su formación (Millar, 2002).

La vía de señalización WNT es una de las más importantes vías requeridas para la organogénesis embrionaria, la mutación de los genes WNT causan serios cambios fenotípicos y está involucrado a la génesis del cáncer. Existen 3 vías de WNT, la vía Wnt/Ca²⁺ el cual involucra la activación de la fosforilaza C (PLC) y conduce a una elevada expresión de los niveles de Ca²⁺ y la activación de la proteína quinasa C (Kühl *et al.*, 2000), la vía de Wnt/PCP (planar cell polarity) que actúa a través de las proteínas GTPasas como RhoA y Rac y activa la quinasa c-Jun amino-terminal (JNK, c-Jun amino-terminal kinase) (Peifer y McEwen, 2002) y la vía WNT/ β -catenina (o Canónica WNT), la más estudiada, involucrada con regulación de la estabilización y la expresión de la β -catenina la cual participa en un complejo co-activador en el núcleo con factores de transcripción de la familia LEF/TCF y otras para regular la expresión (Cadigan y Nuse, 1997, Cadigan y Liu, 2006, Willert y Jones, 2006)

Normalmente la β -catenina se degrada rápidamente en el citoplasma, pero cuando se activa la vía canónica WNT gracias a su interacción con miembros de la familia de receptores Frizzled (FZD) y los coreceptores LRP (Low-density lipoprotein receptor-related protein) inhibe su degradación (Millar, 2002). La activación de esta vía ocurre en varios procesos de organogenesis y ocurre también en la Primera Señal Dérmica, esta activación resulta en la fosforilación de LRP y la inactivación de la GSK3- β (Glycogen synthase kinase three beta) responsable de la fosforilación de la β -catenina para su degradación, por lo que se acumula en el citoplasma y pasa hacia el núcleo donde forma complejos activos con miembros de la familia de factores de transcripción LEF/TCF (lymphoid enhancer-binding factor/T cell factor) lo que conduce a modular muchos genes que se involucran al ciclo y a la proliferación celular (Millar, 2002; Andl *et al.*, 2002; Filipovich, 2011).

La vía WNT/ β -catenina es esencial en la formación del folículo piloso, envuelto en diferentes estados de la morfogénesis incluyendo la formación de la placoda, condensación del mesenquima, diferenciación del eje del pelo hasta las fases de su ciclo (Millar 2002), no solo la activación de esta vía sino también la disminución de su expresión es esencial para el éxito de la morfogénesis del pelo (Närhi, 2008). Esta señal es requerida para la iniciación de la formación de todos los tipos de folículos pilosos, la Primera Señal Dermal (Andl *et al.*, 2002). Las células madres progenitoras en la matriz del folículo son las que mantienen activa la vía WNT y la estabilización de la β -catenina durante la fase de anagen, mientras que en la matriz de la precorteza inhiben la proliferación e inician la diferenciación terminal (Schneider *et al.*, 2009).

La proteína Dkk1 (proteína relacionada a Dickkopf) es un potente inhibidor de esta vía, inhibe los coreceptores LRP requeridos para la activación de la vía Wnt/ β -catenina. La expresión ectópica de Dkk1 en la piel de ratones transgénicos bloquearon la formación de la placoda antes de la señal de diferenciación, indicando que esta señal precede y es requerida para la activación de otros reguladores para el desarrollo del folículo piloso (Andl *et al.*, 2002) mientras que la expresión forzada de la β -catenina resulta en una acelerada y excesiva formación de la placoda, en la cual se observa que se estimula las proteínas BMP. La actividad de la vía WNT/ β -catenina en el epitelio causa una disminución de la expresión de la misma vía en el mesenquima la cual induce, directa o indirectamente, a la expresión de algunos genes de la familia BMP, por lo que las vías WNT/ β -catenina y BMP podrían estar envueltas en el correcto espaciado del folículo piloso (Närhi *et al.*, 2008).

Hay que notar que la maquinaria diseñada para la formación del folículo piloso en mamíferos difiere en algunos aspectos a la maquinaria para la formación de los folículos de la vibrisa, por ejemplo el factor de transcripción Lef1 es esencial para la iniciación del folículo de la vibrisa en ratón mientras que no lo es para la formación del folículo piloso, lo que sugiere que la vía Wnt/ β -catenina es requerida en diferentes estados en el desarrollo del folículo piloso y el folículo de la vibrisa (Millar, 2002).

Otra vía de señalización importante es la SHH. A diferencia de la vía WNT/ β -catenina, la vía SHH no es esencial para la iniciación del desarrollo del folículo, pues su expresión se da en fases posteriores donde su expresión es esencial para la regulación de la proliferación folicular y formación de la papila dérmica (St-Jacques *et al.*, 1998).

En ausencia de la vía SHH, su receptor Patched1 (PTC1) inhibe la actividad de Smoothed (SMO) y los factores de transcripción GLI son fosforilados por la proteína quinasa A (PKA) por lo que se inhibe su activación. En presencia de SHH, la PTC1 se desliga de SMO, la activación de SMO permite inhibir la fosforilación de la familia de transcripción GLI, las proteínas GLI1 y GLI2. La actividad de los factores de transcripción GLI es también regulada por la inhibición de SHH, cuando se activan las proteínas GLI uno de sus blancos es la transcripción de PTC1, en cual, cuando esta sobreexpresado, atenúa la respuesta de SHH previniendo la interacción entre SHH y su receptor PTC1 (Ruiz i Altaba *et al.*, 2002; Dorsam y Gutkind, 2007).

Una de las fases de crecimiento más importantes es la anagen, donde parte de su regulación ocurre a través de la regulación de la acumulación del factor de transcripción GLI. El efecto de la vía SHH en las células multipotentes epiteliales en la piel son controladas por dos caminos: restringiendo cuando y donde la señal es enviada y permitiendo la recepción de la señal solo en etapas particulares del ciclo de vida del folículo. PTC1, GLI1, GLI2 y GLI3 se acumulan en la papila dérmica en la anagen, la expresión de estos genes sirven como indicadores para ver si la vía SHH se ha activado, en la catagen la expresión de estos decrece mientras que en la telogen no se ha detectado (Oro y Hoggins, 2003). Parte de la regulación de la vía SHH es gracias a la vía WNT/ β -catenina, ya que la expresión de SHH en el folículo piloso está regulada por algunos miembros de la vía WNT/ β -catenina (Reddy *et al.*, 2001).

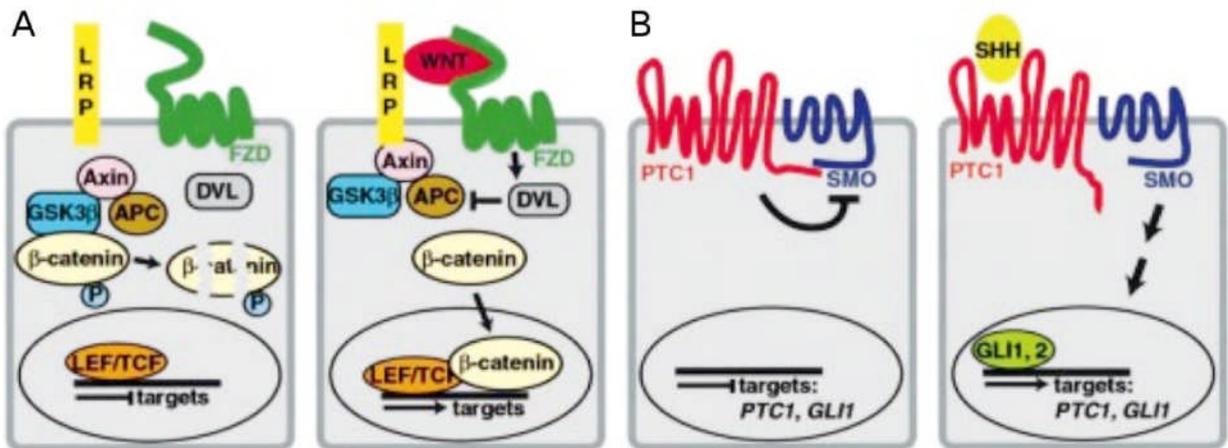


Figura 6. Esquema de las vías de señalización WNT y SHH. (A) En la ausencia de la vía WNT, la β -catenina es fosforilada y degradada por un complejo de proteínas. Con la activación de la vía WNT, la maquinaria de degradación es inhibida y la β -catenina es transportada al núcleo donde forma un complejo de transcripción con factores ligantes al DNA. (B) En ausencia de la vía SHH, la actividad de SMO (Smoothened) es inhibida por acción del receptor de la vía SHH, Patched1 (PTC1). Con la activación de la vía SHH, se activa el SMO y se transcriben genes incluyendo PTC1 y GLI1 (figura tomada de Millar, 2002).

2.7. Los microsatélites en las regiones codificantes.

Los microsatélites o repeticiones de secuencias simple (SSR) son motivos de 1 a 6 nucleótidos, distribución en el genoma no es aleatoria, probablemente por sus efectos en la organización de la cromatina, regulación de la actividad génica, recombinación, DNA replicación, ciclo celular, sistema de reparación mismatch (MMR), pudiendo proveer una ventaja evolutiva para la rápida adaptación en nuevos ambientes (Li *et al.*, 2002, Li *et al.*, 2004). La frecuencia con que se presentan estas repeticiones en las regiones codificantes parece está limitado para que no se interrumpa el marco de lectura ya que en los diferentes mamíferos el motivo más presente son los tripletes (Wren *et al.*, 2000). La frecuencia con la que se encuentran los tripletes en las regiones codificantes es el doble con las que se encuentran en los intrones y las regiones intergénicas, excepto en el cromosoma Y (Subramanian *et al.*, 2003).

El efecto de los SSR en las regiones codificantes ha sido ampliamente estudiado en las enfermedades, la expansión y contracción de estas repeticiones son críticas para la actividad génica, enfermedades como en Huntington, atrofia dentato-rubro-pálido-luisiana, atrofia muscular espinobulbar y ataxia espinocerebelosa todas son causadas por la expansión del codon CAG codificante para el aminoácido glutamina. Se ha encontrado que muchos cáncer son causados por la inactivación de algunos genes MMR (hMSH2; hMLH1, hMSH6) esto ocurre por el cambio de lectura causado por (A)_n en los exones (Duval *et al.*, 2002).

En la gran familia de las KAPs se ha observado que son proteínas muy polimórficas, principalmente causado por un número variable de regiones ricas en cisteína (Shimomura *et al.*, 2005), las frecuencias alélicas de algunos miembros de las KAP1 y KAP4 muestran alelos polimórficos específicos para algunas poblaciones humanas (Shimomura *et al.*, 2002; Shimomura *et al.*, 2005). Así también en algunas queratinas presentan polimorfismos, por ejemplo la K10 presenta un codón polimórfico (GGC)_n que codifica la glicina, donde se reconocen variaciones de 114 nucleótidos o 38 aminoácidos (Korge *et al.*, 1992). Solo los genes *KRTAP6* y *KRTP8*, muestran alelos que los han asociado al tamaño del diámetro en la lana de ovejas (Parsons *et al.*, 1994b).

La búsqueda de marcadores SSR en regiones codificantes es una gran herramienta que permitirá hacer estudios para ver si existe una relación entre los genes que están involucrados a la formación de la fibra y sus polimorfismos.

2. 8. Marcadores de secuencias expresadas (EST).

Uno de las formas para estudiar que genes están involucrados en determinada característica y/o respuesta en un eucarionte es mediante el estudio de todo el conjunto de genes que están codificando proteínas en determinado momento y tejido, esto es el transcriptoma (el ARN mensajero). En organismos no modelos, como es el caso de la alpaca, el estudio de transcriptomas resulta un método rápido y barato comparado con el secuenciamiento de genomas, que es muy caro y demasiado laborioso, sobre todo en organismos con un genoma grande, como plantas superiores, peces, aves y mamíferos quienes tienen una baja densidad de genes a comparación de los procariontes (Nielsen, 2011).

