



UNIVERSIDAD PERUANA
CAYETANO HEREDIA
ESCUELA DE POSGRADO VÍCTOR ALZAMORA CASTRO

Carcterización de marcadores genéticos en genes que codifican a proteínas asociadas a queratina y evaluación de la asociación del gen KRTAP11-1 al diámetro de fibra en alpaca (*Vicugna pacos*) siguiendo una aproximación de gen candidato

**TESIS PARA OPTAR EL GRADO DE
MAGISTER EN BIOQUIMICA Y BIOLOGIA
MOLECULAR**

Bach. Francesco Foppiano Florez Estrada

LIMA – PERÚ

2016

PhD. Patricia Herrera Velit

Asesora

Jurado de tesis

Dr. Jorge Arevalo Zelada

Presidente

Dr. Michael Talledo Albuja

Vocal

Mg. Teresa Barreto Gaviria

Secretaria

Agradecimientos

A la PhD. Patricia Herrera Velit y el Dr. Jose Espinoza por darme la confianza de poder participar en este proyecto de investigación y estar siempre presente para resolver cualquier duda.

Al Dr. Jorge Rodriguez por ser una ayuda indispensable durante toda la investigación.

A mis compañeros en la Unidad de Biotecnología Molecular por su constante apoyo.

A mi familia por siempre apoyarme en cada etapa de mi vida.

A Juan Agapito del IPEN (Instituto Peruano de Energía Nuclear) por la ayuda prestada.

A CONCYTEC (Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica) que financió mis estudios de Maestría.

Al Programa Nacional de Innovación para la Competitividad y Productividad (Innovate Perú), proyecto “Una aproximación de genómica de poblaciones aplicada a la filogenia y mapeo de rasgos productivos en alpaca”, contrato N° 119-FINCYT-IA-2013 - PE

Indice

Caratula

Asesor

Indice

Lista de abreviaturas

Lista de tablas

Lista de figuras

Lista de anexos

Abreviaturas de los aminoácidos

Resumen

Abstract

I.	Introducción	1
	1.1 Planteamiento del problema	3
	1.2 Marco teórico	6
	1.2.1 ¿Qué es la lana?	6
	1.2.2 Aspectos morfológicos del folículo y la fibra	7
	1.2.2.1 El folículo y su desarrollo	7
	1.2.2.2 Estructuras del folículo	9
	1.2.2.2.1 La papila dermal	11
	1.2.2.2.2 La zona germinativa	11
	(Matriz del folículo)	
	1.2.2.2.3 Envoltura interna de la raíz (IRS)	12
	1.2.2.2.4 Envoltura externa de la raíz (ORS)	12
	1.2.2.3 Estructuras de la fibra	13
	1.2.2.3.1 Cutícula	13
	1.2.2.3.2 CórTEX	13
	1.2.2.3.3 Médula	15
	1.2.3 Queratinas	16

1.2.3.1	Formación de filamentos intermedios	18
1.2.4	Proteínas asociadas a queratina (KAPs)	19
1.2.5	Características de la fibra de alpaca	22
1.2.5.1	Aspectos morfológicos de la fibra de alpaca	22
1.2.5.1.1	Estructura folicular	22
1.2.5.1.2	Cutícula	23
1.2.5.1.3	Córtex	24
1.2.5.2	Aspectos bioquímicos de la fibra de alpaca	25
1.2.5.3	Características con valor comercial de la fibra de alpaca	25
1.2.5.3.1	Diámetro de fibra	25
1.2.5.3.2	Peso del vellón	27
1.2.5.4	Factores que afectan las características de valor comercial	28
1.2.6	Rasgos cuantitativos y mapeo de QTLs	30
1.2.7	Marcadores genéticos moleculares	33
1.2.7.1	Polimorfismos de nucleótido simple (SNP)	34
1.2.8	Parámetros genéticos de los marcadores moleculares	36
1.2.8.1	Equilibrio de Hardy – Weinberg	36
1.2.8.2	Heterocigosidad	37
1.2.8.3	Contenido de información polimórfica (PIC)	37
1.2.8.4	Desequilibrio de ligamiento (LD)	37
II.	Objetivos	39
2.1	Objetivo general	39
2.2	Objetivos específicos	39
III.	Hipotesis	39
IV	Componente I. Validación de marcadores en genes KRTAP en alpaca	40
4.1	Materiales y Métodos	40

4.1.1 Determinación de genes candidatos a estar relacionados a características de fibra	40
4.1.2 Obtención de secuencias <i>in silico</i> a partir del genoma de alpaca (Vicugna-Pacos 2.0.1)	40
4.1.3 Caracterización <i>in silico</i> de genes candidatos KRTAP en alpaca	41
4.1.4 Diseño de cebadores y estandarización de PCR	41
4.1.5 Muestras para la validación de marcadores	43
4.1.6 Análisis de secuencias	43
4.2 Resultados	45
4.2.1 Determinación de genes candidatos a estar relacionados a características de fibra	45
4.2.2 Obtención de secuencias <i>in silico</i> a partir del genoma de alpaca	46
4.2.3 Caracterización <i>in silico</i> de los genes candidatos en alpaca	47
4.2.3.1 KRTAP1-2	47
4.2.3.2 KRTAP6-1	49
4.2.3.3 KRTAP9-2	51
4.2.3.4 KRTAP11-1	54
4.2.3.5 KRTAP13-1	56
4.2.4 Diseño y prueba de cebadores	58
4.2.4.1 KRTAP1-2	58
4.2.4.1.1 Diseño y prueba del primer par de cebadores (1F y 1R)	58
4.2.4.1.2 Diseño y prueba del segundo par de cebadores (1NF y 1NR)	61
4.2.4.2 KRTAP6-1	63
4.2.4.2.1 Diseño y prueba del primer par de cebadores (6F y 6R)	63

4.2.4.2.2	Diseño y prueba del segundo par de cebadores (6NF y 6NR)	66
4.2.4.3	KRTAP9-2	68
4.2.4.3.1	Diseño y prueba del primer par de cebadores (9F y 9R)	68
4.2.4.3.2	Diseño y prueba del segundo par de cebadores (9NF y 9NR)	70
4.2.4.3.3	Diseño y prueba del tercer par de cebadores (9NNF y 9NNR)	72
4.2.4.4	KRTAP11-1	74
4.2.4.4.1	Diseño de cebadores	74
4.2.4.5	KRTAP13-1	76
4.2.4.5.1	Diseño de cebadores	76
4.2.5	Amplificación y secuenciamiento de las muestras para validación de marcadores	79
4.2.6	Identificación de sitios polimórficos y validación de marcadores	79
4.2.6.1	KRTAP1-2	80
4.2.6.2	KRTAP6-1	84
4.2.6.3	KRTAP9-2	88
4.2.6.4	KRTAP11-1	91
4.2.6.5	KRTAP13-1	94
3.2.6	Desequilibrio de ligamiento entre marcadores de diferentes genes	98
V.	Componente II. Evaluación de asociación del gen KRTAP11-1 al diámetro de fibra en alpaca huacaya	100
5.1	Materiales y Métodos	100
5.1.1	Selección de gen candidato para evaluar la asociación al diámetro de fibra en alpaca.	100
5.1.2	Muestras para el estudio de asociación	100

5.1.3	Análisis de la muestra	101
5.1.3.1	Cuantificación de las muestras de ADN	101
5.1.3.2	Parentesco entre las muestras	101
5.1.3.3	Estructuración poblacional	102
5.1.3.4	Distribución normal del diámetro de fibra en la muestra	102
5.1.3.5	Poder estadístico de la prueba	102
5.1.4	Modelo de asociación	103
5.1.5	Amplificación y secuenciamiento de las muestras	104
5.2	Resultados	105
5.2.1	Selección de KRTAP11-1 como gen candidato para evaluar la asociación al diámetro de fibra en alpaca	105
5.2.2	Análisis de la muestra	106
5.2.2.1	Parentesco	106
5.2.2.2	Estructuración poblacional	106
5.2.2.3	Distribución normal del diámetro de fibra en la muestra seleccionada para la evaluación de asociación	107
5.2.2.4	Poder estadístico de la prueba	107
5.2.3	Amplificación de las muestras	108
5.2.4	Evaluación de la asociación del gen KRTAP11-1 y el diámetro de fibra	109
VI.	Discusión	110
VII.	Conclusiones	134
VIII.	Recomendaciones	135
IX.	Referencias Bibliográficas	136
X.	Anexos	147

Lista de abreviaturas

ADN	Acido desoxirribonucleico
BMP	Proteínas morfogénicas óseas
DH	Déficit de heterocigotos
dNTP's	2'-deoxinucleotidos 5'-trifostato
EGFR	Receptor del factor de crecimiento epidérmico
EH	Exceso de heterocigotos
EHW	Equilibrio de Hardy- Weinberg
EST	Marcador de secuencia expresada
FAO	Organización de las Naciones Unidas para la Agricultura y la Alimentación
FGFR	Receptor del factor de crecimiento de fibroblasto
GWAS	Estudio de asociación del genoma completo
H	Heterocigosidad
h²	Heredabilidad en sentido estricto
H²	Heredabilidad en sentido amplio
H_E	Heterocigosidad esperada
HGT	Proteínas con alto contenido en glicina y tirosina
H_o	Heterocigosidad observada
HS	Proteínas con alto contenido de sulfuro
IF	Filamento intermedio
Indel	Inserción/Delección
IRS	Envoltura interna de la raíz
K	Queratina
KAP	Proteína asociada a queratina
KRT	Gen de queratina
KRTAP	Gen de proteína asociada a queratina
LD	Desequilibrio de ligamiento
LRT	Prueba de cociente de verosimilitud
mM	Milimolar

ng	Nanogramo
nt	Nucleótidos
NTP	Norma Técnica Peruana
ORF	Marco de lectura abierto
ORS	Envoltura externa de la raíz
P	Folículo primario
pb	Par de bases
PCR	Reacción en cadena de la polimerasa
PIC	Contenido de información polimórfica
pM	Picomolar
QTDT	Prueba de desequilibrio de transmisión cuantitativa
S	Folículo secundario
S/P	Ratio folículos secundarios: folículos primarios
S_D	Folículo secundario derivado
SNP	Polimorfismo de nucleótido simple
T_m	Temperatura de alineamiento
UBM	Unidad de Biotecnología Molecular
UHS	Proteínas con muy alto contenido de sulfuro
UTR	Región no traducida
V_A	Varianza genética aditiva
V_D	Varianza genética de dominancia
V_E	Varianza ambiental
V_F	Varianza fenotípica
V_G	Varianza genética
V_I	Varianza genética de interacción
ΔG	Energía libre de Gibbs
μl	Microlitro