En este aspecto existen varios métodos para analizar el transcriptoma, uno de los métodos son la generación de Marcadores de Secuencias Expresadas (EST), estas son lecturas de fragmentos de secuencia de ARNm derivados de un solo paso de secuenciación generado de librerías de clones aleatorios de ADN complementario (ADNc), estos representan una porción expresada del genoma, importante para identificar genes con función conocida, genes nuevos y nuevas variantes de transcritos por splicing alternativo (Parkinson, 2009).

El proceso para generar una librería de ESTs empieza con la purificación de ARNm y luego transformado a ADN complementario (ADNc), insertándose en un vector para generar la librería de ADNc; clones individuales son seleccionados aleatoriamente y pasados por un paso de secuenciación obteniéndose la secuencia en dato crudo que se puede analizar por computadora. El resultado se analiza con varios pasos, analizando la calidad de cada secuencia, enmascaramiento de la cola poli-A, un mínimo de longitud y posteriormente subirlo a alguna base de datos, como la del NCBI (*National Center of Biotechnology Information*) este último es el más utilizado, después de este primer proceso del dato crudo de secuenciamiento se procede a reconstruir el mRNA, un paso llamado ensamblaje de ESTs, cuyo objetivo es reducir la redundancia de secuencias y obtener la mayor longitud del mRNA, después una búsqueda de homologías y su anotación funcional (Parkinson, 2009).

Actualmente se tiene una vasta información sobre los factores tanto morfológico, citológico como moleculares sin embargo este conocimiento está más estudiado en humanos y en menor grado en otros organismos modelos como las ovejas, ratones y otras especies de importancia para el hombre. En organismos no modelos como las alpacas se tiene un conocimiento tanto morfológicamente como

citológicamente de algunas características de la fibra, sin embargo casi nada se sabe sobre los factores moleculares que están involucrados.

Los estudios referentes a los genes que están involucrados al desarrollo de la fibra del pelo en diversos organismos se encuentran en aumento, en los últimos años se han descubierto nuevos alelos en diferentes organismos y se han descubierto nuevas funciones por lo que se ha re-annotado nuevos genes y proteínas, esto favorece a los nuevos análisis *in silico* ya que se cuenta con una mayor información en las bases de datos para la comparación.

En los resultados se reconocieron los transcritos que están involucrados con la síntesis de la fibra en piel de alpaca así también se detectaron marcadores SSR-ESTs en estos genes para que futuras investigaciones puedan determinar si estas tienen un rol importante para algunas características de la fibra como son su diámetro o longitud, así también estos marcadores podrán transferirse a otras especies cercanas como la vicuña para poder hacer estudios de comparación entre las diferentes características de las fibras.

3. HIPÓTESIS

Hipótesis.

En la librería de ESTs de piel de alpaca se puede asignar de manera *in silico* nuevas funciones y caracterizar marcadores SSR asociados a la síntesis de fibra a través de la comparación de bases de datos del NCBI de nucleótidos, proteínas y del Gene Ontology.

Hipótesis nula.

En la librería de ESTs de piel de alpaca no se puede asignar de manera *in silico* nuevas funciones y caracterizar marcadores SSR asociados a la síntesis de fibra a través de la comparación de bases de datos del NCBI de nucleótidos, proteínas y del Gene Ontology.

4. OBJETIVOS

Objetivo general:

Reconocer transcritos asociados a la síntesis de fibra a través de diferentes bases de datos y caracterizar marcadores EST-SSR en los genes asociados a la síntesis de fibra de alpaca.

5. MATERIALES Y MÉTODOS

5.1. Materiales.

La librería de ESTs utilizada fue extraída de la base de datos del Centro Nacional para la Información Biotecnológica (National Center for Biotechnology Information, NCBI), pertenecientes a la piel de la región del cuello, donde se encuentra gran densidad folicular, de una alpaca hembra blanca del tipo Huacaya de 3 años de edad de una granja de la Universidad de Agricultura de Shanxi (Shanxi, China) publicados por Fan *et al.*, 2011. Se extrajeron todos los ESTs de alpaca obteniendo un total de 7286 secuencias (Tabla 4). Las secuencias escogidas del NCBI se encuentran preprocesadas, es decir que fueron eliminadas las secuencias del vector y enmascaradas las zonas de baja calidad y repetitivas, por lo que no se tuvo que hacer estos pasos y se procedió a ensamblar las secuencias. Toda la metodología se realizó en el sistema operativo GNU/Linux en una computadora con características de corei7 de 16 GB de memoria RAM.

Tabla 4. Número de ESTs de piel de cuello de alpaca huacaya hembra blanca usados en el presente estudio extraídas de la base de datos del NCBI. (Fan *et al.*, 2011).

Número de Accesoión	Número de secuencias
EH218976-EH219025	50
ES263542-ES263641	100
ES444048-ES444167	120
ES605553-ES605692	140
ES608886-ES609085	200
ES768454-ES768853	400
EV254357-EV254756	400
EV552651-EV554902	2252
EX159842-EX162087	2246
EX656048-EX656750	703
EY413719-EY414393	675
Total	7286

5.1. Métodos.

Para determinar las funciones de los genes de más interés, el ranking de expresión o genes más expresados y la búsqueda de microsatélites dentro de los genes (EST-SSR) se siguió la siguiente metodología: se ensamblaron las secuencias ESTs para reducir la redundancia y obtener la mayor longitud del ARNm, secuencia llamada unigen, a partir de este paso se pudo determinar la distribución de ESTs en cada unigen y determinar un ranking de expresión, así también se hizo la búsqueda de funciones, comenzando con la búsqueda de homologías en la base de datos del NCBI y posteriormente en la base de datos del Gene Ontology, teniendo una lista de todos los genes con lo que fue posible identificar los genes de interés, por último se identificaron microsatélites (SSR) en todos los unigenes por lo que se pudo identificar que genes de interés tenían marcadores EST-SSR (Figura 7).

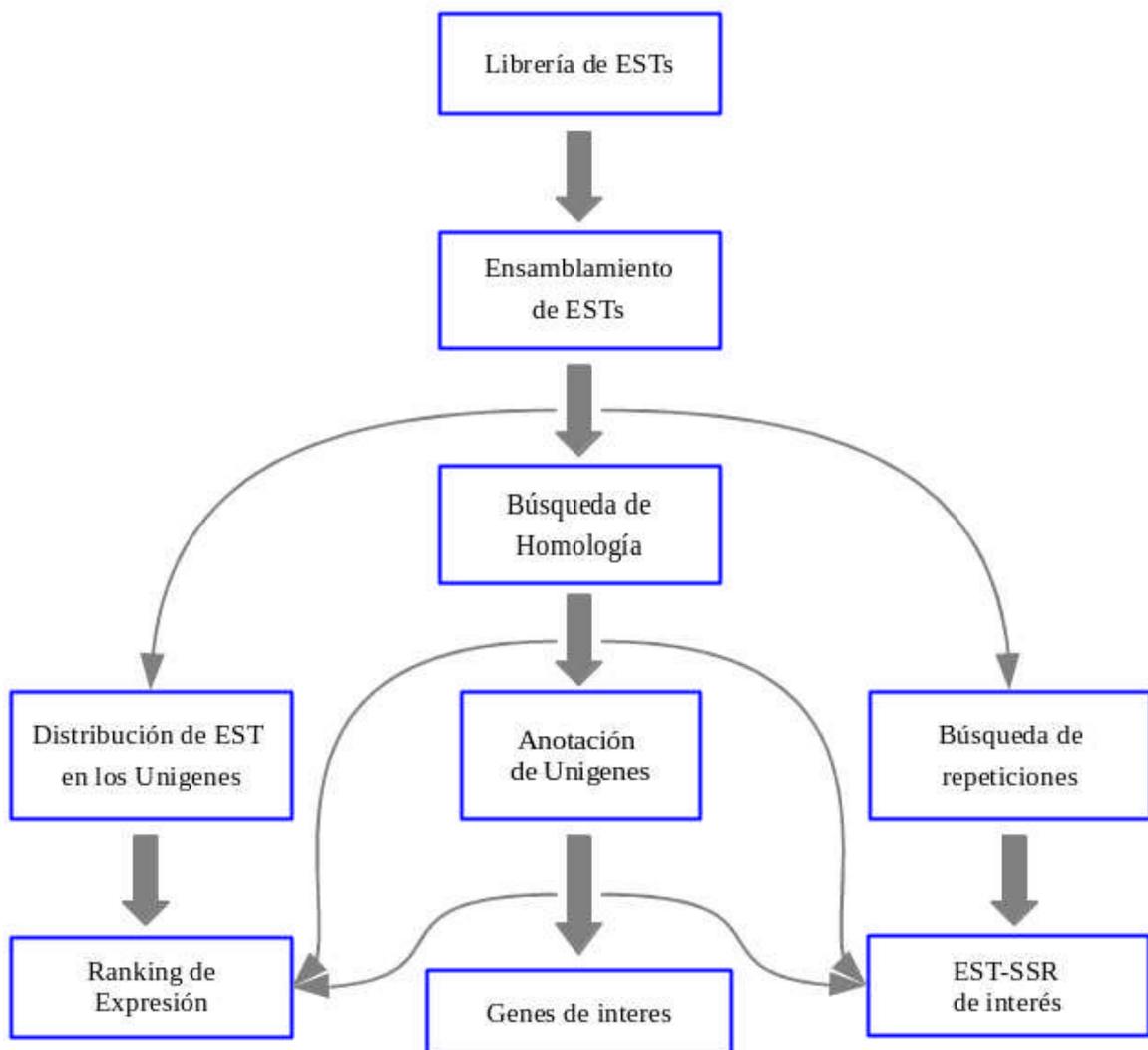


Figura 7. Diagrama de flujo. Mostrando al final los genes más expresados, los genes de interés y los marcadores EST-SSR.

5.2. Ensamblaje de los ESTs.

Con el objetivo de reducir la redundancia en las librerías de ESTs y obtener la mayor longitud de los mRNA, los ESTs fueron ensamblados con el programa CAP3 (Huang y Madan, 1999). El ensamblaje fue hecho tomando en cuenta que el alineamiento tenga como mínimo 50 bp como longitud de solapamiento y un 80% mínimo de similitud del solapamiento.

El resultado de los 7286 ESTs generó contigs (ESTs que han solapado en una región) y singletons (ESTs que no solapan con ningún otro EST). Estas secuencias, contigs más los singletons, forman en conjunto los unigenes. Este conjunto de unigenes solo se trabajó con las secuencias que tenían una longitud mayor o igual a 100 bp ya que la longitud de las secuencias es importante para hallar una buena homología en las bases de datos. Para este objetivo se diseñó un pequeño script por el intérprete de comandos de línea Bourne Again SHell (BASH) que a su vez utiliza otros 2 scripts escritos en PERL "get_fasta_stats.pl" y "extract_fasta.pl" (<http://www.genome.ou.edu/informatics.html>); el primero ve las estadísticas de cada secuencia, entre ellas la longitud y el segundo permite extraer las secuencias deseadas.

5.3. Distribución de los ESTs en los Unigenes.

Obtenidos los Unigenes mayores o iguales a 100 pb se procedió a tener un perfil de la expresión en el tejido para conocer cuántos mRNA están expresándose mayoritariamente. Esto se puede saber debido al número de EST ensamblados en cada Unigen, así cada Unigen nos da una idea de la expresión de cada mRNA. Para esto se hizo un cuadro representativo de la distribución de los EST en los Unigenes, se diseñó un script en BASH el cual toma el archivo del resultado del CAP3 en el ensamblamiento para tener una lista de las distribuciones de los ESTs en los Contigs.

5.4. Homología en la base de datos.

Se descargó toda la Base de Datos de Nucleótidos y Proteínas del NCBI. Se buscó la homología por medio de los programas BLASTN y BLASTX de manera local, con un E-value máximo de $1e^{-40}$ y $1e^{-10}$ respectivamente, y entre los hits reconocidos que se seleccionen como máximo a los 40 primeros hits.

El tamaño de las bases de datos de nucleótidos y proteínas del NCBI no es igual, por lo que en las 2 hay información que no es compartida. La base de datos de

nucleótidos es mucho más grande que el de las proteínas y muchos de los genes involucrados en la fibra solo se conoce su mRNA, muchas veces incompleto, por lo que los resultados de la comparación entre estas 2 bases de datos son importantes.

5.5. Anotación de Secuencias.

El Proyecto Ontología Génica (The Gene Ontology, GO) provee una clasificación y vocabulario estructurado y controlado de muchos dominios para la biología molecular y celular, los cuales describen la función del gen y las relaciones entre ellos. Es una poderosa herramienta para anotar y analizar las funciones de un gran número de genes y producto, su base de datos (GO Database) provee tres ontologías, función molecular, proceso biológico y componente celular. Estos representan el primer nivel GO, cada una subdividida lo que representaría el segundo nivel y estas siguen subdividiéndose, lo que nos permite representar desde lo más general hasta lo más específico usando términos de la ontología. (Gene Ontology Consortium 2004)

Para su anotación y clasificación se utilizó el programa Blast2GO (Conesa *et al.*, 2005) por los parámetros por defecto. Se utilizó el resultado de la comparación contra la base de datos de proteínas con el BLASTX, se mapeó cada hit en su base de datos y anotó la función más aceptada, también permite generar un diagrama con la distribución de las funciones. Posteriormente el análisis de los gráficos se hizo a un segundo nivel GO.

Además de la anotación por comparación con la base de datos del GO, que es una anotación automática, se realizó una anotación curada manualmente de los unigenes que no se identificaron en esta base de datos. Muchas de estas secuencias no identificadas en la base de datos del GO fueron identificadas en la base de datos de nucleótidos y proteínas. En el paso anterior, la búsqueda de homologías, se generaron dos listas, una de la comparación contra la base de datos de nucleótidos y otra contra la base de datos de proteínas, los unigenes que no fueron reconocidos en la base de datos del GO fueron separados y solo se conservaron los que fueron similares con secuencias de referencia (RefSeq), estas son secuencias curadas depositadas en el NCBI, separadas en dos bases de datos curados, RefSeq de transcriptos y RefSeq de proteínas.

Primero se observó qué unigenes no fueron reconocidos por el la base de datos del GO, luego se separaron y se observó quienes hicieron hit con las SeqRef de proteínas y que tenían una identidad del 70%, posteriormente los unigenes que no

fueron reconocidos en la RefSeq de proteínas, fueron identificados quienes hicieron hit con la RefSeq de transcritos y con un 90% de identidad.

5.6. Búsqueda de microsatélites.

La búsqueda de microsatélites o SSR se realizó mediante el script en perl MISA (Thiel *et al.*, 2003). La búsqueda se hizo dentro de los Unigenes mayores de 100 pb y se utilizó los siguientes parámetros: un mínimo de 7 repeticiones para los dinucleótidos y un mínimo 3 repeticiones para los trinucleótidos, tretranucleótidos, pentanucleótidos y hexanucleótidos según Bakhtiarizadeh *et al.*, 2012 y Chen *et al.*, 2015. Los resultados dieron una lista de las secuencias identificadas con microsatélites luego esta se comparó con los resultados de la lista de distribución de ESTs en los unigenes y los resultados de la búsqueda de funciones.

6. RESULTADOS.

6.1. Ensamblaje de los ESTs.

Inicialmente la librería de ESTs contenía un número de 7286 secuencias, los cuales tenían una longitud entre 30 a 1080 pb en donde el 78% de las secuencias (5677 ESTs) tienen una longitud entre 100 y 300 pb, teniendo un promedio de 271 pb (Figura 8).

Se ensamblaron 5842 ESTs en 1188 contigs y 1444 singletons. Estos contigs y singletons forman juntos los 2632 unigenes (Tabla 5), indicando que la redundancia en la librería (número de ESTs ensamblados sobre el total de ESTs) fue del 80%. Cada uno de los unigenes puede representar diferentes transcritos de un mismo mRNA.

Los unigenes tienen una longitud entre 30 a 1694 pb, donde los contigs poseen en promedio una longitud mucho mayor (~325 bp) que los singletons (~275 bp) (Figura 9). La longitud de las secuencias es importante para hallar una buena homología en las bases de datos, por lo que solo se trabajó con los unigenes que tenían una longitud mayor o igual a 100 pb, un total de 2428 unigenes. El 70% de las secuencias (1831 unigenes) tuvieron una longitud entre 100 a 300 pb.

En los contigs se encontraron 5 secuencias con una longitud mayor a 1000 pb, donde la secuencia de mayor longitud (1694 pb) es también el segundo contigs con más ESTs, por lo que es una de los transcritos más expresados.

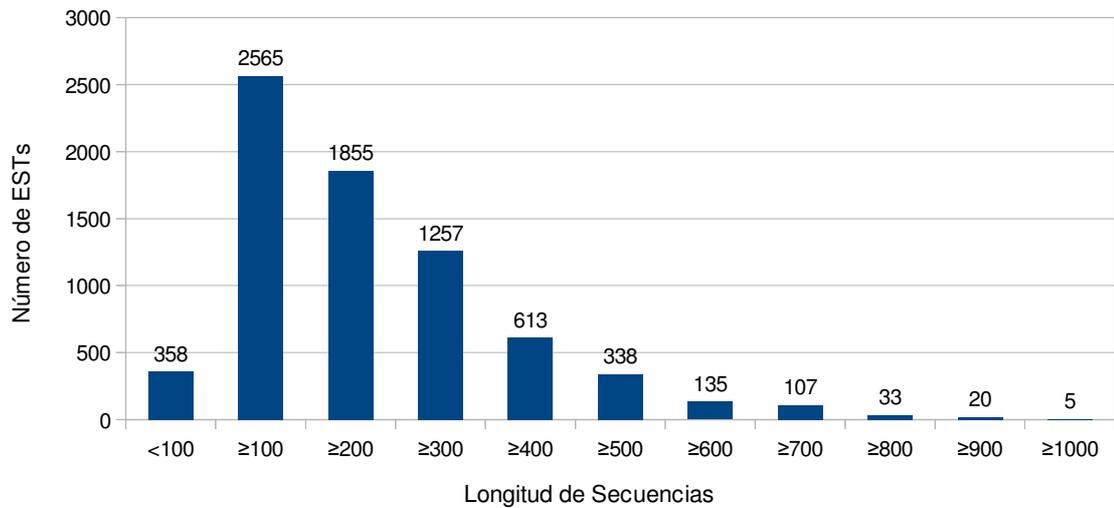


Figura 8. Distribución de las longitudes de las secuencias en la librería de ESTs de alpaca.

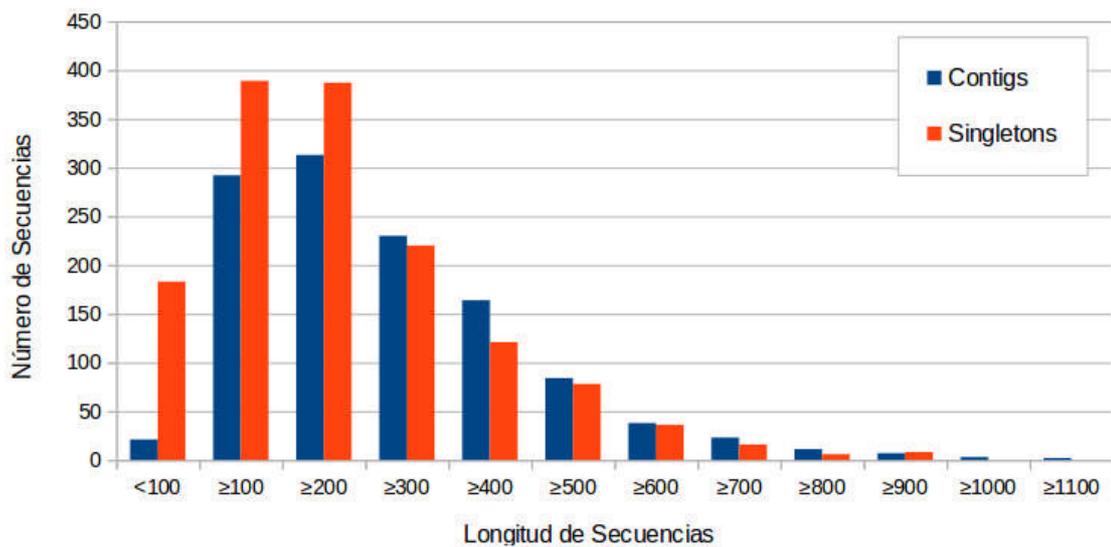


Figura 9. Distribución de las longitudes de las secuencias ensambladas de alpaca. Contigs representados en azul y los singletons en rojo.

Tabla 5. Número de ESTs generados por el ensamblaje de secuencias de alpaca. Se trabajó solo con las secuencias de longitud mayores a 100 pb.

	Número de secuencias totales	Número de secuencias mayores/iguales a 100 pb
Contigs	1188	1167
Singletons	1444	1261
Unigenes	2632	2428

6.2. Distribución de los ESTs en los Unigenes.

La gran mayoría de unigenes están conformadas por pocos ESTs, el 51,9 % está formado por un solo ESTs (singletons), el 22,6 % por 2 ESTs y 5 Unigenes contienen más de 100 ESTs (Figura 10).

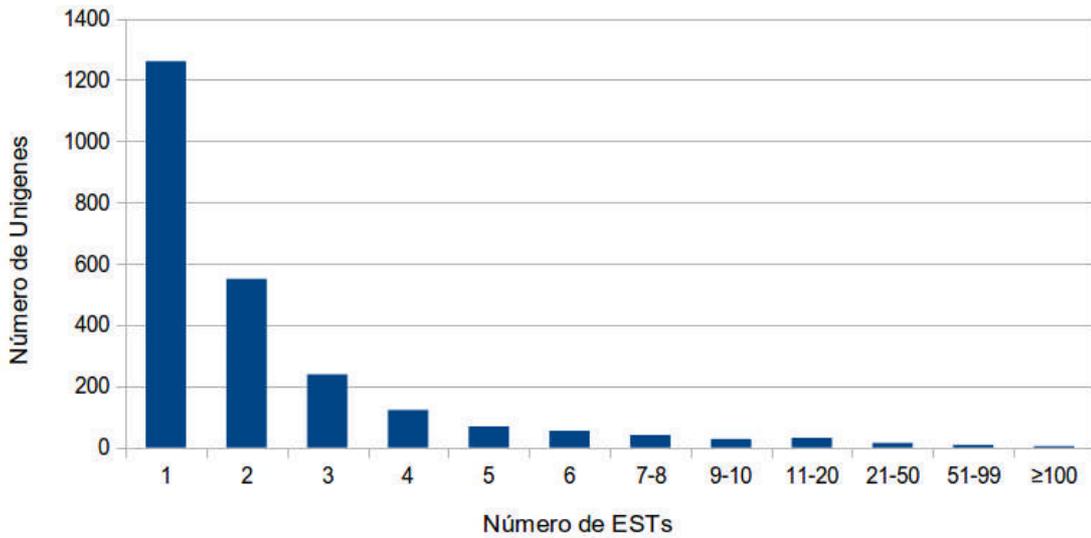


Figura 10. Distribución de los ESTs dentro de los Unigenes.

Una relación entre los 2428 unigenes o posibles tipos de mRNA y los 7286 ESTs o copias de cada mRNA muestra que solo el 10% de los tipos de mRNA representa el 50% del total de copias (Figura 11). Confirma la buena calidad de la librería de ESTs de piel de alpaca debido a que esto es lo que sucede normalmente en la célula, donde la mayoría del total de mRNA en mamíferos está representado por tan solo unos pocos tipos de mRNA (Perdew *et al.*, 2006, Parkinson 2009).