Lista de Tablas

- Tabla 1.** Distribución de la alpaca en el Perú de acuerdo a la región y la variedad.
- Tabla 2.** Porcentaje de producción de fibra de alpaca de acuerdo a las categorías establecidas Según Norma Técnica Peruana (NTP 231.301:2004)
- Tabla 3.** Asignación de familias a los tres grupos principales de KAPs.
- Tabla 4.** Clasificación de la calidad de fibra de alpaca en el mercado. Según Norma Técnica Peruana (NTP 231.301:2004)
- Tabla 5.** Reportes de diámetro de fibra y coeficiente de variación para alpaca huacaya de distintas regiones del Perú y países en el mundo.
- Tabla 6.** Cálculos de heredabilidad en sentido estricto (h^2) para el diámetro de fibra de diferentes regiones del Perú y países del mundo
- Tabla 7.** Concentraciones finales de los reactivos utilizados en la estandarización de las PCRs.
- Tabla 8.** Numero de alpacas por región utilizadas para la validación de los marcadores.
- Tabla 9.** Genes de la familia KRTAP identificados en biblioteca de cDNA
- Tabla 10.** Región del genoma de alpaca donde se ubicaron las secuencias pertenecientes a los distintos genes.
- Tabla 11.** Tamaño de los ORFs encontrados y tamaño y posición de las secuencias a partir de las cuales se diseñaron los cebadores.
- Tabla 12.** Información de los cebadores diseñados para amplificar el gen KRTAP1-2.
- Tabla 13.** Condiciones a las cuales se estandarizaron las PCR para los cebadores diseñados en el gen KRTAP1-2.
- Tabla 14.** Información de los cebadores diseñados para amplificar el gen KRTAP6-1.
- Tabla 15.** Condiciones a las cuales se estandarizaron las reacciones de PCR para los cebadores en el gen KRTAP6-1.
- Tabla 16.** Información de los cebadores diseñados para amplificar el gen KRTAP9-2.
- Tabla 17.** Condiciones a las cuales se estandarizaron las PCR para los cebadores diseñados en el gen KRTAP9-2.
- Tabla 18.** Información del par de cebadores (11F y 11R) diseñados para amplificar el gen KRTAP11-1.
- Tabla 19.** Condiciones a las cuales se estandarizó la PCR para los cebadores 11F y 11R.
- Tabla 20.** Información del par de cebadores (13F y 13R) diseñados para amplificar el gen KRTAP13-1.
- Tabla 21.** Condiciones a las cuales se estandarizó la PCR para los cebadores 13F y 13R.
- Tabla 22.** Selección de cebadores para la validación de los marcadores en los 5 genes.

Tabla 23. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 10 sitios polimórficos identificados en el fragmento del gen KRTAP1-2.

Tabla 24. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP1-2.

Tabla 25. Ubicación de los sitios polimórficos dentro del gen KRTAP1-2 y los cambios en la secuencia de nucleótidos y de aminoácidos

Tabla 26. Índices de variabilidad genética para cada loci en el gen KRTAP1-2 así como la evaluación del equilibrio de Hardy – Weinberg.

Tabla 27. Variantes a nivel proteico de acuerdo a los haplotipos identificados para la proteína KAP1-2.

Tabla 28. Determinación de desequilibrio de ligamiento entre los 10 marcadores del fragmento del gen KRTAP1-2.

Tabla 29. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 9 sitios polimórficos identificados en el fragmento del gen KRTAP6-1.

Tabla 30. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP6-1.

Tabla 31. Índices de variabilidad genética para cada loci en el gen KRTAP6-1 así como la evaluación del equilibrio de Hardy – Weinberg.

Tabla 32. Determinación de desequilibrio de ligamiento entre los 9 posiciones polimórficas del fragmento del gen KRTAP6-1.

Tabla 33. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 3 sitios polimórficos identificados en el fragmento del gen KRTAP9-2.

Tabla 34. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP9-2.

Tabla 35. Ubicación de los sitios polimórficos dentro del gen KRTAP9-2 y los cambios en la secuencia de nucleótidos y de aminoácidos

Tabla 36. Índices de variabilidad genética para cada loci en el gen KRTAP9-2 así como la evaluación del equilibrio de Hardy – Weinberg.

Tabla 37. Variantes a nivel proteico de acuerdo a los haplotipos identificados para la proteína KAP9-2.

Tabla 38. Determinación de desequilibrio de ligamiento entre las 3 posiciones polimórficas del fragmento del gen KRTAP9-2.

Tabla 39. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 6 sitios polimórficos identificados en el fragmento del gen KRTAP11-1.

Tabla 40. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP11-1.

Tabla 41. Ubicación de los sitios polimórficos dentro del gen KRTAP11-1 y los cambios en la secuencia de nucleótidos y de aminoácidos

Tabla 42. Índices de variabilidad genética para cada loci en el gen KRTAP11-1 así como la evaluación del equilibrio de Hardy – Weinberg.

Tabla 43. Variantes a nivel proteico de acuerdo a los haplotipos identificados para la proteína KAP11-1.

Tabla 44. Determinación de desequilibrio de ligamiento entre las 6 posiciones polimórficas del fragmento del gen KRTAP11-1.

Tabla 45. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 6 sitios polimórficos identificados en el fragmento del gen KRTAP13-1.

Tabla 46. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP13-1.

Tabla 47. Ubicación de los sitios polimórficos dentro del gen KRTAP13-1 y los cambios en la secuencia de nucleótidos y de aminoácidos

Tabla 48. Índices de variabilidad genética para cada loci en el gen KRTAP13-1 así como la evaluación del equilibrio de Hardy – Weinberg.

Tabla 49. Variantes a nivel proteico de acuerdo a los haplotipos identificados para la proteína KAP13-1.

Tabla 50. Determinación de desequilibrio de ligamiento entre las 4 posiciones polimórficas del fragmento del gen KRTAP13-1.

Tabla 51. Determinación de desequilibrio de ligamiento entre las 32 posiciones polimórficas identificadas en los 5 genes distintos.

Tabla 52. Determinación de los coeficientes de regresión y su significancia estadística (P-value).

Lista de Figuras

Figura 1. Las dos variedades de alpaca: huacaya (izquierda) y suri (derecha)

Figura 2. Esquema del desarrollo de los dos tipos de folículo en oveja merino. A) Desarrollo del folículo primario; Flecha negra indica la posición del musculo erector y la flecha azul indica la posición de la glándula sudorípara. B) Desarrollo del folículo secundario. (Imagen modificada de Rogers, 2006).

Figura 3. Izquierda. Estructura interna del folículo piloso. Derecha. Esquema de las diferentes capas celulares. En rojo la zona germinativa, en blanco la envoltura externa de la raíz (ORS), en azul, celeste y gris la envoltura interna de la raíz (IRS), en rosado la cutícula de la fibra, en amarillo el córtex de la fibra, en morado la medula de la fibra, células con forma de escama representa a la papila dermal (Imágenes reproducidas de Rogers, 2004)

Figura 4. Esquema de la estructura interna de la fibra. En la imagen A se observa una fibra no medulada. En la imagen B se observa una fibra medulada. Traducción: Filamento intermedio (*Intermediate filament*), macrofibrilla (*Macrofibril*), exocutícula (*Exocuticle*), endocutícula (*endocuticle*), célula de la medula (*Medullary cell*), espacio intercelular (*Intercellular space*), (Imagen reproducida de Powell and Rogers, 1997).

Figura 5. Micrografía electrónica del córtex. A. Se observa un corte transversal de un fibra teñida con un tratamiento de OsO₄, uranilo acetato e hidróxido de plomo. Se observa la segmentación del córtex, O indica las células ortocorticales, P indica las células paracorticales y nr indica remanentes del núcleo. C Acercamiento de células paracorticales donde se observan las macrofibrillas con empaquetamiento hexagonal. D Acercamiento de las células ortocorticales se observa el empaquetamiento de las microfibrilla en forma de espiral, mayor empaquetamiento por ausencia de matriz entre los filamentos intermedios (Imagen reproducida de Rogers et al., 2006).

Figura 6. Se observa la estructura de una alfa queratina. Los segmentos helicoidales son 1A, 1B, 2A, 2B mientras que las regiones linker son L1, L12, L2. Las Heptades hacen referencia a los 7 residuos que conforman una vuelta en la hélice. (Imagen reproducida de Höcker, 2002)

Figura 7. Estructuras intermedias previas a la formación de un filamento intermedio. De izquierda a derecha se observa un monómero de queratina, un heterodimero, un protofilamento, una protofibrilla y un filamento intermedio. (Imagen reproducida de Wang et al., 2015)

Figura 8. En la figura se puede observar una macrofibrilla, dentro de esta se aprecian agregados de proteínas que corresponden a las KAPs (flecha negra). (Imagen reproducida de Rogers, 2004)

Figura 9. Micrografía de una sección de piel de alpaca huacaya. Se observan dos paquetes foliculares, cada uno presenta un folículo primario asociado a glándulas sudoríparas y varios folículos secundarios. (Imagen reproducida de Moore *et al.*, 2015)

Figura 10. Micrografía de la topografía de fibras finas. A la izquierda se observa el patrón de las células cuticulares en alpaca huacaya y a la derecha el patrón de las células cuticulares de oveja merino. (Imagen reproducida de Thomas *et al.*, 2012)

Figura 11. Micrografía electrónica del córtex de una fibra de alpaca. A la izquierda se observan las células ortocorticales (O) con el empaquetamiento de las macrofibrillas en forma de anillos. También se observan el grupo celular que presenta un empaquetamiento intermedio (H). A la derecha se observan las células paracorticales con el empaquetamiento de las macrofibrillas pseudohexagonal. (Imagen reproducida de Thomas *et al.*, 2012).

Figura 12. Descomposición de la varianza fenotípica y determinación de la heredabilidad en sentido amplio y estricto.

Figura 13. Identificación de SNPs por secuenciamiento directo de producto de PCR. En la caja 1 en la primera secuencia se observa un individuo homocigoto para el alelo A, en la segunda secuencia se observa un individuo heterocigoto AG y en la tercera secuencia se observa un individuo homocigoto para el alelo G. En la caja 2 se puede observar en las tres secuencias los errores de secuenciamiento a los que está sujeta esta técnica (Imagen reproducida de Vignal *et al* 2002).

Figura 14. Alineamiento entre la proteína KRTAP1-2 reportada en oveja y la proteína predicha del gen KRTAP1-2 en alpaca.