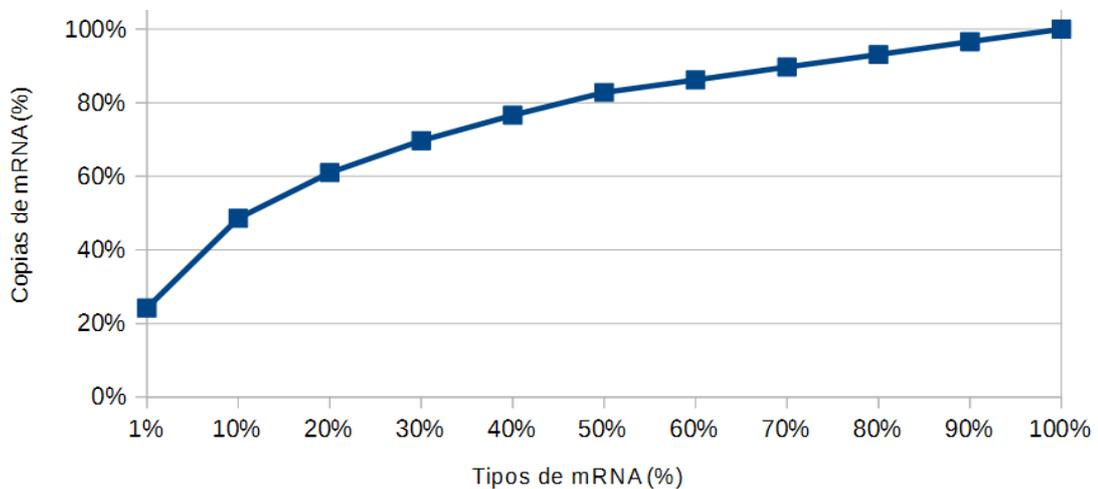


Figura 11. Relación entre los tipos de mRNA o unigenes con la cantidad de copias del mRNA total o ESTs.

6.3. Homología en las bases de datos.

Se realizó la búsqueda en las bases de datos del NCBI de nucleótidos y proteínas, de los 2428 Unigenes, 1666 (~69%) fueron reconocidas. 1660 unigenes fueron reconocidas solo en la base de datos de nucleótidos, con una identidad mínima de 74%, y 637 solo en la de proteínas, con una identidad mínima de 42%, 631 fueron reconocidas tanto por ambas bases (Tabla 6).

La gran mayoría de secuencias, tanto en los contigs como los singletons, fueron reconocidas en las bases de datos, en los contigs el 71% de fue reconocido (830 secuencias) mientras en los singletons lo fue el 66% (836 secuencias) donde más del 50% en ambos casos tienen una longitud entre 200 y 400 bp (Figura 12).

Tabla 6. Número de hits en las bases de datos (BD) de nucleótidos y proteínas del NCBI.

Base de datos	Unigenes
Reconocidas en las BD	1666
Reconocidas solo en la BD de Nucleótidos	1660
Reconocidas solo en la BD de Proteínas	637
Reconocidas en ambas bases de datos	631

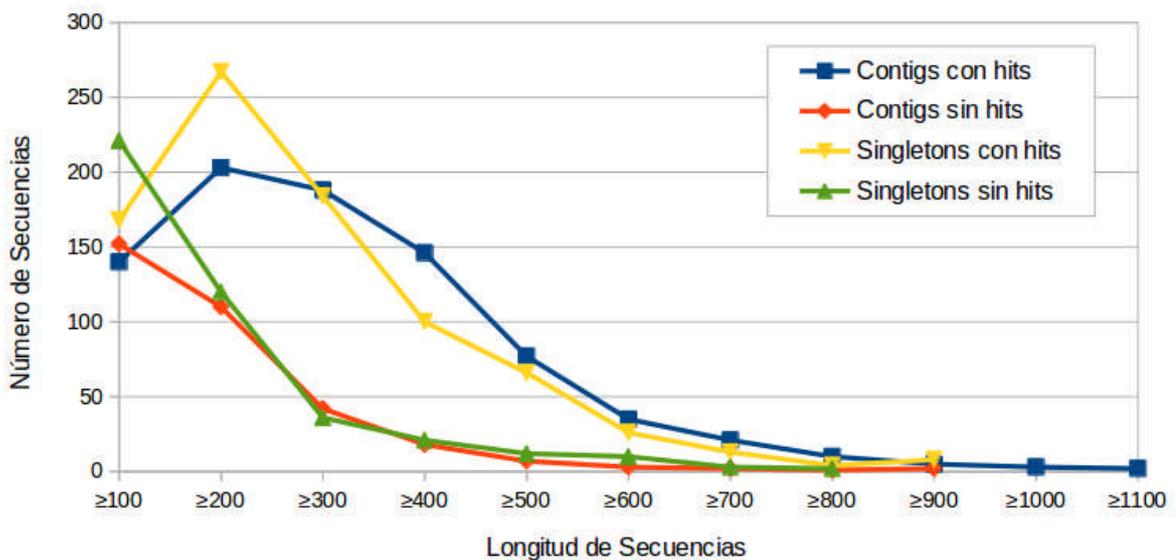


Figura 12. Distribución de las longitudes de los Unigenes con hits y sin hits en las bases de datos de nucleótidos y proteínas.

6.4. Anotación de secuencias.

Los resultados de la búsqueda de homología en la base de datos de proteínas del NCBI se utilizaron para hacer la anotación mediante el programa Blast2GO. De las 637 Unigenes reconocidas por el BLASTX, se obtuvieron 541 unigenes anotados (~85%) por el programa Blast2GO, estas fueron distribuidas en las tres principales categorías GO (Gene Ontology), 1685 secuencias en Componentes Celulares (CC), 2882 secuencias en Procesos Biológicos (PB) y 916 secuencias en Función Molecular (FM) (Figura 13).

Como cada Unigene puede representar un mRNA estos pueden ejercer más de una función por lo que la mayoría de las secuencias están en más de una categoría GO. Así una queratina del tipo II está en dos diferentes procesos biológicos: desarrollo de la epidermis y organización del citoesqueleto, por lo que el número de secuencias en los CC, PB y FM parecería ser mayor que las secuencias anotadas.

En las funciones moleculares las tres primeras categorías con más secuencias fueron las relacionadas a la unión molecular (401 secuencias) entre ellas las proteínas de ligación al citoesqueleto, a receptores, a enzimas, a RNA, a DNA y a regiones nucleicas regulatorias; la actividad catalítica (213) como actividad de hidrolasas, de transferasas y oxidoreductasas y por último en los que tenían una actividad molecular estructural (125 secuencias), entre ellas las constituyentes estructurales de los ribosomas, del citoesqueleto y de la epidermis.

En los procesos biológicos destacaron los procesos celulares (459 secuencias) que se encuentran los relacionados a la comunicación celular, el ciclo celular y la división celular; y los procesos metabólicos (426) donde se encuentran los procesos metabólicos de macromoléculas y de metabolitos primarios.

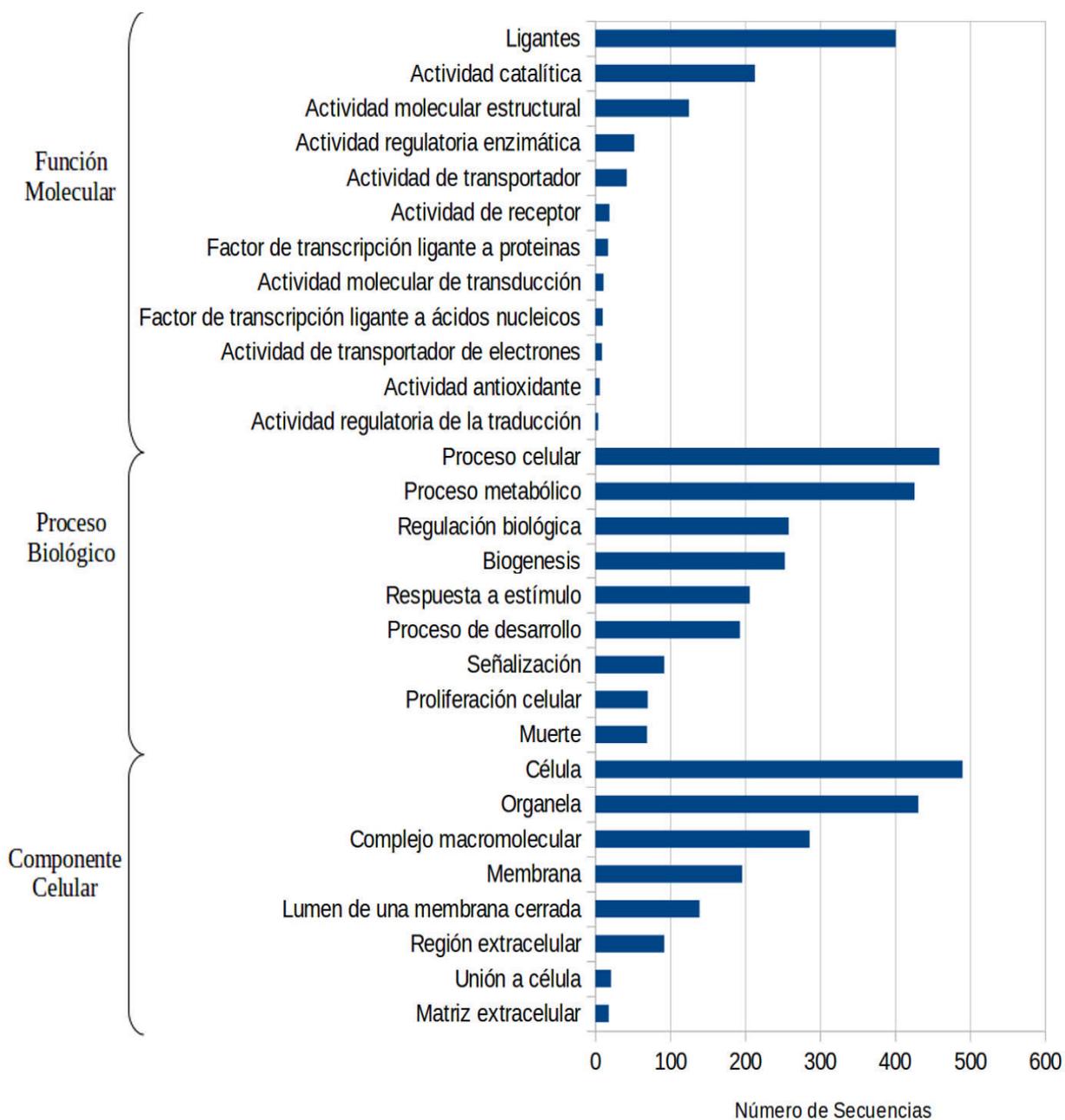


Figura 13. Principales anotaciones GO de los ESTs de piel de alpaca. Los datos están separados por sus tres categorías. Se encontraron 1685 secuencias en Componentes Celulares (CC), 2882 secuencias en Procesos Biológicos (PB) y 916 secuencias en Función Molecular (FM).

De los unigenes que no fueron reconocidos por la base de datos GO (1887 secuencias), 33 fueron reconocidas solo en la base de datos de RefSeq de proteínas y 647 solo por la RefSeq de transcritos. Estos unigenes revisados se añadieron a la lista de anotación obteniendo en total 1221 unigenes anotados y con 1207 unigenes de función desconocida (Tabla 7).

Tabla 7. Resumen de los unigenes anotados.

Unigenes anotados y comparados contra	Número de unigenes
El Gene Ontology	541
RefSeq de proteínas	33
RefSeq de transcritos	647
Total unigenes anotados	1221
Total unigenes con función desconocida	1207

6.5. Búsqueda de microsatélites.

Se identificaron 673 microsatélites repartidos en 501 unigenes, los que representan el 20,6% del total, mostrando que las repeticiones trinucleotidos son los más abundantes (Tabla 8) y que la repetición AGG/CCT es la más frecuente (16,05 %) seguida de AAG/CTT (11,14 %) y AGC/CTG (11,14%).

Tabla 8. Los tipos de repeticiones con sus frecuencias en los unigenes.

Repetición	Porcentaje
Dinucleotidos	11,74 %
Trinucleotidos	64,64 %
Tetranucleotidos	18,28 %
Pentanucleotidos	3,27 %
Hexanucleotidos	2,08 %

6.6. Genes altamente expresados.

De los resultados de la distribución de los ESTs en los unigenes y la identificación de genes, tenemos genes que se encuentran más expresados en la piel de alpaca (Tabla 9). Se observa que entre los más expresados, el factor de replicación 5 (*RFC5*) y la proteína que une Acil-CoAs (*ACBP*) fueron solo identificadas por la RefSeq de transcritos y proteínas, respectivamente.