Figura 15A. Probable secuencia del gen KRTAP1-2 en alpaca. En letras negritas se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, en verde se muestra la secuencia de la biblioteca de cDNA (región 3'UTR) y en rojo la secuencia señal de poliadenilación (AAUAAA). **15B.** Secuencia traducida de aminoácidos de la proteína KAP1-2 en alpaca obtenida a partir de la región codante identificada *in silico*.

Figura 16. Alineamiento entre la proteína KRTAP6-1 reportada en oveja y la proteína predicha del gen KRTAP6-1 en alpaca.

Figura 17A. Probable secuencia del gen KRTAP6-1 en alpaca. En letras negritas se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, y en rojo la secuencia señal de poliadenilación (AAUAAA). **17B.** Secuencia traducida de aminoácidos de la proteína KAP6-1 en alpaca obtenida a partir de la región codante identificada *in silico*.

Figura 18. Alineamiento entre la proteína KRTAP9-2 reportada en humano y la proteína predicha del gen KRTAP9-2 en alpaca.

Figura 19A. Probable secuencia del gen KRTAP9-2 en alpaca. En letras negritas se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, y en rojo la secuencia señal de poliadenilación (AAUAAA). **19B.** Secuencia traducida de aminoácidos de la proteína KAP9-2 en alpaca obtenida a partir de la región codante identificada *in silico*.

Figura 20. Alineamiento entre la proteína KRTAP11-1 reportada en oveja y la proteína predicha del gen KRTAP11-1 en alpaca.

Figura 21A. Probable secuencia del gen KRTAP11-1 en alpaca. En letras negritas se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, en verde se muestra la secuencia de la biblioteca de cDNA (región 3'UTR) y en rojo la secuencia señal de poliadenilación (AAUAAA). **21B.** Secuencia traducida de aminoácidos de la proteína KAP11-1 en alpaca obtenida a partir de la región codante identificada *in silico*.

Figura 22. Alineamiento entre la proteína KRTAP13-1 reportada en humano y la proteína predicha del gen KRTAP13-1 en alpaca.

Figura 23A. Probable secuencia del gen KRTAP13-1 en alpaca. En letras negritas se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, en verde se muestra la secuencia de la biblioteca de cDNA parte de la región 3'UTR y en rojo la secuencia señal de poliadenilación (AAUAAA). **23B.** Secuencia traducida de aminoácidos de la proteína KAP11-1 en alpaca obtenida a partir de la región codante identificada *in silico*.

Figura 24. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 1F y 1R a diferentes concentraciones de magnesio y temperatura, todos los productos se encuentran en un tamaño cercano a 800 pb. El pozo con la M hace referencia al marcador de peso molecular (100bp plus), El pozo con la B hace referencia a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 1000 pb.

Figura 25. Cromatograma con patrón anómalo de la secuencia de KRTAP1-2 obtenido con los cebadores 1F y 1R. Se observa un gran número de picos superpuestos, esto se debe que el cromatograma corresponde a un individuo heterocigoto para la inserción/delección.

Figura 26. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 1NF y 1NR a diferentes concentraciones de magnesio y temperatura, las bandas se encuentran entre 984 y 861 pb. El pozo con la M hace referencia al marcador de peso molecular (123bp), el pozo en blanco entre muestras corresponde a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 984 pb.

Figura 27. Secuencia a partir de la cual se diseñaron ambos pares de cebadores y las posiciones de estos dentro de la misma. En amarillo el primer par de cebadores (1F y 1R), en rojo el segundo par de cebadores (1NF y 1NR), en negrita la región codante y un asterisco indica la posición de la inserción/delección.

Figura 28. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 6F y 6R (KRTAP6-1) a diferentes concentraciones de magnesio y temperatura, se observa una banda cerca a las 800 pb. El pozo con la M hace referencia al marcador de peso molecular (100bp plus), El pozo con la B hace referencia a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 1000 pb.

Figura 29. Cromatograma con patrón anómalo de la secuencia de KRTAP6-1 obtenido con los cebadores 6F y 6R. Se observa un gran número de picos superpuestos, esto se

debe que el cromatograma corresponde a un individuo heterocigoto para la inserción/delección.

Figura 30. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 6NF y 6NR a diferentes concentraciones de magnesio y temperatura, las bandas se encuentran entre 861 y 984 pb. El pozo con la M hace referencia al marcador de peso molecular (123bp), el pozo en blanco entre muestras corresponde a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 984 pb.

Figura 31. Secuencia a partir de la cual se diseñaron ambos pares de cebadores y las posiciones de estos dentro de la misma. En amarillo el primer par de cebadores (6F y 6R), en rojo el segundo par de cebadores (6NF y 6NR), en negrita se muestra la región codante y un asterisco indica la posición de la inserción/delección.

Figura 32. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 9F y 9R (KRTAP9-2) a diferentes temperaturas, se observan bandas entre 400 y 500 pb. El pozo con la M hace referencia al marcador de peso molecular (100bp plus), El último pozo contiene la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 500 pb. Se observa una distorsión en la corrida del gel a causa de un alto voltaje.

Figura 33. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 9NF y 9NR (KRTAP9-2) a diferentes temperaturas. El pozo con la M hace referencia al marcador de peso molecular (100bp plus), los espacios en blanco entre las muestras corresponden a las reacciones en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 800 pb.

Figura 34. Cromatograma con patrón anómalo de la secuencia de KRTAP9-2 obtenido con los cebadores 9NF y 9NR. Se observa un gran número de picos superpuestos, esto se debe que el cromatograma corresponde a un individuo heterocigoto para la inserción/delección.

Figura 35. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 9NNF y 9NNR a diferentes concentraciones de magnesio y temperatura. El pozo con la M hace referencia al marcador de peso molecular (123bp), el pozo en blanco entre muestras corresponde a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 861 pb

Figura 36. Secuencia a partir de la cual se diseñaron los tres pares de cebadores y las posiciones de estos dentro de la misma. En verde el primer par de cebadores (9F y 9R), en amarillo el segundo par de cebadores (9NF y 9NR), en rojo el tercer par de cebadores (9NNF y 9NNR) y un asterisco indica la posición de la inserción/delección.

Figura 37. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 11F y 11R (KRTAP11-1) a diferentes concentraciones de magnesio y temperatura. El pozo con la M hace referencia al marcador de peso molecular (100bp plus), El pozo con la B hace referencia a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 1100 pb.

Figura 38. Secuencia a partir de la cual se diseñó el par de cebadores y las posiciones de estos dentro de la misma. En amarillo la posición de los cebadores 11F y 11R.

Figura 39. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 13F y 13R (KRTAP13-1) a diferentes concentraciones de magnesio y temperatura. El pozo con la M hace referencia al marcador de peso molecular (100bp plus), El pozo con la B hace referencia a la reacción en blanco y la flecha blanca indica la banda del marcador de peso molecular que corresponde a 1000 pb.

Figura 40. Secuencia a partir de la cual se diseñó el par de cebadores y las posiciones de estos dentro de la misma. En amarillo la posición de los cebadores 13F y 13R.

Figura 41 Diagrama de barras obtenido de Structure considerando los 10 microsatelites y los 152 individuos, asumiendo 7 *clusters* ($K = 7$). En el eje X se muestran los 152 individuos analizados, cada color corresponde a uno de los 7 *clusters*. Se puede observar que a lo largo del eje x no se observa ninguna diferencia en las proporciones de cada *cluster*, esto indica la ausencia de estructuración poblacional

Figura 42. Se muestra la relación entre el poder de la prueba y el efecto del gen en función de las frecuencias alélicas del marcador seleccionado. En el eje Y se observa la proporción entre resultados positivo y el total de resultados (Poder estadístico). En el eje X se muestra el efecto del gen expresado como la heredabilidad en sentido estricto (h^2). (Fuente: Diego Veliz Otani, no publicado)

Lista de anexos

Anexo 1. Secuencias a partir de las cuales se identificaron los genes en el genoma de referencia de alpaca.

Anexo 2. Relación de las 96 muestras utilizadas para la validación de los marcadores

Anexo 3. Alineamientos entre las proteínas predichas para KRTAP6-1 y KRTAP11-1 de alpaca y las proteínas ortólogas en humano.

Anexo 4. Cromatogramas de los genotipos para cada una de las posiciones polimórficas identificadas

Anexo 5. Resultados del análisis de estructuración poblacional

Anexo 6. Datos de entrada para el estudio de asociación

Abreviaturas de aminoácidos

A	Ala	Alanina
C	Cys	Cisteina
D	Asp	Aspartato
E	Glu	Glutamato
F	Phe	Fenilalanina
G	Gly	Glicina
H	His	Histidina
I	Ile	Isoleucina
K	Lys	Lisina
L	Leu	Leucina
M	Met	Metionina
N	Asn	Asparagina
P	Pro	Prolina
Q	Gln	Glutamina
R	Arg	Arginina
S	Ser	Serina
T	Thr	Treonina
V	Val	Valina
W	Trp	Triptofano
Y	Tyr	Tirosina

Resumen

La fibra de alpaca (*Vicugna pacos*) es un recurso socioeconómico fundamental para las poblaciones que habitan las regiones altoandinas del Perú. Por ello es importante la implementación de actividades y programas de mejoramiento genético con marcadores genéticos de ADN y genes relacionados a la formación de la fibra y a las diversas características de importancia económica. El presente trabajo tiene como objetivo la identificación y caracterización de marcadores genéticos de polimorfismos de nucleótido únicos (SNPs) en las secuencias de cinco genes de la familia de proteínas asociadas a queratina; KTRAP1-2, KRTAP6-1, KRTAP9-2, KRTAP11-1 y KRTAP13-1 en alpaca y la evaluación de la asociación entre el gen KRTAP11-1 y el diámetro de la fibra.. Se identificaron 27 marcadores genéticos SNPs en los 5 genes estudiados, los cuales fueron caracterizados en muestras de alpacas provenientes de las regiones de Puno, Cusco, Huancavelica, Apurimac, Arequipa y Junin (n = 96). Todos los marcadores presentan frecuencias alélicas mayores al 5% en el alelo menor y valores de heterocigosidad intermedios a altos ($H_o = 0.2 - 0.48$). No se encontró una asociación significativa del gen KRTAP11-1 y el diámetro de fibra ($p = 0.8$ y $p = 0.723$) en una muestra de 152 alpacas huacaya blancas de la región de Puno.

Abstract

Alpaca (*Vicugna pacos*) fibre is an important socio-economic resource of Andean regions of Peru. In order to implement genetic enhancement programs, it is important to work with genes and molecular markers involved in the formation of the fibre and other features of economic relevance. This study aims to identify and characterize SNP genetic markers in five genes of the Keratin Associated Proteins family in alpaca (KRTAP1-2, KRTAP6-1, KRTAP9-2, KRTAP11-1 and KRTAP13-1) and to evaluate the association between KRTAP11-1 and fibre diameter. Within the 5 genes analyzed, 27 SNP genetic markers were identified using 96 alpacas from Puno, Cusco, Junin, Apurimac, Huancavelica and Arequipa. All markers presented allele frequencies higher than 5% on the least prevalent allele; and values of heterozygosity ranging from intermediate to high ($H_o = 0.2 - 0.48$). No significant association between KRTAP11-1 and fibre diameter ($p = 0.8$ and $p = 0.723$) was observed in a sample of 152 Huacaya white alpacas.

I. Introducción

La alpaca (*Vicugna pacos*) pertenece, junto a la vicuña, la llama y el guanaco al grupo de camélidos sudamericanos. La vicuña y el guanaco son especies silvestres mientras que la alpaca y la llama son especies domesticadas (Quispe *et al.*, 2009a). Originalmente la distribución de estos animales se limitaba a la región del altiplano de Sudamérica, sin embargo hoy en día es posible encontrar alpacas y llamas en Australia, Estados Unidos, Polonia entre otros países (Quispe *et al.*, 2009a). La alpaca tiene su origen hace aproximadamente 6000 años. Un estudio con marcadores mitocondriales reveló que el origen de este animal se dio por procesos de domesticación a partir de la vicuña, por lo que se le asignó al género *Vicugna* (Kadwell *et al.*, 2001; Marin *et al.*, 2007).

Existen dos variedades distintas de alpaca, las huacaya y las suri (Figura 1). Estas se diferencian en el tipo de fibra que producen. La fibra de la huacaya se caracteriza por presentar una curvatura pronunciada (rizo), mientras que en el caso de las suri la curvatura es menor y la fibra tiene una apariencia más lustrosa (Presciuttini *et al.*, 2010). Otra diferencia importante es la abundancia de estas dos variedades, la alpaca huacaya se encuentra mucho más representada en el Perú (alrededor del 85% de alpacas) (Quispe *et al.*, 2009) (Tabla 1). Para la herencia de este carácter se han propuesto varios modelos, el más aceptado propone la existencia de dos genes. Se ha determinado también que el carácter suri es dominante sobre el carácter huacaya (Presciuttini *et al.*, 2010).



Figura 1. Las dos variedades de alpaca: huacaya (izquierda) y suri (derecha) (Fuente: Peruhandicraft Inc.)

Ambas variedades son criadas principalmente para la producción de fibra, aunque también se consume su carne (Quispe *et al.*, 2009a). La fibra de alpaca es considerada hoy en día como un artículo de lujo a nivel mundial, por lo cual su producción y uso no se limitan solo a Sudamérica (Wang *et al.*, 2003). Una de las características más importantes de esta fibra es, que es considerada hipo alérgica debido a que no contiene lanolina, cera que se encuentra en lana ovina (Panigrahi y Kushwaha, 2013). Otra característica importante radica en su estructura interna, la cual presenta espacios aéreos. Estos espacios hacen que la fibra tenga una baja conductividad térmica, por lo que las prendas de alpaca son muy útiles para climas fríos (Czaplicki, 2012). Por otro lado cuenta con una amplia gama de colores naturales, 22 en total. Esta característica es importante debido al aumento en los últimos años de la demanda por prendas con colores naturales (Morante *et al.*, 2009).

La fibra de alpaca crece a partir de un tejido llamado folículo piloso (Orwin, 1979). Este tejido tiene una estructura muy conservada en los distintos taxones de los mamíferos (Rogers, 2004). Sin embargo el pelo o la fibra pueden variar considerablemente de especie a especie, estas diferencias se deben en parte a las familias de proteínas estructurales que

la componen (Khan *et al.*, 2014). Dos familias de proteínas se ven expresadas principalmente en las células estructurales de la fibra, las alfa queratinas y las proteínas asociadas a queratinas (KAPs) (Powell y Rogers, 1997). Las KAPs han demostrado ser proteínas con una alta variabilidad a diferencia de la queratinas (Khan *et al.*, 2014).

Con el fin de mejorar la calidad de la fibra, su producción y otras características de interés comercial, es importante determinar qué genes están relacionados a estas. De esta forma se pueden implementar planes de mejoramiento genético (Quispe *et al.*, 2013). Una aproximación para el establecimiento de estos planes es a través del uso de marcadores genéticos. Los marcadores genéticos son herramientas que permiten caracterizar la relación genotipo – fenotipo, mediante estudios de asociación o de ligamiento (Cardon y Bell, 2001). De los diferentes tipos de marcadores que existen, los marcadores de polimorfismo de nucleótido simple (SNP) han demostrado ser muy útiles para esta clase de estudios (Balding, 2006). Esto debido a que son muy frecuentes en el genoma y se pueden encontrar con mayor facilidad en regiones codantes (Vignal *et al.*, 2002).

El presente estudio tiene como objetivo identificar marcadores del tipo SNP en genes que codifican para las proteínas asociadas a queratinas (KAPs) en alpaca, para luego seleccionar un gen candidato y evaluar si éste se encuentra asociado al diámetro de fibra.

1.1 Planteamiento del problema

El Perú es el productor número uno a nivel mundial de fibra de alpaca, produciendo alrededor de 3399 toneladas anuales y teniendo alrededor de 800 mil productores (Quispe *et al.*, 2009a). Esto se debe a que posee aproximadamente 3 millones de alpacas, constituyendo alrededor del 80% de la población mundial (Tabla 1) (FAO, 2005). Por lo general las alpacas son criadas en las regiones altoandinas del país, por lo que representa un recurso socio económico fundamental para estas regiones (Tabla 1). El sistema de

producción de alpaca es de tipo extensivo, utilizándose campos nativos de pastoreo. Por lo general los rebaños son mixtos incluyendo ovinos y llamas (Quispe, 2005). Existen dos características importantes en la producción de fibra de alpaca: el peso del vellón, que representa la cantidad de producción; y el diámetro de fibra que refleja la calidad de ésta. Para los criadores de alpaca, lo ideal es tener vellones pesados y con fibras que presenten un diámetro de fibra pequeño y homogéneo (Quispe *et al.*, 2013).

Tabla1. Distribución de la alpaca en el Perú de acuerdo a la región y la variedad.

Region	Huacaya	%	Suri	%	Total	%
Puno	1 392 600	56.5	289 319	66.6	1 681 919	58
Cusco	304 797	12.4	41 431	9.5	346 228	11.9
Junin	47 620	1.9	7 970	1.8	55 590	1.9
Arequipa	207 810	8.4	26 561	6.2	234 371	8.1
Ayacucho	113 332	4.6	16 174	3.7	129 506	4.5
Apurimac	66 744	2.7	18 204	4.2	84 948	2.9
Huancavelica	306 968	12.4	23 660	5.4	330 628	11.4
Lima	26 333	1.1	11 377	2.6	37 710	1.3
Total	2 466 204	100	434 696	100	2 900 900	100

Fuente: FAO, 2005

A pesar de representar un recurso tan importante para el Perú, la cantidad en la producción (peso de vellón) al igual que la calidad (diámetro de fibra) se ha ido deteriorando con el tiempo (De los Ríos, 2006). La baja producción se debe, en parte, a la falta de adopción de tecnología moderna, las esquilas por lo general son realizadas con tijeras manuales, mecánicas u otros implementos más rudimentarios (Quispe, 2005; observación del autor). En cuanto a la pérdida de calidad, se ha observado un aumento en el diámetro de fibra probablemente causado por eventos de hibridación con llamas (Quispe *et al.*, 2013; De los Ríos, 2006; Wheeler, 2012). De los Ríos encontró que solo un 30% de la producción de fibra peruana corresponde a fibra categorizada como fina (Tabla 2). Por otro lado un hallazgo arqueológico de alpacas momificadas, encontró que en la época antes de la colonia existía una variedad de alpaca con un promedio de diámetro de fibra de 17 μm ,

similar al reportado en fibra cashmere, esta variedad hoy en día se ha perdido (Wheeler, 1995).

Tabla2. Porcentaje de producción de fibra de alpaca de acuerdo a las categorías establecidas Según Norma Técnica Peruana (NTP 231.301:2004)

Categoría	Diametro	Frecuencia de producción en el Peru
Alpaca Baby	Hasta 23 μm	12%
Alpaca Fleece	23.1 a 26.5 μm	22%
Alpaca Medium Fleece	26.6 a 29 μm	46%
Alpaca Huarizo	29.1 a 31.5 μm	20%
Alpaca Gruesa	> 31.5 μm	-

Fuente: De los Ríos, 2006

Con fines de mejorar la producción y la calidad, se deben adoptar tecnologías más avanzadas para la recolección de la fibra como también para su clasificación. Además es importante la implementación de planes de mejoramiento genético orientados a disminuir el diámetro de fibra y aumentar la producción (Quispe *et al.*, 2009b). Para esto es necesario la generación de un gran número de marcadores genéticos, preferentemente en genes estructurales y reguladores de la fibra. En el caso de la alpaca el número de marcadores existente es muy escaso.

En cuanto a la información genética que se tiene de la alpaca, se cuenta con un genoma en el Genbank a una cobertura de 22X publicado por Washington University Genome Sequencing Center y Broad Institute [número de acceso: *PRJNA30567*]. Este genoma ya se encuentra ensamblado en *Scaffolds* y anotado de manera automática. Se ha empezado un mapa físico en base a este genoma, donde algunos genes y *scaffolds* han sido asignados a cromosomas (Avila *et al.*, 2014). También se han reportado algunos marcadores del tipo SNP asociados a color de fibra (Feeley y Munyard, 2009; Guridi *et al.*, 2011) y marcadores de tipo microsatélite (Sasse *et al.*, 2000; McPhartlan *et al.*, 1998; Obrique *et al.*, 1998; Penedo *et al.*, 1998; Penedo *et al.*, 1999a; Penedo *et al.*, 1999b; Bustamante,

2003). Además, en el Genebank se encuentran 7,286 EST reportados por investigadores de la Universidad Agrícola de Shanxi, bajo la biblioteca ASCD.

En este estudio se generó marcadores en uno de los principales grupos de proteínas estructurales de la fibra en alpaca, de acuerdo a reportes previos de asociación a distintas características en otras especies productoras de fibra. Además de la generación de marcadores, se evaluó si existe asociación entre el gen KRTAP11-1 y el diámetro de fibra siguiendo una aproximación de gen candidato.

1.2 Marco teórico

1.2.1 ¿Qué es la lana?