También se observó que algunos unigenes correspondían al mismo gen, en el caso del transcripto que codifica la *Tweak* hubo 4 unigenes, todos identificados en el GO. Para el transcripto *RFC5* hubo 3 unigenes, todos identificados en la RefSeq de transcritos. Para los genes que codifican *COL1A1*, *COL1A2* y *SPARC*, todos tuvieron 2 unigenes por cada uno, uno identificado por el Gene Ontology y otro por la RefSeq de transcritos, por lo que en la columna de “N° ESTs” en la Tabla 9 aparece la suma de todas las secuencias de ESTs que se identificaron con el mismo gen.

Tabla 9. Genes altamente expresados en la piel de alpaca. BD, columna de las bases de datos donde se identificó la función, GO (Gen Ontology), Trans (RefSeq de transcritos), Prot (RefSeq de proteínas), N° ESTs (número de ESTs)

Gen	Nombre del gen	N° ESTs	BD
<i>TNFSF12</i>	Tweak, factor de necrosis tumoral, ligando, miembro 12.	235	GO
<i>RFC5</i>	Factor de replicación C, subunidad 5	144	Trans.
<i>COL1A2</i>	Gene del colágeno $\alpha 1$	143	GO/Trans
<i>RNA pol II</i>	ARN polimerasa II	136	GO
	Desconocido	118	
<i>SPARC</i>	Proteína secretada ácida y rica en cisteína	95	GO/Trans
<i>COL1A1</i>	Gene del colágeno $\alpha 1$	89	GO/Trans
	Desconocido	88	
<i>sFRP4</i> ¹	Proteína secretada relacionada con Frizzled	79	GO
	Desconocido	79	
	Desconocido	72	
	Desconocido	64	
	Desconocido	62	

6.7. Genes de interés en la síntesis de fibra.

Los resultados obtenidos generaron una lista de 22 genes de mayor interés relacionados a la síntesis de la fibra (Tabla 10), muchos mostraron tener microsatélites.

Tabla 10. Genes identificados involucrados en la síntesis de la fibra.

¹ Los genes identificados en la base de datos de nucleótidos.

Gen	Nombre del gen	SSR
<i>KRT10</i>	Queratina 10	
<i>KRT25</i>	Queratina 25	
<i>KRT31</i> ¹	Queratina 11	
<i>KRT14</i> ¹	Queratina 14	
<i>KRT15</i> ¹	Queratina 15	(GTG) ₃
<i>KRT5</i> ¹	Queratina 5	(GTG) ₃ , (CCT) ₃
<i>KRT78</i> ¹	Queratina 78	
<i>KRT81</i>	Queratina 81	(GAA) ₃
<i>KRT85</i> ¹	Queratina 85	(TGC) ₃
<i>KRTAP3-2</i> ¹	Proteína asociada a queratina 3-2	(TGC) ₃
<i>KRTAP10</i>	Proteínas asociadas a queratinas 10	
<i>KRTAP11-1</i> ¹	Proteína asociada a queratina 11-1	
<i>TNFSF12</i>	Tweak, factor de necrosis tumoral, ligando, miembro 12.	(AATA) ₃
<i>sFRP4</i>	Proteína secretada relacionada con Frizzled	
<i>RACK1</i>	Receptor de la actividad de la proteína quinasa C1	
<i>COL1A1</i>	Gene del colágeno α1	(TGC) ₃
<i>COL1A2</i>	Gene del colágeno α2	
<i>DAB2IP</i>	disabled homolog 2-interacting protein Proteína activadora de GTPasa, subfamilia Ras.	(AAG) ₃
<i>GLI1</i>	Proteína de dedo de cinc	(CCT) ₃
<i>BMP1</i>	Proteína morfogénica osea 1	
<i>BMP4</i>	Proteína morfogénica osea 4	(TGG) ₃ , (CA) ₁₀
<i>CTNNB1</i>	Beta catenina 1	

7. DISCUSIÓN

7.1. Identificación de la cantidad de tipos y copias de mRNA.

La identificación de transcritos de especies no modelos, como es el caso de la alpaca, a través de la generación de librerías de ESTs y su secuenciación han sido ampliamente utilizados para explorar la diversidad de los transcriptomas de los eucariontes. Típicamente han sido usados para identificar genes nuevos y/o de interés, secuencias conservadas evolutivamente, niveles relativos de expresión génica, análisis de transcritos alternativos y filogenómica de organismos en los cuales el genoma completo no es económicamente disponible por lo que la secuenciación de ESTs es mucho más factible (Parkinson 2009, Buzdin y Lukyanov, 2007).

Debido a que los ESTs son el resultado de un solo paso de secuenciación, ellos tienen una tasa de error cerca del 3% (Boguski *et al.*, 1993), pueden contener sustituciones, deleciones o inserciones comparado con el mRNA de donde provienen, además que representan solo un fragmento del mRNA. Por eso es importante el ensamblaje, para obtener una buena calidad de secuencias.

La mayoría de los unigenes resultaron que presentan pocos ESTs, el 51% con uno y el 22% con dos (Figura 10), en los resultados se observó que de los 2428 tipos de mRNA identificados, unigenes, que representan 7286 copias en total de mRNA, cantidad de ESTs, solo el 10% de ellos (264 tipos de mRNA) representan el 50% de total de copias, así el 50% de los tipos representan más del 80% del total de copias (Figura 11). Esto se debe a que la mayoría del total de mRNA en mamíferos está representado por tan solo unos pocos tipos de mRNA (Perdew *et al.*, 2006) confirmando una buena calidad de la librería de ESTs de piel de alpaca puesto que es lo que normalmente sucede en la célula (Parkinson 2009).

Que la mayoría de tipos de mRNA haya tenido solo una expresión de unas cuantas copias también involucra su longitud (Figura 12). La mayoría presentan una longitud corta, el 50% de ellos entre 200 a 400 pb, debido a que fueron ensamblados por pocas ESTs. Esto es común en los trabajo de análisis de ESTs y en general con trabajo de análisis de transcriptomas (Parkinson, 2009).

Existen secuencias que solo fueron identificadas por la RefSeq de transcritos y proteínas, como la *RFC5* y la *ACBP*. Así también unigenes que correspondían al mismo gen, como la *Tweak* que en un inicio el unigene identificado para este gen presentaba 162 ESTs, el más expresado, sin embargo habían 3 unigenes más

identificados para el mismo gen por lo que sumados los 4 unigenes sumaban 235 ESTs o replicas para este gen. El mismo caso se encontró para el gen *RFC5* que se identificaba en 3 unigenes y *COL1A1*, *COL1A2* y *SPARC* todos identificados en 2 unigenes, en todos ellos el número de ESTs originales aumentó gracias a la comparación de varias bases de datos. El hecho de que el mismo gen haya sido identificado por varios unigenes es porque estos unigenes son similares en distintas regiones del mRNA y no se sobrelaparon por lo que no lograron constituir un mismo unigen en el paso de ensamblaje, esto suele pasar en los análisis de ESTs (Parkinson, 2009).

Debido a un análisis de procesos automáticos y su posterior análisis manual (análisis curado), muchos genes aumentaron su expresión demostrando que es necesario compararlo no sólo con la base de datos del GO si no también con otras bases de datos curadas.

7.2. Identificación de las funciones de los mRNA.

A nivel molecular solo se tiene información en humanos, algo en ovejas, y prácticamente nada en alpacas. En el estudio que hicieron Fan *et al.* (2011) obtuvieron 3504 unigenes de los cuales encontraron 863 genes. En el presente estudio se encontró un menor número de unigenes (2428), sin embargo se identificó un mayor número de funciones, 1221 genes, mientras que 1207 no se encontraron función alguna, esto debido que la comparación no solo se realizó con la base de datos del NCBI de proteínas y del GO como realizaron Fan *et al.* (2011) sino a la también a la base de datos del NCBI de nucleótidos. Estas secuencias sin identificar pueden tratarse de genes desconocidos, por lo que no se reconocen en las bases de datos, sin embargo también existe la posibilidad de que estas no se hayan reconocido por que existen errores debido a su secuenciación.

Es de notar que casi en su totalidad los unigenes identificados en la base de datos de proteínas fueron también identificados en la base de datos de nucleótidos (Tabla 6) más lo contrario no ocurrió. Esto demuestra la necesidad de la búsqueda de homologías en ambas bases de datos para un mejor análisis, pues la mayoría de las secuencias identificadas en la fibra del pelo en diversos animales modelos, como humanos, ovejas y cabras esta solo disponible en secuencias de nucleótidos. Es por eso que si bien la ontología génica nos permite ordenar la gran cantidad de datos y obtener de una manera rápida la función de cada secuencia, esta se basa en la comparación con la base de datos de proteínas. Por lo tanto las secuencias que no

tienen un homólogo en esta base de datos no estarán presentes en el análisis de ontologías, pero no significa que las secuencias que no encontraron un homólogo en la base de datos de proteínas no lo hagan en la base de datos de nucleótidos, puesto que esta última es mucho más grande. Pero aún mucho más importante que su tamaño es la información relacionada a este estudio, ya que se observó que la gran mayoría de investigaciones que sirvieron de referencia para este estudio fueron hechas a nivel de mRNA. Por eso (Tabla 7) se presentan datos del análisis de la ontología como la comparación con la base de datos de proteínas, que no llegaron a ser encontrados en la base de datos GO, y nucleótidos. Debido a esta metodología se identificaron transcritos de interés que también identificaron Fan *et al.* (2011) (*COL1A1* y *COL1A2*, *RNA pol II*, *SPARC*, *sFRP4*, *GLI*, *BMP1*, *BMP4*, β -*catenina*) y que no habían sido identificados (*Tweak*, *RFC5*, *ACBP*, *RACK1*, *DAB2I*).

7.3. Identificación de genes asociados al desarrollo y la calidad de la fibra.

La complejidad de los factores que se ven asociados al desarrollo de la fibra hace difícil determinar qué genes son determinantes para la calidad de la misma, incluso algunos genes identificados y asociados a la calidad de fibra tienen un efecto pleiotropico, esto dificulta la selección de genes primordiales para la calidad. Un estudio de mutaciones en genes "halo-hair", causante de una pilosidad extrema en ovejas Romney, ideal para un sector de la industria en alfombras, muestra que esta conlleva a la formación de cuernos en las hembras, un número de efectos en la morfología del folículo piloso y un profundo efecto en las características de la fibra (Purvis y Franklin, 2005), así que estos genes son interesantes en el contexto de entender la función de los genes en el desarrollo del folículo y la fibra. También se han reportado estudios en mutaciones de genes de ovejas que causan una lustrosidad en la fibra como la que se aprecia en suri, una característica apreciable pues se prefiere la lustrosidad al rizado de la fibra, sin embargo esta presenta una reducción significativa de la densidad del folículo y el peso del vellón en ovejas (Blair, 1990) mas no se puede extrapolar al modelo de alpaca, pues estas características no aparecen en suri.

7.4. Identificación de las Queratinas y KAPs.

Actualmente se ha identificado un gran número de los componentes moleculares en el pelo y en el folículo piloso en humanos y otros mamíferos modelos. Entre las proteínas que cobra mayor interés son los filamentos intermedios de

queratinas que están asociados al pelo (hKIFs) y las proteínas asociadas a queratinas (KAPs) por ser las más expresadas (Rogers, 2004).

A diferencia del estudio de Fan *et al.* (2011) donde encontraron 3 queratinas (*KRT10*, *KRT14* y *KRT15*), en el presente estudio se encontraron 9 queratinas, 5 transcritos para las queratinas tipo I (*KRT10*, *KRT25*, *KRT31*, *KRT14* y *KRT15*) y 4 del tipo II (*KRT81*, *KRT85*, *KRT5*, *KRT78*) además de encontrar 3 transcritos para las KAPs (*KRTAP3-2*, *KRTAP10*, *KRTAP11-1*). Es de saber que en muchos miembros de las KPAs, incluida la familia KAP10, se han identificado varios elementos potenciadores (enhancers), tales como ETS, proteínas de choque térmico, AML y miembros de la familia HOX (Rogers *et al.*, 2004). Estos no se mostraron muy expresados pues existen varios factores importantes involucrado en ello, por ejemplo la región de la toma de la muestra, se sabe que la densidad folicular en la alpaca disminuye en la dirección antero-posterior y dorso-ventral, donde la región con mayor densidad es el cuello (más de 20 folículos/mm²) y se sabe que existe una relación negativa entre la densidad y el diámetro, a una mayor densidad un menor diámetro (Arana, 1972, Bustinza, 2001); aparte otros factores como la generación de la librería, el secuenciamiento, así como la metodología usada para el análisis de datos (Parkinson, 2009).