El término lana o vellón hace referencia por lo general al conjunto de fibras provenientes de la oveja, aunque el término se aplica también a otras especies productoras de fibra como por ejemplo cabra, alpaca, yak, entre otros (Höcker, 2002). La fibra en las especies productoras de lana crece a partir de folículos de la misma manera como crece el cabello humano (Rogers, 2006). Sin embargo, existen diferencias estructurales entre las fibras de diferentes especies (Thomas *et al.*, 2012). Al igual que en el caso de la oveja, la lana de alpaca protege al animal contra condiciones ambientales extremas (radiación solar, granizo, lluvias entre otros) (Grigg *et al.*, 2004).

Los humanos han domesticado varias especies (ovejas, alpaca, yak) para la producción de lana con fines orientados a la industria textil. Hoy en día la industria textil representa un rubro importante para la economía de varios países, por lo que el estudio de la lana es muy importante.

1.2.2 Aspectos morfológicos del folículo y la fibra

1.2.2.1 El folículo y su desarrollo

El folículo es un tejido característico de los mamíferos y su estructura se ha mantenido constante en los diferentes taxones que conforman esta clase. La función de este tejido es la formación del pelo, el cual de acuerdo a la especie puede presentar adaptaciones especiales como por ejemplo fibra o escamas modificadas (Khan *et al*, 2014).

La primera fase del desarrollo del folículo se da por una proliferación de células epidérmicas que forman la placoda folicular; debajo de está, células dermales comienzan a agregarse. Ambas líneas celulares crecen en dirección a la dermis (Rogers, 2006). Progresivamente las células dermales migran internándose en el brote epitelial, formando la pre – papila. Conforme el folículo va alargándose y descendiendo, las células epiteliales envuelven la pre-papila formando así la papila dermal y el bulbo del folículo (Rogers, 2006). A partir del bulbo se originan el resto de estructuras, mediante migración y diferenciación celular (Rogers, 2006).

Los primeros folículos en formarse se denominan folículos primarios (P), seguidos por los folículos secundarios (S) y luego los folículos secundarios derivados (SD) (Hardy y Lyne, 1956). Los folículos primarios se diferencian por presentar una glándula sudorípara y un músculo erector (Figura 2). Los folículos secundarios carecen de estas estructuras, pero a partir de ellos crecen varios folículos secundarios derivados (Figura 2) (Rogers, 2004).

El volumen de la lana va acorde al número folículos secundarios. En oveja Merino se determinó que los folículos primarios empiezan a desarrollarse en el feto a los 70 días, mientras que los folículos secundarios se desarrollan alrededor del día 85. En oveja los folículos primarios ya se encuentran desarrollados al momento de nacer el animal

(Rogers, 2006). Se ha determinado también que los folículos primarios dan origen a fibras más anchas y por lo general meduladas, mientras que los folículos secundarios producen fibras más finas (Moore *et al*, 2015).

Una vez que el folículo está maduro éste pasa por distintas fases que conforman el ciclo del folículo. Por lo general este ciclo consta de tres fases: Anagen, catagen y telogen (Rogers, 2006). La primera fase se caracteriza por la formación de la fibra, la segunda por una involución del folículo donde mediante apoptosis parte del bulbo folicular se destruye; y la tercera fase es donde la fibra se desprende. Una vez que la fibra se desprende del folículo comienza nuevamente el ciclo en la fase anagen (Paus y Cotsarelis, 2004). Animales como el ratón o el conejo presentan las tres fases mientras que animales productores de fibra presentan folículos predominantemente en una fase anagen muy larga (Rogers, 2006).

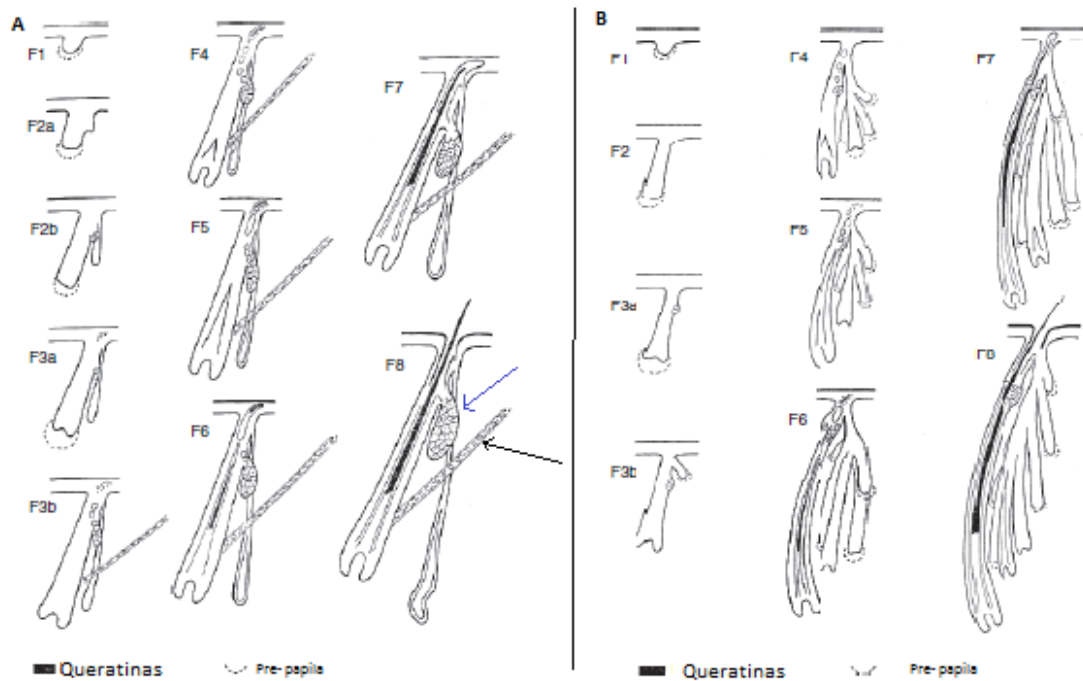


Figura 2. Esquema del desarrollo de los dos tipos de folículo en oveja merino. A) Desarrollo del folículo primario; Flecha negra indica la posición del músculo erector y la flecha azul indica la posición de la glándula sudorípara. B) Desarrollo del folículo secundario. (Imagen modificada de Rogers, 2006).

1.2.2.2 Estructuras del folículo

Las estructuras principales del folículo y la fibra son: La zona germinativa, la papila dérmica, la envoltura externa de la raíz (ORS), la envoltura interna de la raíz (IRS), el córtex, la cutícula y la médula (Figuras 3 y 4).

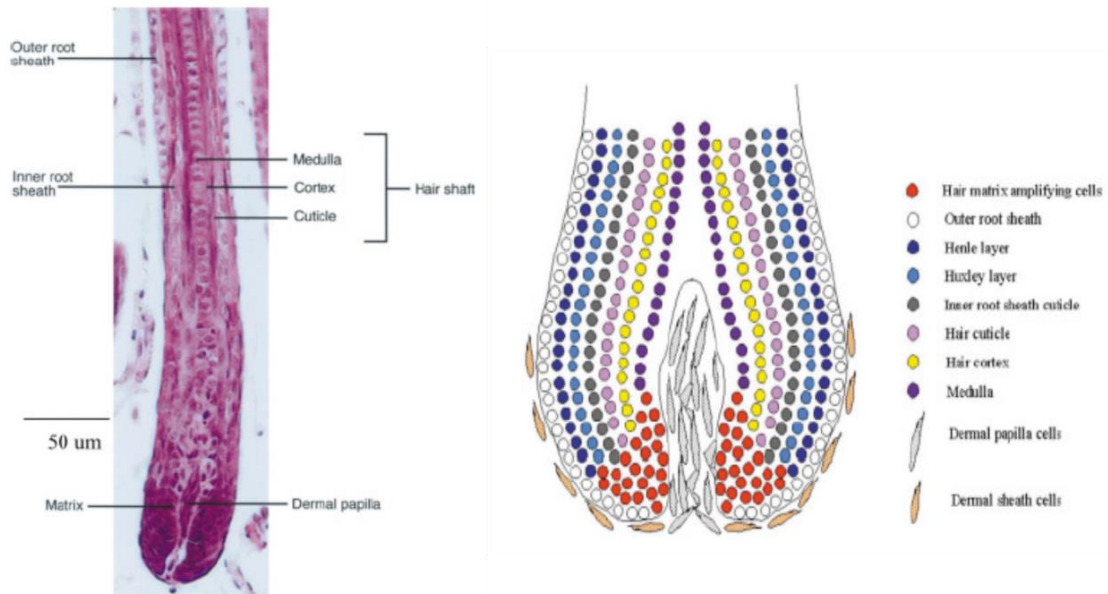


Figura 3. Izquierda. Estructura interna del folículo piloso. Derecha. Esquema de las diferentes capas celulares. En rojo la zona germinativa, en blanco la envoltura externa de la raíz (ORS), en azul, celeste y gris la envoltura interna de la raíz (IRS), en rosado la cutícula de la fibra, en amarillo el córtex de la fibra, en morado la medulla de la fibra, células con forma de escama representa a la papila dermal (Imágenes reproducidas de Rogers, 2004)

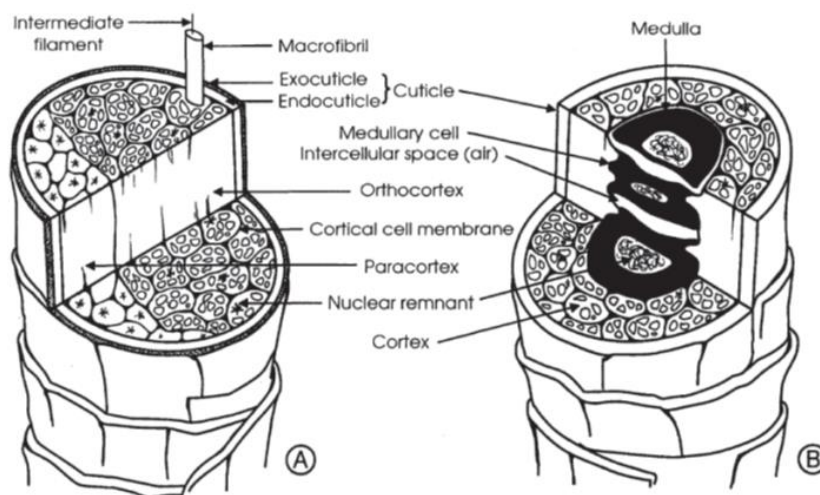


Figura 4. Esquema de la estructura interna de la fibra. En la imagen A se observa una fibra no medulada. En la imagen B se observa una fibra medulada. Traducción: Filamento intermedio (*Intermediate filament*), macrofibrilla (*Macrofibril*), exocutícula (*Exocuticle*), endocutícula (*endocuticle*), célula de la medulla (*Medullary cell*), espacio intercelular (*Intercellular space*), (Imagen reproducida de Powell and Rogers, 1997).

1.2.2.2.1 La papila dermal

La papila dermal está compuesta por células mesenquimales y se encuentra en la zona basal del folículo (Orwin, 1979). Esta región regula el número de células de la zona germinativa y provee de estímulos para el crecimiento de la fibra, diferenciación celular y transición entre las distintas etapas del ciclo folicular (Rogers, 2004). El número de células presente en la papila dermal está correlacionado con el tamaño del folículo y el diámetro de la fibra (Elliot *et al.*, 1999). Se ha observado que células de la papila dermal pueden recuperar folículos o inducir la neogenesis folicular (Jahoda *et al.*, 1993; McElwee *et al.*, 2003). Se han encontrado varias proteínas que modulan la diferenciación de células y la transición entre fases (anagena, catagena y telogena); por ejemplo, la actividad de la fosfatasa alcalina (Handjiski *et al.*, 1994) Otras moléculas involucradas en la señalización de esta región son: Adenosina, BMPs, Sonic Hedgehog, Wnts, Notch, BMPR1A, EGFR, FGFR, TGFR24 (Rogers, 2004; Hwang *et al.*, 2012; Xavier *et al.*, 2013).

1.2.2.2.2 La zona germinativa (matriz del folículo)

Esta zona se ubica en el bulbo del folículo (base) y envuelve a la papila dermal (Orwin, 1979). Las células presentes en esta zona tienen como destino la división mitótica (proliferación celular) o la diferenciación celular (Rogers 2004). Al diferenciarse, las células comienzan a migrar hacia la parte superior del folículo dando origen a la envoltura interna de la raíz (IRS), células corticales y células de la cutícula (Rogers, 2006). Una proteína involucrada en el proceso de diferenciación es β_1 - Integrina, se ha observado una alta expresión en células epidermales que al diferenciarse disminuye significativamente. La expresión de esta proteína evita que las células comiencen la migración que conducirá a su diferenciación (Zhu *et al.* 1999).

1.2.2.2.3 Envoltura interna de la raíz (IRS)

La envoltura interna de la raíz está constituida por 3 diferentes capas; la capa de Henle (externa), la capa de Huxley (intermedia) y la capa de la cutícula de la IRS (interna) (Orwin, 1979). La IRS se ubica entre la envoltura externa y la cutícula de la fibra. Las células que constituyen esta envoltura se originan a partir de la zona germinativa (Rogers, 2004). El proceso de diferenciación de la IRS, al igual que el de las células de la fibra, consta de una excesiva producción de queratinas (queratinización). Sin embargo la IRS está caracterizada también por la síntesis de tricohialina, una proteína que no está presente en el córtex. Estas diferencias se deben a que esta envoltura es degradada cuando la fibra emerge a la superficie, la presencia de tricohialina posibilita esa degradación al ser más sensible a la acción de proteasas (Powell y Rogers., 1997; Rogers, 1964). Se ha identificado que el factor de transcripción GATA-3 cumple un rol muy importante en la diferenciación de esta línea celular (Kaufman *et al.*, 2003). La IRS es responsable por la topografía de la superficie de la fibra, además acompaña a está durante su crecimiento (Rogers, 2004).

1.2.2.2.4 Envoltura externa de la raíz (ORS)

Esta envoltura se encuentra rodeando a la IRS. La ORS no se origina a partir de la zona germinativa (Orwin, 1979). Esta estructura es continua con la epidermis, sin embargo la capa de células adyacente a la IRS presenta ciertas características que la diferencian de la epidermis. Una de estas diferencias es la presencia de filamentos intermedios de queratina orientados de tal forma que rodean al folículo (Rogers, 2004). Esta capa se conoce como “la capa acompañante” pero aún no se conoce su función. Una hipótesis es que la IRS se mueve con respecto a la ORS (Rogers 2004). Se ha determinado que la ORS provee de células germinativas a la matriz del folículo (Rufaut *et al.*, 2012). Por otro lado se ha visto que en caso de lesión, estas migran fuera del folículo (Paus y Cotsarelis, 1999).

1.2.2.3 Estructuras de la fibra

1.2.2.3.1 Cutícula

La cutícula es la estructura que se encuentra en la periferia de la fibra, esta rodea y protege al córtex. Varias características físicas de la fibra dependen de ésta como, por ejemplo, la hidrofobicidad (Orwin, 1979). El ancho de la cutícula puede ser de una célula o más de acuerdo a la especie. En oveja merino esta consta de una única capa mientras que en alpaca se han encontrado que varía entre dos o tres (Thomas *et al* 2012). Las células que las conforman tienen una apariencia escamosa sobreponiéndose una sobre otra. La forma de estas células permite la diferenciación entre fibras, cada especie presenta un patrón único (Thomas *et al.*, 2012). Estas células tienen su origen en la matriz del folículo y pasan por un evento de diferenciación diferente al resto debido a la baja expresión de filamentos intermedios de queratina en comparación al córtex e IRS (Rogers, 2006). Dentro de la cutícula se han identificado 3 capas: la endocutícula, la exocutícula y la capa “A”. Estas se diferencian en cuanto a sus componentes, la endocutícula contiene residuos del citoplasma, la exocutícula contiene dos proteínas principales KAP5 y KAP10 mientras que la capa “A” presente como componentes principales las proteínas lorocrina e involucrina (Rogers, 2004). Adheridos a la capa “A” se han identificado una capa de ácidos grasos, siendo el mayor componente en oveja el ácido graso 18-metileicosanoico (Jones and Rivet, 1997).

1.2.2.3.2 Córtex

El córtex constituye el volumen de la fibra y es responsable por propiedades como la elasticidad, el diámetro y el rizo de ésta (Orwin, 1979). Está compuesto por las células corticales, las cuales tienen su origen en la matriz del folículo. Las células de la matriz migran hacia la parte superior del folículo sufriendo un proceso de diferenciación

conocido como queratinización (Rogers 2004). Este proceso consiste en la excesiva producción de filamentos de queratina y proteínas asociadas a queratina (KAPs), hasta el punto en que la célula muere a causa de la gran concentración de estas proteínas (Plowman *et al.*, 2015). En primer lugar se expresan las queratinas formando filamentos intermedios. Una vez que hay una cantidad grande de estas proteínas, las células comienzan a sintetizar las KAPs. Se ha observado que la presencia de KAPs y queratinas varía de acuerdo al estadio de diferenciación en el cual se encuentra la célula (Plowman *et al.*, 2015). Una vez que estas células están diferenciadas constan de microfibrillas de queratinas embebidas en una matriz compuesta por KAPs. Entre las células corticales se han identificado dos clases diferentes, las células paracorticales y ortocorticales (Figura 5) (Orwin, 1979). La diferencia radica en el empaquetamiento que presentan los filamentos intermedios de queratina. Las células paracorticales presentan filamentos con un menor grado de empaquetamiento y una mayor presencia de proteínas de la matriz (KAPs) en comparación con las células ortocorticales (Plowman *et al.*, 2007). En cuanto a las proteínas de la matriz, se ha determinado que las células paracorticales presentan más KAPs con residuos de cisteína mientras que las células ortocorticales presentan más residuos de glicina y tirosina. En oveja merino se ha asociado una mayor presencia de células paracorticales con un diámetro más pequeño de la fibra (Plowman *et al.*, 2007).

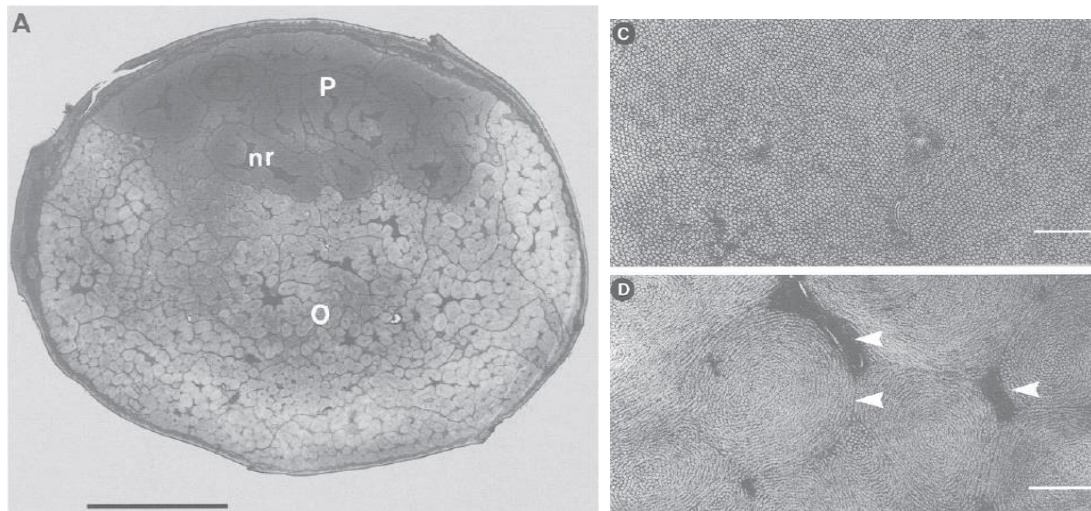


Figura 5. Micrografía electrónica del córtex. A. Se observa un corte transversal de un fibra teñida con un tratamiento de OsO_4 , uranilo acetato e hidróxido de plomo. Se observa la segmentación del córtex, O indica las células ortocorticales, P indica las células paracorticales y nr indica remanentes del núcleo. C Acercamiento de células paracorticales donde se observan las macrofibrillas con empaquetamiento hexagonal. D Acercamiento de las células ortocorticales se observa el empaquetamiento de las microfibrilla en forma de espiral, mayor empaquetamiento por ausencia de matriz entre los filamentos intermedios (Imagen reproducida de Rogers *et al.*, 2006).

1.2.2.3.3 Médula

La médula puede o no estar presente en el folículo, encontrándose principalmente en folículos primarios. En el caso esté presente, se ubica en la parte central de la fibra (Rogers, 2004). La proteína estructural más importante de esta estructura es la trichialina (Powell y Rogers, 1997).