En la identificación de estas se probó que la metodología usada para la identificación de funciones no es suficiente solo compararlo con la base de datos Gene Ontology, a pesar de que esta base de datos permite la ontología génica y la distribución de funciones ordenadamente, algunas de las funciones identificadas se hicieron con los resultados de la comparación con la base de datos de nucleótidos y proteínas del NCBI (Tabla 9) que no llegaron a ser reconocidas por el Gene Ontology.

Si bien la identificación no mostró que está entre los más expresados, la expresión de muchos de sus miembros está en la cutícula de la fibra y por lo que están involucrados a su estructura son los más importantes a identificar. Estudios prueban que los genes *KRTAP6* y *KRTAP8* poseen alelos que ayudan a tener un menor diámetro (Parsons *et al.*, 1994b), por lo que la identificación de KAPs para probarlos en futuros estudios es de mucha importancia.

7.5. Identificación de factores de necrosis tumoral (TNF).

En el presente trabajo se identificó el gen para la proteína Tweak, es el miembro 12 de la superfamilia de los factores de necrosis tumoral (ligandos)

(*TNFSF12*). Los factores de necrosis tumoral (TNF) constituyen una superfamilia de 19 ligandos y 29 precursores, estos son importantes para numerosos procesos incluyendo la apoptosis. La señal TNF, regula la formación y la función de la placoda, e interacciona con WNT en la comunicación mesénquima-epitelial durante la primera señal (Laurikkala *et al.*, 2002). La Tweak, una proteína transmembrana, tienen varias funciones entre ellas la estimulación del crecimiento celular y un inductor débil de la apoptosis, esta se expresa solo en la capa basal de la epidermis y se le asocia al efecto directo pro-apoptotico en los queratinocitos humanos (Sabour Alaoui *et al.*, 2012) por lo que regula la formación del folículo piloso.

7.6. Identificación de factores asociados a la vía de señalización Wnt/ β -catenina.

Los ligantes Wnt (Wnt1, Wnt3a, Wnt8a y Wnt8b) son críticos para la señalización durante el desarrollo, muchos factores modulan esta actividad, uno de los primeros factores que se identificaron fueron las proteínas secretadas relacionadas a Frizzled (sFRP, secreted Frizzled-related proteins) las cuales contienen un dominio rico en cisteína (CRD, cysteine-rich domain) el cual posee un gran parecido en identidad y estructura con el CRD presente en las proteínas receptoras de Wnt, las proteínas Frizzled. Esto sugiere que se puede ligar a Wnt y pueda inhibir la señal por lo que actúa como un modulador endógeno para la vía WNT.

Poco es conocido sobre la especificidad de los miembros de la familia sFRP para varios Wnt, actualmente se conocen 5 variantes, además de que en muchos órganos no están todos presentes y algunos juegan un rol más importante que otros, dependiendo de qué órgano se trate, por ejemplo en pollos en el desarrollo del tubo neural, sFRP1 y sFRP2 pueden inhibir la actividad de Wnt3a mientras sFRP3 no lo hace (Galli *et al.*, 2006), en la glándula mamaria en ratas, la sFRP4 inhibe también a Wnt3a (Constantinou *et al.*, 2008), en el desarrollo del riñón en ratas sFRP1 y sFRP2 compiten localmente para regular la señal Wnt durante la organogénesis renal, así mientras sFRP1 se liga a Wnt4 e inhibe su función, sFRP2 no lo inhibe pero restablece el desarrollo de la inhibición por sFRP1 (Yoshino *et al.*, 2001), en células madres hematopoyéticas de ratón sFRP1 y sFRP2 actúan como inhibidores de Wnt3a (Nakajima *et al.*, 2009).

No se conoce el rol que juega la familia sFRP en el desarrollo del folículo piloso, además de que su expresión varía según el tejido y en muchos estudios son enfocados a enfermedades donde su expresión también varía. En el presente estudio se identificó una secuencia homóloga al gen *sFRP4* el cual posee una alta identidad,

99% contra la base de datos de nucleótidos y 100% con la bases de datos de proteínas, además de estar entre los contigs con mayor número de secuencias ensambladas, 79 ESTs, por lo su expresión puede ser fundamental para la organogénesis del folículo piloso. En un estudio reportado sobre la piel de humanos donde el sFRP4 promueve la diferenciación celular de la epidermis, este es expresado en condiciones normales y su expresión se incrementa durante la diferenciación de los queratinocitos y en la apoptosis (Maganga *et al.*, 2008).

La célula constantemente sintetiza y degrada la β -catenina, excepto cuando se activa la vía canónica Wnt, un complejo de proteínas que se encarga de la degradación de la β -catenina por lo que es importante para mantener los niveles bajos de la β -catenina en ausencia de la señal exógena Wnt, mediante la fosforilización y la ubiquitinación de la β -catenina. Este complejo de destrucción lo conforman la APC (adenomatous polyposis coli), Axin, CK1 α , la GSK3 β y la RACK1 (Kimelman y Xu, 2006; Deng *et al.*, 2012). El mecanismo de función es aún poco claro. Se identificó la presencia de una secuencia homóloga al gen *RACK1*, la baja-expresión de *RACK1* promueve el crecimiento celular mientras que su sobre-expresión lo atenúa, por lo que se piensa que posee una función de regulación negativa de la vía Wnt/ β -catenina mediante la estabilización del complejo de destrucción de la β -catenina (Deng *et al.*, 2012).

Así también se identificó la presencia de *DAB2IP*, gene que codifica la proteína activadora de GTPasa de la subfamilia Ras. Se ha observado que si se silencia a este gen, se observa la acumulación de la β -catenina en el citoplasma y su translocación al núcleo, este genera el complejo DAB2IP-PP2A-GSK-3 β el cual facilita la activación de GSK3 β a través de la desfosforilización (Xie *et al.*, 2010).

7.7. Identificación de colágenos.

Es conocido que los colágenos son muy abundantes en la piel, estos comprenden aproximadamente el 70% de su peso seco en humanos, entre ellos los predominantes son los colágenos tipo I y III donde el tipo I es el más abundante el cual comprende entre el 80% al 90% del total del colágeno (Ala-Kokko *et al.*, 1987). El colágeno de tipo I es el producto de dos genes α 1(I) [*col1a1*, gen del colágeno α 1(I)] y α 2(I) [*col1a2*, gen del colágeno α 2(I)] estos contribuyen a la formación de fibrillas (Trojanowska *et al.*, 1998). Se encontró que la expresión de *COL1A1* es abundante en la piel de alpaca, lo cual es también consistente con estudios anteriores en piel de humanos (Cheng *et al.*, 2011). La *WNT3A* se ve involucrada en el incremento de expresión de algunos genes incluidos el *COL1A1* (Zhang *et al.*, 2013).

7.8. Identificación de señales proteínas morfogénicas oseas.

Durante la formación del folículo piloso embrionario, la interacción epitelial-mesenquimal promueve la especificación y agrupación de células dermales especializadas, llamadas células de la papila dérmica, las cuales transmiten señales a las células madres epiteliales multipotentes para iniciar la morfogénesis (Millar, 2002), las proteínas BMP tienen un fuerte efecto positivo en la papila dérmica por lo que regula tanto la proliferación como la diferenciación del folículo (Rendl *et al.*, 2008) y en este aspecto la expresión de *BMP4* tienen el rol más importante, su expresión epitelial en el bulbo está restringida a la matriz distal, incluso se cree que este podría, junto con *BMP2* y *NOGGIN*, crear tres zonas de actividad BMP en la papila dérmica (Kulesa *et al.*, 2000). Su expresión en piel de ovejas merino está más restringido a la zona epitelial circundante al folículo en la anagen (Botchkarev y Sharov, 2004), por lo que la *Bmp4* no está entre las más expresadas tal y como se observó en los resultados.

7.9. Identificación de microsatélites.

La distribución de los microsatélites en el genoma no es aleatoria, probablemente por el efecto en la regulación genética, recombinación, replicación del DNA, etc, haciendo que estos se encuentren bajo una fuerte presión selectiva a diferencia de otras regiones debido a su importancia que cobre en el fenotipo (Li *et al.*, 2004).

Como muestra la Tabla 8, los microsatélites trinucleotidos fueron los más abundantes. Este tipo de repeticiones es la más común en las regiones codificantes (Li *et al.*, 2004). Se encontraron microsatélites en el 20,6% de todos los unigenes, este porcentaje es muy variado entre los mamíferos, pues se han encontrado en el 4% de unigenes en vacas (Yan *et al.*, 2008) y 3,8% en cabras (Xu *et al.*, 2013). Este alto porcentaje en alpacas podría deberse a que solo se cuenta con un número determinado de ESTs de un solo tejido mientras que en otros mamíferos se cuenta con muchos más ESTs y de una variedad de tejidos.

Se encontraron microsatélites en 10 genes de interés: *COL1A1*, *TNFSF12* (TWEAK), *DAB2IP*, *GLI1*, *BMP4*, *KRT81*, *KRT5*, *KRT15*, *KRT85* y *KRTAP3-2*. No hay reportes que señalen microsatélites en estos genes por lo que encontrarlos tendría una aplicación muy útil para estudios posteriores especialmente en la determinación de su relación con el diámetro de la fibra. Por ejemplo se tiene como referencia que

muchos miembros de las KAPs poseen polimorfismos (Rogers, 2004), pero solo en dos de ellos se han encontrado alelos asociados a un menor diámetro, *KRTAP6* y *KRTAP8* (Parsons *et al.*, 1994b) además que estos están ligados 15 centimorgans (Parsons *et al.*, 1994a) los alelos de este último corresponden a un microsatélite con motivo dinucleotido CA (Wood *et al.*, 1992).

8. CONCLUSIONES.

- Se asignaron 1221 funciones génicas de los cuales se identificaron 22 genes involucrados en la síntesis de la fibra de la alpaca.
- Se identificaron 358 nuevas funciones en la librería de ETS de alpaca depositados en el NCBI.
- Se identificaron 9 nuevos genes en los mayores componentes moleculares de la fibra: 6 queratinas (*KRT5*, *KRT25*, *KRT31*, *KRT78*, *KRT81* y *KRT85*) y 3 KAPs (*KRTAP3-2*, *KRTAP10* y *KRTAP11-1*).
- Se encontraron que 4 genes involucrados a la síntesis de la fibra poseen una alta expresión a nivel del mRNA (*TNFSF12*, *sFRP4*, *COL1A1* y *COL1A2*).
- Se encontraron 501 secuencias con marcadores microsatélites de los cuales 10 genes están relacionados a la síntesis de fibra (*COL1A1*, *TNFSF12*, *DAB2IP*, *GLI1*, *BMP4*, *KRT81*, *KRT5*, *KRT15*, *KRT85* y *KRTAP3-2*).

9. RECOMENDACIONES.

Hacer un cluster de las funciones previamente al paso de ensamblamiento, ya que en algunos casos algunos unigenes tienen la misma función, por lo que corresponden al mismo ARNm pero regiones no tan cercanas como para que se ensamblen en un solo unigen. En el estudio se hizo manualmente, pero al ver la lista de genes identificados, se recomienda desde el inicio usar un programa que haga un cluster de funciones para no hacerlo manualmente.

La búsqueda de homologías puede hacerse con una base de datos curadas como el UniProt o RefSeq, esto hará que haya un menor número de genes identificados pero más fiables en cuanto a la función predicha. En este estudio los parámetros usados en los programas fueron estrictos, aun así se recomendaría usar base de datos curadas.