1.2.3 Queratinas

Las queratinas son una familia de proteínas estructurales presentes en un gran número de vertebrados. Existen dos grupos grandes de queratinas, las alfa- queratinas y las beta-queratinas. Las primeras están presentes solo en mamíferos mientras que las segundas se expresan en aves y reptiles (Vandergh y Bossuyt, 2012). Las alfa queratinas (queratinas para el resto del texto) pertenecen a la superfamilia de proteínas capaces de formar filamentos intermedios (IF), diferenciándose del resto en que la unidad fundamental del IF es un heterodimero (Powell y Rogers, 1997). Por otro lado entre las proteínas capaces de formar IFs, las queratinas son el grupo más diverso (Powell y Rogers, 1997).

Se pueden clasificar en dos grupos, las del tipo I y las del tipo II. Las del tipo I son más pequeñas y de carácter ácido mientras que las del tipo 2 son más grandes y de carácter básico y neutro (Rogers *et al.*, 2006). Cada heterodimero consta de una queratina del tipo I y una del tipo II. Además de esta clasificación, estas proteínas pueden agruparse en queratinas suaves y duras. Las queratinas suaves son aquellas que se expresan en células epiteliales mientras que las queratinas duras son aquellas que en mamíferos forman uñas, pelos, cuernos etc (Fraser *et al.*, 1972). Hasta el momento se han identificado 54 queratinas distintas, de estas 11 del tipo I y 6 del tipo II se ven expresadas en el pelo del humano (Schweizer *et al.*, 2006).

Las queratinas se encuentran en un rango de tamaño de 40 – 70 kDa. Las del tipo I expresadas en pelo humano tienen un tamaño aproximado entre 40 – 48 KDa, mientras que las del tipo II se encuentran entre 59 – 62 kDa en humanos (Yu *et al.*, 1993). La estructura de la proteína consta de 4 dominios α - hélice separados por 4 dominios no helicoidales. Las regiones N terminales y C terminales son pequeñas, no helicoidales y contienen un segmento final, un segmento altamente conservado y uno altamente variable. En el caso de las queratinas suaves el segmento conservado de la región C

terminal está ausente (Figura 6) (Powell y Rogers, 1997). Según la última nomenclatura propuesta por Schweizer y colaboradores, las queratinas del tipo I expresadas en pelo humano son K31 – K40 y las del tipo II son K81 – k86 (Schweizer *et al.*, 2006).

En cuanto a los genes que codifican a estas proteínas, se ha visto que en diferentes especies se encuentran agrupados en dos *clusters* dentro del genoma. Estos dos *clusters* se han mantenido conservados en diferentes especies (Zimek y Weber, 2005; Hesse *et al.*, 2003). Los genes que codifican para queratinas tipo I tienen 7 intrones y un tamaño entre 4 y 5 Kb, mientras que los genes que codifican a las del tipo II poseen 8 intrones y un tamaño entre 7 y 9 kb (Powell y Rogers, 1997) En el humano todas las queratinas del tipo I, menos K18 se encuentran en el cromosoma 17q21 en una región de 977kb. Todos los genes muestran una misma orientación; sin embargo, dentro del *cluster* se encuentran separados por una porción que codifica proteínas asociadas a queratinas. Las queratinas del tipo II se encuentran en el humano en el cromosoma 12q13 en una región de 783 kb y tienen una misma orientación (Langbein *et al.*, 1999; Langbein *et al* 2001).

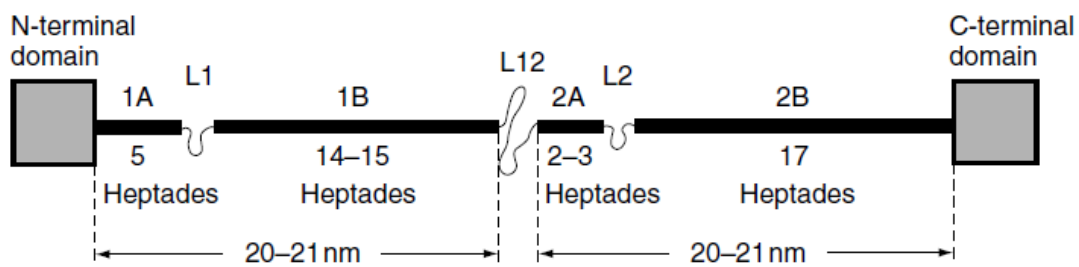


Figura 6. Se observa la estructura de una alfa queratina. Los segmentos helicoidales son 1A, 1B, 2A, 2B mientras que las regiones linker son L1, L12, L2. Las Heptades hacen referencia a los 7 residuos que conforman una vuelta en la hélice. (Imagen reproducida de Höcker, 2002)

1.2.3.1 Formación de filamentos intermedios

Para la formación de los filamentos intermedios es necesario cantidades equimolares de queratinas del tipo I y del tipo II. Una vez expresadas, una queratina de cada tipo se agrupa para formar un heterodímero. Los monómeros se juntan de manera antiparalela y se mantienen unidos por efecto hidrofóbico, además son estabilizados por puentes salinos entre residuos específicos presentes en los monómeros. El dímero formado adopta una forma de superhélice con giro hacia la izquierda (Höcker, 2002). Dos heterodímeros se empaquetan para formar un protofilamento, luego dos protofilamentos se unen para formar una protofibrilla y finalmente 4 protofibrillas constituyen un filamento intermedio que posee un diámetro de aproximadamente 7 nm (Wang *et al.*, 2015) (Figura 7). Una microfibrilla está conformada por 500-800 IFs; y, en una célula cortical se pueden encontrar aproximadamente entre 5 a 20 microfibrillas (Höcker, 2002). El mecanismo para el ensamblado de todas estas estructuras aún no se conoce pero la presencia de proteínas conocidas como proteínas asociadas a queratinas (KAPs) podría jugar un papel importante en este proceso (Fujimoto *et al.*, 2014).

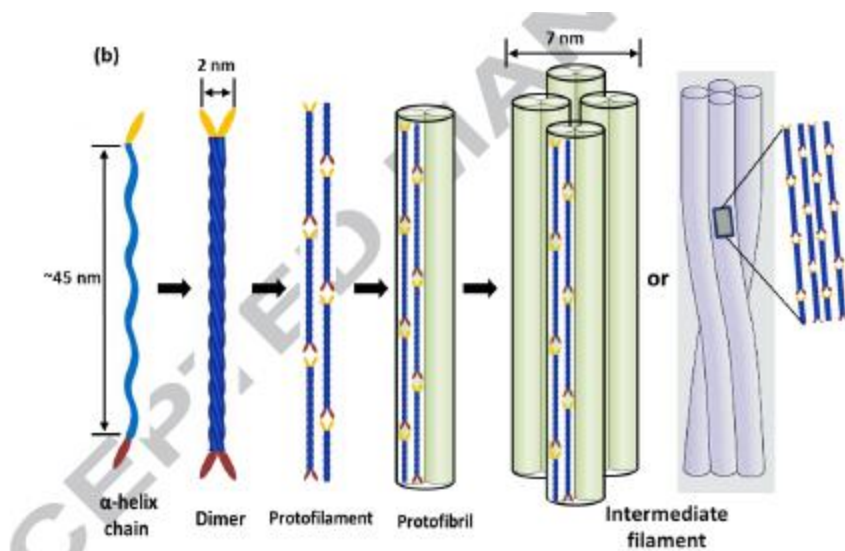


Figura 7. Estructuras intermedias previas a la formación de un filamento intermedio. De izquierda a derecha se observa un monómero de queratina, un heterodímero, un protofilamento, una protofibrilla y un filamento intermedio. (Imagen reproducida de Wang *et al.*, 2015)

1.2.4 Proteínas asociadas a queratina (KAPs)

Las proteínas asociadas a queratina (KAPs) son proteínas presentes en el pelo de todos los mamíferos. Estas constituyen la matriz que rodea a los IFs de queratina en las células corticales y de la cutícula (Rogers *et al.*, 2006). Esta familia fue por primera vez descrita al aislar proteínas del folículo de oveja (Gillespie y Board 1972; Gillespie, 1972). Desde entonces se han caracterizado en varias especies, identificando alrededor de 100 proteínas distintas (Khan *et al.*, 2014).

Las KAPs se dividen en dos grupos principales, las que presentan un alto contenido de sulfuro y las que presentan un alto contenido en glicina y tirosina (HGT). El primer grupo puede separarse en dos subgrupos, las proteínas con alto contenido de sulfuro (HS) con menos de 30% de residuos de cisteína y las de muy alto contenido de sulfuro (UHS) que presentan más del 30% de residuos de cisteína (Gillespie y Broad 1972; Gillespie, 1972). De acuerdo a la nueva nomenclatura presentada por Powel y Rogers, estas pueden ser agrupadas también en familias, y hasta el momento se han identificado 30 familias

diferentes (Tabla 3). La asignación de un miembro a una determinada subfamilia se realiza por homología de secuencias (Rogers *et al.*, 2007; Wu *et al.*, 2008).

Tabla 3. Asignación de familias a los tres grupos principales de KAPs.

Clasificación en base a %de residuos	Clasificación en familias
HS	KAP1 - 3, KAP10 - 16, KAP23 - 27
UHS	KAP4, KAP5, KAP9, KAP17
HGT	KAP6 - 8, KAP18 - 22

Estas proteínas son pequeñas, su tamaño varía aproximadamente entre 10 - 30 kDa. En cuanto a la estructura primaria en el caso de HS/UHS se ha visto que las secuencias presentan ciertas combinaciones de aminoácidos que se observan en repeticiones pentapeptídicas (Parry *et al.* 2006). En el caso de HGT estas repeticiones están ausentes (Rogers *et al.*, 2006). Falta realizar estudios en cuanto a su estructura tridimensional.

Aún no se conoce con exactitud su función; sin embargo se ha determinado que estas proteínas se unen a los IF de queratina (Figura 8). Estas uniones son a través de puentes disulfuro, pero también se dan mediante puentes de hidrógeno (Fujimoto *et al.*, 2014; Rogers *et al.*, 2006). Por otro lado se ha visto que la unión con las proteínas de queratina es de manera selectiva y que generalmente se da en el extremo N- terminal de las queratinas, el cual es muy conservado en estas proteínas (Fujimoto *et al.*, 2014). También se ha observado que la presencia de estas pueden ser importantes para un correcto ensamblaje de los IF (Fujimoto *et al.* 2014). Por último, hay reportes de una regulación estricta a nivel post-transcripcional (Fujimoto *et al.*, 2014). Estas proteínas se expresan en gran cantidad en la cutícula y en el córtex, sin embargo hay reportes de una baja expresión en la matriz del folículo (Rogers *et al.*, 2006).