El genoma de la alpaca ya está disponible 22X, mapear los genes en el genoma para así obtener toda la secuencia del ARNm y tener sus regiones UTR e intrones, ya que en estos podrían encontrarse motivos que regulen la expresión de los genes identificados.

10. BIBLIOGRAFIA

ADELSON, DL; CAM, GR; SESILVA, U; and Franklin, IR. Gene expression in sheep skin and wool (hair). *Genomics* 2004, vol. 83, n° 1, p. 95-105.

ANDL, T; REDDY, ST; GADDAPARA, T and MILLAR, SE. WNT signals are required for the initiation of hair follicle development. *Developmental Cell* 2002, vol 2, n° 5, p. 643-653.

ALA-KOKKO, L; RINTALA, A and SAVOLAINEN, ER. Collagen gene expression in keloids: analysis of collagen metabolism and type I, III, IV, and V procollagen mRNAs in keloid tissue and keloid fibroblast cultures. *The Journal of Investigative Dermatology* 1987, vol 89, n° 3, p. 238-44.

ARANA, BL. "Distribución de la densidad folicular en la piel de la alpaca y su relación con el diámetro de la fibra". Tesis Título Profesional de Ingeniero Zootecnico, UNALM, Lima 1972.

BADAJOS, AL. "Determinación de la finura de fibra de alpaca asociado a la relación folículo secundario/folículo primario (S/P) entre las razas Suri y Huacaya". Tesis para Título de Médico Veterinario, UNMSM, Lima 2007.

BAKHTIARIZADEH, MR; AREFNEJAD, B; EBRAHIMIE, E and EBRAHIMI, M. Application of functional genomic information to develop efficient EST-SSRs for the chicken (*Gallus gallus*). *Genetics and Molecular Research*, 2012, vol. 11, n° 2, p 1558-1574

BOGUSKI, MS; LOWE, TM and TOLSTOSHEV, CM. dbEST database for "expressed sequence tags". *Nature Genetics* 1993, vol. 4, n° 4, p. 332-333.

BOTCHKAREV, VA and SHAROV, AA. BMP signaling in the control of skin development and hair follicle growth. *Differentiation* 2004 vol. 72, n° 5, p. 512-526.

BLAIR, HT. Inheritance of a major gene for excessively lustrous wool in sheep. *Journal of Hereditary* 1990, vol. 81, n° 3, p. 220-222.

BRAY M, REVELL DK, BAWDEN CS and HYND PI. Keratin gene expression in Merino sheep with divergent wool growth. *Australian Journal Agricultural Research* 2005, vol. 56, n° 3, p. 203-210.

- BRENES, ER; KRYSSIA, M; PÉREZ, F and VALLADARES, K. El Cluster de los Camélidos en Perú: Diagnóstico Competitivo y Recomendaciones Estratégicas. Proyecto Andino de Competitividad, Documentos de Trabajo, Instituto Centroamericano de Administración de Empresas, INCAE 2001, p 14-15.
- BUSTINZA, AV. La Alpaca, Conocimiento del Gran Potencial Andino. Edit. Universidad Nacional del Altiplano, Puno, Perú, 2001.
- BUZDIN, A and LUKYANOV, S. Nucleic Acids Hybridization Modern Applications. The Netherlands: Springer 2007, 314 p. ISBN 978-1-4020-6039-7.
- CADIGAN, KM and NUSE, R. Wnt signaling: a common theme in animal development. *Genes & Development* 1997, vol. 11, n° 24, p. 3286-3305.
- CADIGAN, KM and LIU, YI. Wnt signaling: complexity at the surface. *Journal of Cell Science* 2006, vol. 119, n° 3, p. 395-402.
- CECCHI, T; VALBONESI, A; PASSAMONTI, P; FRANK, E and RENIERI, C. Quantitative variation of melanins in llama (*Lama glama L.*). *Small Ruminant Research* 2007, vol. 71, n° 1, p. 52-58.
- CHEN, W; LIU, Y-X and JIANG, G-F. *De novo* Assembly and Characterization of the Testis Transcriptome and Development of EST-SSR Markers in the Cockroach *Periplaneta americana*. *Scientific Reports* 2015, vol. 5 n° 1, p. 11144.
- CHENG, W; YAN-HUA, R; FANG-GANG, N and GUO-AN, Z. The content and ratio of type I and III collagen in skin differ with age and injury. *African Journal Biotechnology*, 2011, vol. 10, n° 13, p. 2524-29.
- CONESA, A; GÖTZ, S; GARCÍA-GÓMEZ, JM; TEROL, J; TALÓN, M and ROBLES, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005, vol. 21, n° 18, p. 3674-3676.
- CONSTANTINO, T; BAUMANN, F; LACHER, MD; SAURER, S; FRIIS, R and DHARMARAJAN, A. SFRP-4 abrogates Wnt-3a-induced beta-catenin and Akt/PKB signalling and reverses a Wnt-3a-imposed inhibition of in vitro mammary differentiation. *Journal of Molecular Signaling* 2008, vol. 3, n° 10, p. 1-14.
- DE LOS RÍOS E. Producción Textil de Fibras de Camélidos Sudamericanos en el Área Altoandina de Bolivia, Ecuador y Perú. Informe de la Organización De Las Naciones Unidas Para El Desarrollo Industrial, 2006, p 6-11.

- DENG, Y; YAO, F; LI, J; MAO, Z; HU, P; LONG, L; LI, G; JI, X; SHI, S; GUAN, D; FENG, Y; CUI, L; LI, D; LIU, Y; DU, X; GUO, M; XU, L; LI, E; WANG, H and XIE, D. RACK1 Suppresses Gastric Tumorigenesis by Stabilizing the Catenin Destruction Complex. *Gastroenterology* 2012, vol. 142, n° 1, p.812-823.
- DORSAM, RT and GUTKIND, JS. G-protein-coupled receptors and cancer. *Nature Reviews Cancer* 2007, vol. 7, n° 1, p. 79-94.
- DUVAL, A and HAMELIN, R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Research* 2002, vol. 62 n° 2, p. 2447-2454.
- FAN R, DONG Y, CAO J, BAI R, ZHU Z, LI P, ZHANG J, HE X, LÜ L, YAO J, MONDAL M, SMITH GW, and DONG C. Gene expression profile in white alpaca (*Vicugna pacos*) skin. *Animal* 2011 vol. 5, n° 8, p. 1157-1161.
- FILIPOVICH, A; GEHRKE, I; POLL-WOLBECK, SJ and KREUZER, KA. Physiological inhibitors of Wnt signaling. *European Journal of Haematology* 2011, vol. 86, n° 6, p. 453-465.
- GALLI, LM; BARNES, T; CHENG, T; ACOSTA, L; ANGLADE, A; WILLERT K; NUSSE R and BURRUS, LW. Differential inhibition of Wnt-3a by Sfrp-1, Sfrp-2, and Sfrp-3. *Developmental Dynamics* 2006, vol. 235, n° 3, p. 681-90.
- GENE ONTOLOGY CONSORTIUM. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Research* 2004, vol. 1, n° 32, p. D258-D261.
- GONG, H; ZHOU, H; MCKENZIE, GW; YU, Z; CLERENS, S; DYER, JM; PLOWMAN, JE; WRIGHT, MW; ARORA, R; BAWDEN, CS; CHEN Y, LI, J and HICKFORD, JG. An updated nomenclature for keratin-associated proteins (KAPs). *International Journal of Biological Sciences* 2012, vol. 8, n° 2, p. 258-264.
- HARDY, MH and LYNE, AG. The pre-natal development of wool follicles in merino sheep. *Australian Journal of Biological Sciences* 1956, vol. 9, n° 3, p. 423-441.
- HARDY, MH. The secret life of the hair follicle. *Trends in Genetics: TGI* 1992, vol. 8, n° 2, p. 55-61.
- HUANCA, T; APAZA, N and LAZO, A. Evaluación del diámetro de fibra en alpacas de las comunidades de los distritos de Cojata y Santa Rosa - Puno. Sitio Argentino de Producción Animal, APPA-ALPA - Cuzco, Perú. 2007, p 1-8.

- HUANG, X and MADAN, A. CAP3: A DNA sequence assembly program. *Genome Research* 1999, vol. 9, n° 9, p 868-877.
- HYND, PI and MASTERS, DG. Nutrition and Wool Growth. Eds. M Freer and H. Dove, *Sheep Nutrition*, CAB International, New Zeland 2002, p 175,179.
- KIMELMAN, D and XU, W. β -Catenin destruction complex: insights and questions from a structural perspective. *Oncogene* 2006, vol. 25, n° 1, p. 7482-7491.
- KORGE, BP; GAN, SQ; MCBRIDE, OW; MISCHKE, D and STEINERT, PM. Extensive size polymorphism of the human keratin 10 chain resides in the C-terminal V2 subdomain due to variable numbers and sizes of glycine loops. *Proc Natl Acad Sci USA*. 1992, vol. 89, n° 3, p. 910-914.
- KÜHL, M; SHELDAHL, M; Park, JR; Miller, JR and Moon, RT. The Wnt/Ca²⁺ pathway: a new vertebrate Wnt signalling pathway takes shape. *Trends in Genetics: TIG* 2000, vol. 16, n° 7, p. 279-283.
- KULESSA, H; TURK, G and HOGAN, BL. Inhibition of Bmp signaling affects growth and differentiation in the anagen hair follicle. *The EMBO Journal* 2000, vol. 19, n° 24, p. 6664-6674.
- LACOLLA, D; GARCÍA, M; CORREDERA, C and BUEY, V. Estructura histológica de la piel de los camélidos sudamericanos. *Ciencia Veterinaria* 2010, vol. 12, n° 1, p. 8-15.
- LANGBEIN, L and SCHWEIZER, J. Keratins of the human hair follicle. *International Review of Cytology*, 2005, vol. 243, n° 1, p. 1-78.
- LAURIKKALA, J; PISPA, J; JUNG, HS; NIEMINEN, P; MIKKOLA, M; WANG, X; SAARIALHO-KERE, U; GALCERAN, J; GROSSCHEDL, R and THESLEFF, I. Regulation of hair follicle development by the TNF signal ectodysplasin and its receptor Edar. *Development* 2002, vol. 129, n° 10, p. 2541-53.
- LEGUÉ, E and NICOLAS, JF. Hair follicle renewal: organization of stem cells in the matrix and the role of stereotyped lineages and behaviors. *Development* 2005, vol. 18, n° 4, p. 4143-4154.
- LI, YC; KOROL, AB; FAHIMA, T and NEVO, E. Microsatellites within genes: structure, function and evolution. *Molecular Biology and Evolution* 2004, vol. 21, n° 6, p. 991-1007.

LI, YC; KOROL AB, FAHIMA T, BEILES A and NEVO E. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*. 2002, vol. 11, n° 12, p. 2453-2465.

MAGANGA, R; GILES, N; ADCROFT, K; UNNI, A; KEENEY, D; WOOD, F; FEAR, M and DHARMARAJAN, A. Secreted Frizzled related protein-4 (sFRP4) promotes epidermal differentiation and apoptosis. *Biochemical and Biophysical Research Communications* 2008, vol. 377, n° 2, p. 606-611.

MCGREGOR, BA and BUTLER, KL. Sources of variation in fiber diameter attributes of Australian alpacas and implications for fleece evaluation and animal selection. *Australian Journal of Agricultural Research*, 2004 vol. 55, n° 1, p. 433-442.

MILLAR, SE. Molecular mechanisms regulating hair follicle development. *The Journal of Investigative Dermatology* 2002, vol. 118, n° 2, p 216-25.

NAKAJIMA, H; ITO, M; MORIKAWA, Y; KOMORI, T; FUKUCHI, Y; SHIBATA, F; OKAMOTO, S and KITAMURA, T. Wnt modulators, SFRP-1, and SFRP-2 are expressed in osteoblasts and differentially regulate hematopoietic stem cells. *Biochemical and Biophysical Research Communications* 2009, vol. 390, n° 1, p. 65-70.

NÄRHI, K; JÄRVINEN, E; BIRCHMEIER, W; TAKETO, MM; MIKKOLA, ML and THESLEFF, I. Sustained epithelial beta-catenin activity induces precocious hair development but disrupts hair follicle down-growth and hair shaft formation. *Development* 2008, vol. 135, n° 6, p. 1019-1028.

NIELSEN, H. RNA Methods and Protocols. Copenhagen: Humana Press, 2011. 329 p. ISBN 978-1-58829-913-0.

ORIA, I; QUICAÑO, I; QUISPE, E and ALFONSO, L. Variabilidad del color de la fibra de alpaca en la zona altoandina de Huancavelica-Perú. *Animal Genetic Resources Information* 2009, vol. 45, n° 1, p. 79-84.

ORO, AR and HIGGINS, K. Hair cycle regulation of Hedgehog signal reception. *Developmental Biology* 2003, vol. 255, n° 1, p. 238-248.

PARKINSON, J. Expressed Sequence Tags (ESTs) Generation and Analysis. Toronto: Humana Press, 2009. 331 p. ISBN. 978-1-58829-759-4.

PARSONS, YM; PIPER, LR and COOPER, DW. Linkage relationships between keratin-associated protein (KRTAP) genes and growth hormone in sheep, *Genomics* 1994a, vol. 20, n° 1, p. 500–502.

PARSONS, YM; COOPER, DW and PIPER, LR. Evidence of linkage between high-glycine-tyrosine keratin gene loci and wool fibre diameter in a Merino half-sib family. *Animal Genetics*, 1994b, vol. 25, n° 2, p. 105-108.

PEIFER, M and MCEWEN, DG. The ballet of morphogenesis: unveiling the hidden choreographers. *Cell* 2002, vol. 109, n° 3, p. 271-274.

PÉREZ-CABAL, MA; CERVANTES, I; MORANTE, R; BURGOS, A; GOYACHE, F and GUTIÉRREZ, JP. Analysis of the existence of major genes affecting alpaca fiber traits. *Journal of Animal Science* 2010, vol. 88, n° 12, p. 3783-3788.

PERDEW, GH; VANDEN HEUVEL, JP and PETERS, JM. Regulation of Gene Expression Molecular Mechanism. New Jersey: Humana Press, 2006. 333p. ISBN 1-58829-265-7.

PURVIS, IW and FRANKLIN, IR. Major genes and QTL influencing wool and quality: a review. *Genetics, selection, evolution* 2005, vol. 37, n° 1, p. S97-107.

QUISPE, EC; RODRÍGUEZ, TC; IÑIGUEZ, LR and MUELLER, JP. Producción de fibra de alpaca, llama, vicuña y guanaco en Sudamérica. *Animal Genetic Resources Information* 2009, vol. 45, n° 1, p. 1-14.

REDDY, S; ANDL, T; BAGASRA, A; LU, MM; EPSTEIN, DJ; MORRISEY, EE and MILLAR, SE. Characterization of Wnt gene expression in developing and postnatal hair follicles and identification of Wnt5a as a target of Sonic hedgehog in hair follicle morphogenesis. *Mechanism of Development* 2001, vol. 107, n° 2, p. 69-82.

RENDL, M; POLACK, L and FUCHS, E. BMP signaling in dermal papilla cells is required for their hair follicle-inductive properties. *Genes & Development* 2008, vol. 22, p. 4, n° 1, p. 543-557.

ROGERS, GE. Electron microscope studies of hair and wool. *Annals of the New York Academy of Sciences* 1959, vol. 20, n° 83, p. 378-399.

ROGERS, GE. Hair follicle differentiation and regulation. *The International Journal of Developmental Biology* 2004, vol. 48, n° 3, p. 163-170.

ROGERS, MA; LANGBEIN, L; WINTER, H; BECKMANN, I; PRAETZEL, S and SCHWEIZER, J. Hair keratin associated proteins: characterization of a second high sulfur KAP gene domain on human chromosome 21. *Journal of Investigative Dermatology* 2004, vol. 122, n° 1, p. 147-158.

ROGERS, GE. Biology of the wool follicle: an excursion into a unique tissue interaction system waiting to be re-discovered. *Experimental Dermatology* 2006, vol. 15, n° 12, p. 931-949.

ROGERS, MA; LANGBEIN, L; PRAETZEL-WUNDER, S; WINTER, H and SCHWEIZER, J. Human hair keratin-associated proteins (KAPs). *International Review of Cytology* 2006, vol. 251 n° 1, p. 209-263.

RUIZ I ALTABA, A; PALMA, V and DAHMANE, N. Hedgehog-Gli signalling and the growth of the brain. *Nature Reviews Neuroscience* 2002, vol. 3, n° 1, p. 24-33.

RUSSEL, AJF and REDDEN, HL. The effect of nutrition on fibre growth in the alpaca. *Animal Science* 1997, vol. 64, n° 3, p. 509-512.

SABOUR ALAOUI, S; DESSIRIER, V; DE ARAUJO, E; ALEXAKI, VI; PELEKANOU, V; LKHIDER, M; STATHOPOULOS, EN; CASTANAS, E; BAGOT, M; BENSUSSAN, A and TSAPIS, A. TWEAK affects keratinocyte G2/M growth arrest and induces apoptosis through the translocation of the AIF protein to the nucleus. *PloS One* 2012, vol. 7, n° 3, p. e33609

SCHNEIDER, MR; SCHMIDT-ULLRICH, R and PAUS, R. The Hair Follicle as a Dynamic Miniorgan. *Current Biology* 2009, vol. 19, n° 1, p. R132-R142.

SHIMOMURA, Y and ITO, M. Human Hair Keratin-Associated Proteins. *Journal of Investigative Dermatology Symposium Proceedings* 2005. vol 10, p. 230-233.

SHIMOMURA, Y; AOKI, N; SCHWEIZER, J; LANGBEIN, L; ROGERS, MA; WINTER H and ITO, M. Polymorphisms in the human high sulfur hair keratin-associated protein 1, KAP1, gene family. *The Journal of Biological Chemistry*. 2002. Vol. 277, n° 47, p. 45493-45501.

ST-JACQUES, B; DASSULE, HR; KARAVANOVA, I; BOTCHKAREV, VA; LI, J; DANIELIAN, PS; MCMAHON, JA; LEWIS, PM; PAUS, R and MCMAHON, AP. Sonic hedgehog signaling is essential for hair development. *Current Biology* 1998, vol. 8, n° 19, p. 1058-1068.

- SLOMINSKI, A; WORTSMAN, J; PLONKA, PM; SCHALLREUTER, KU; PAUS, R and TOBIN, DJ. Hair follicle pigmentation. *The Journal of Investigative Dermatology* 2005, vol. 124, n° 1, p. 13-21.
- STENN, KS and PAUS, R. Controls of hair follicle cycling. *Physiological Reviews*, 2001, vol. 81, n° 1, p. 449-494.
- SUBRAMANIAN, S; MISHRA, RK and SINGH, L. 2003. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biology*. 4:R13.
- SUMMER, RMW and BIGHAM, ML. Biology of fiber growth and possible genetic and non-genetic means of influencing fiber growth in sheep and goats - a review. *Livestock Production Science* 1993, vol. 33, n° 1, p. 1-29.
- SCHWEIZER, J; BOWDEN, PE; COULOMBE, PA; LANGBEIN, L; LANE, EB; MAGIN, TM; MALTAIS, L; OMARY, MB; PARRY, DA; ROGERS, MA and WRIGHT, MW. New consensus nomenclature for mammalian keratins. *The Journal of Cell Biology* 2006, vol. 174, n° 2, p. 169-174.
- TROJANOWSKA, M; LEROY, EC; ECKES, B and KRIEG, T. Pathogenesis of fibrosis: type 1 collagen and the skin. *Journal of Molecular Medicine* 1998, vol. 76, n° 3-4, p266-274.
- THIEL, T; MICHALEK, W; VARSHNEY, RK and GRANER, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 2003, vol. 106, n° 1, p. 411-422.
- TORRES DE JASAI, J; VÉLEZ, V; ZEGARRA, J and DÍAZ, G. Caracterización de la histología de la piel de alpaca. APPA - ALPA - Cusco, Perú 2007.
- WANG, X; WANG, L; LIU, X. The Quality and Processing Performance of Alpaca Fibres. A Report for the Rural Industries Research and Development Corporation. Australian Government, 2003.
- WINTER, H; ROGERS, MA; GEBHARDT, M; WOLLINA, U; BOXALL, L; CHITAYAT, D; BABUL-HIRJI, R; STEVENS, HP; ZLOTOGORSKI, A and SCHWEIZER, J. A new mutation in the type II hair cortex keratin hHb1 involved in the inherited hair disorder monilethrix. *Human Genetics* 1997, vol. 101, n° 1, p. 165-169.

WILLERT, K and JONES, KA. Wnt signaling: is the party in the nucleus?. *Genes & Development* 2006, vol. 20, n° 11, p. 1394-1404.

WHEELER, JC. Evolution and present situation of the South American Camelidae. *Biological Journal of the Linnean Society* 1995, vol. 54, n° 3, p. 271-295.

WHEELER, JC; CHIKHI, L; BRUFORD, MW. Genetic Analysis of the Origins of Domestic South American Camelids. Documenting Domestication: New Genetic And Archaeological Paradigms 2004.

WOOD, NJ; PHUA SH and CRAWFORD, AM. A dinucleotide repeat polymorphism at the glycine-and tyrosine-rich keratin locus in sheep. *Animal Genetics* 1992, vol. 23, n° 4 p. 391.

WREN, JD; FORGACS, E; FONDON III, JW; PERTSEMLIDIS, A; CHENG, SY; GALLARDO, T; WILLIAMS, RS; SHOHET, RV; MINNA, JD and GARNER, HR. Repeat polymorphisms within gene regions: phenotypic and evolutionary implications. *American Journal of Human Genetics* 2000. Vol. 67, n° 2, p 345-356.

XIE, D; GORE, C; LIU, J; PONG, RC; MASON, R; HAO, G; LONG, M; KABBANI, W; YU, L; ZHANG, H; CHEN, H; SUN, X; BOOTHMAN, DA; MIN, W and HSIEH, JT. Role of DAB2IP in modulating epithelial-to-mesenchymal transition and prostate cancer metastasis. *Proceedings of the National Academy of Sciences of the United States of America* 2010, vol. 107, n° 6, p. 2485-90.

XU, T; GUO, X; WANG, H; HAO, F; DU, X; GAO, X and LIU, D. Differential gene expression analysis between anagen and telogen of *Capra hircus* skin based on the *de novo* assembled transcriptome sequence. *Gene* 2013, vol. 520, n° 1, p. 30-38

YAN, Q; ZHANG, Y; LI, H; WEI, C; NIU, L; GUAN, S; LI, S and DU, L. Identification of microsatellites in cattle unigenes. *Journal of Genetics and Genomics* 2008, vol. 35, n° 1, p. 261-266.

YOSHINO, K; RUBIN, JS; HIGINBOTHAM, KG; UREN, A; ANEST, V; PLISOV, SY and PERANTONI, AO. Secreted Frizzled-related proteins can regulate metanephric development. *Mechanisms of Development* 2001, vol. 102, n° 1-2, p. 45-55.

YU, ZHIDONG; JANET, EW; OLIVIA, AMW; STEVEN, WG; NAUMAN, JM; PAUL, HM; ALLAN, JN and ALLAN, JP. Annotation of sheep keratin intermediate filament genes and their patterns of expression. *Experimental Dermatology* 2011, vol. 20, n° 7, p. 582-588.

YU, ZHIDONG; GORDON, SW; NIXON, AJ; BAWDEN, CS; ROGERS, MA; WILDERMOTH, JE; MAQBOOL, NJ and PEARSON, AJ. Expression patterns of keratin intermediate filament and keratin associated protein genes in wool follicles. *Differentiation* 2009, vol. 77, n° 3, p. 307-316.

ZHANG, R; OYAJOBI, BO; HARRIS, SE; CHEN, D; TSAO, C; DENG, HW and ZHAO, M. Wnt/ β -catenin signaling activates bone morphogenetic protein 2 expression in osteoblasts. *Bone* 2013, vol. 52, n° 1, p. 145-56.