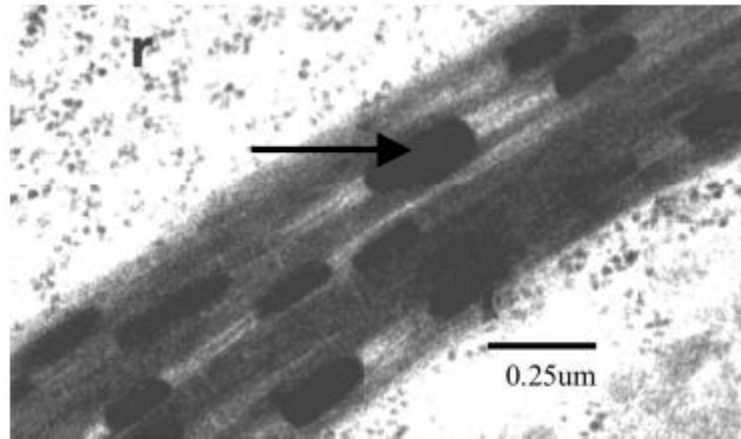


Figura 8. En la figura se puede observar una macrofibrilla, dentro de esta se aprecian agregados de proteínas que corresponden a las KAPs (flecha negra). (Imagen reproducida de Rogers, 2004)

En cuanto a los genes que codifican estas proteínas, existe evidencia que en varias especies se encuentran organizados en *clusters* (Shibuya *et al.*, 2004a; Khan *et al.*, 2014). En el humano se han identificado 5 *clusters* distintos, las UHS/HS están presentes en los 5. Estos se ubican en los cromosomas 17q21.2, 21q22.1, 21q22.3, 11p15.5, 11q13. Mientras que las proteínas HGT se encuentran en el cromosoma 21q22.1, rodeado de dos grupos de genes de proteínas UHS/HS (Zhumbaeva *et al.*, 1992; Rogers *et al.*, 2001; Rogers *et al.*, 2002; Rogers *et al.*, 2004; Shibuya *et al.*, 2004). La nomenclatura para estos genes fue propuesta por Powell y Rogers; los genes tienen el prefijo KRTAP seguido del número de familia (Rogers y Powell 1993). Los genes en promedio son de 1Kb y constan de 1 solo exón (Rogers *et al.*, 2006). La expresión de estos genes dentro del folículo se da en etapas avanzadas de la diferenciación. Estos genes empiezan a expresarse de manera coordinada luego de que las células entran en la fase de queratinización (Rogers *et al.*, 2006).

Se han observado diferencias importantes en los repertorios de KRTAPs que presentan las distintas especies de mamíferos, sobre todo en aquellas donde el pelo ha sufrido ciertas

modificaciones adaptativas. Además la presencia de pseudogenes también varía de acuerdo a la especie; esta gran variabilidad se originó por eventos de duplicación (Khan *et al.*, 2014). En cuanto a los patrones evolutivos, se ha determinado que el grupo de alto contenido de sulfuro ha pasado por una evolución concertada. En el caso de las HGT se han encontrado menos casos de conversión génica y recombinación (Khan *et al.*, 2014). Además del número de genes, las secuencias de estos presentan una gran variabilidad, habiéndose identificado en humano, cabra y oveja polimorfismos por mutaciones puntuales como también inserciones/deleciones de hasta 30 nucleótidos en región codante. Además, varios de estos polimorfismos han sido asociados a distintas características de la fibra con valor comercial (oveja y cabra), como por ejemplo el diámetro (Rogers y Schweizer, 2005; Gong *et al.*, 2011a, Gong *et al.*, 2011b; Itenge – Mweza *et al.*, 2007; Shimomura *et al.*, 2002; Zhou *et al.* 2015; Mckenzie *et al.*, 2010; Zhang *et al.*, 2011; Gong *et al.*, 2015; Yu *et al.*, 2008).

1.2.5 Características de la fibra de alpaca

1.2.5.1 Aspectos morfológicos de la fibra de alpaca

1.2.5.1.1 Estructura folicular

No hay muchos estudios en cuanto a las características del folículo en alpaca. Se ha determinado que la alpaca huacaya presenta agrupaciones foliculares que constan de un folículo primario rodeado por 4 a 20 folículos secundarios (Figura 9) (Antonini *et al.*, 2004; Badajoz, 2009; Ferguson *et al.*, 2012; Moore *et al.*, 2015). En el caso de las suri se han visto agrupaciones similares a las observadas en oveja Merino, donde la agrupación folicular consta de tres folículos primarios rodeados de varios folículos secundarios (Badajoz, 2009; Moore *et al.*, 2015). Por otro lado se ha determinado que la maduración

completa de los folículos secundarios en alpaca se da a los 4 meses después del nacimiento (Antonini *et al.*, 2004).

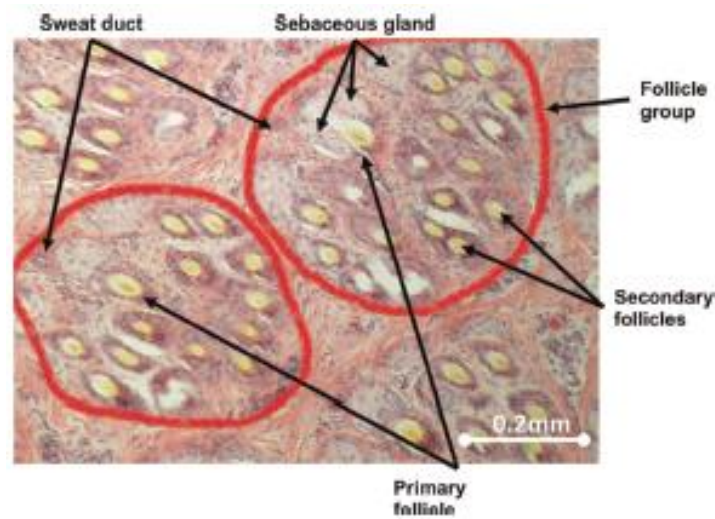


Figura 9. Micrografía de una sección de piel de alpaca huacaya. Se observan dos paquetes foliculares, cada uno presenta un folículo primario asociado a glándulas sudoríparas y varios folículos secundarios. (Imagen reproducida de Moore *et al.*, 2015)

1.2.5.1.2 Cutícula

En cuanto a la estructura de la cutícula en alpaca se ha determinado que consta de 3 a 4 capas celulares (Thomas *et al.*, 2012). Las células presentan un patrón de mosaico irregular similar al observado en oveja (Figura 10) (Thomas *et al.*, 2012). En cuanto al número de células cuticulares, se encontró un promedio de 9.11 células por 100 μm de fibra para alpaca huacaya y 7.57 células por 100 μm de fibra en alpaca suri (Valbonesi *et al.*, 2010). La altura de las células cuticulares fue similar para ambas variedades de alpaca, siendo de 0.5 μm (Valbonesi *et al.*, 2010).

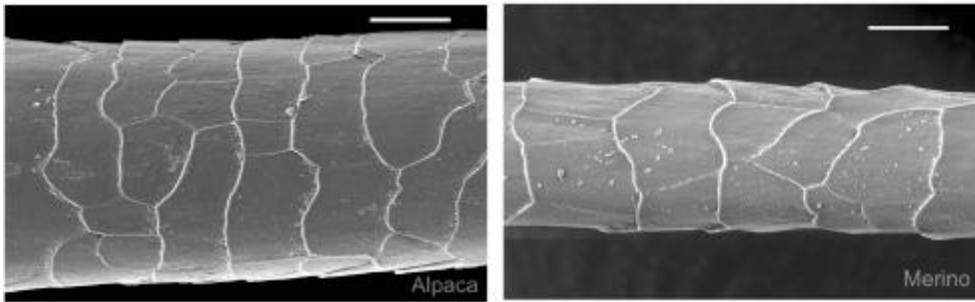


Figura 10. Micrografía de la topografía de fibras finas. A la izquierda se observa el patrón de las células cuticulares en alpaca huacaya y a la derecha el patrón de las células cuticulares de oveja merino. (Imagen reproducida de Thomas *et al.*, 2012)

1.2.5.1.3 Córtez

El córtex de la alpaca presenta la típica estructura bilateral observada en fibra de oveja. Las células ortocorticales presentan las macrofibrillas en forma de anillos (Figura 11) (Thomas *et al.*, 2012). Las células paracorticales presentan los filamentos intermedios empaquetados de manera pseudo-hexagonal (Figura 11) (Thomas *et al.*, 2012). Además de estos dos tipos celulares se encontró un tercer tipo, que presenta un grado intermedio de empaquetamiento de las macrofibrillas (Figura 11) (Thomas *et al.*, 2012).

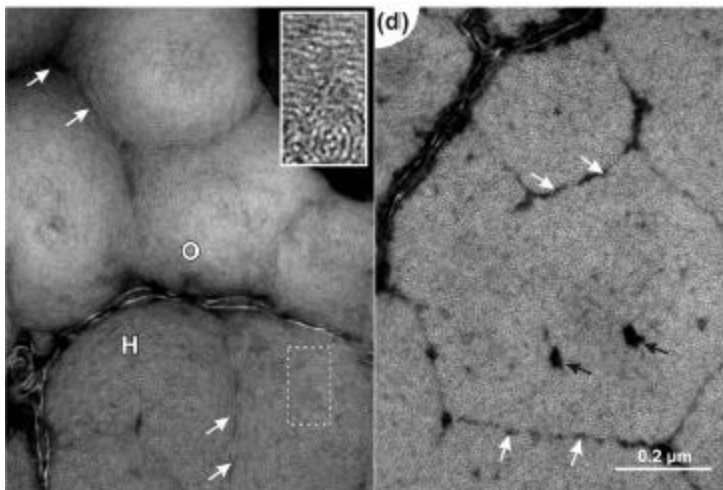


Figura 11. Micrografía electrónica del córtex de una fibra de alpaca. A la izquierda se observan las células ortocorticales (O) con el empaquetamiento de las macrofibrillas en forma de anillos. También se observan el grupo celular que presenta un empaquetamiento intermedio (H). A la derecha se observan las células paracorticales con el empaquetamiento de las macrofibrillas pseudo-hexagonal. (Imagen reproducida de Thomas *et al.*, 2012).

1.2.5.2 Aspectos bioquímicos de la fibra de alpaca

Existen poco reportes de los genes expresados en fibra de alpaca. De una biblioteca de ESTs generada por Fan y colaboradores se identificaron 9 genes de 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) (Fernandez, 2015). Por otro lado a partir de una biblioteca de cDNA generada en la Unidad de Biotecnología Molecular (UBM) de la Universidad Peruana Cayetano Heredia se identificaron transcritos para los genes KRTAP3-3, KRTAP9-2, KRTAP17-1, KRTAP13-1, KRTAP7-1, KRTAP11-1, KRTAP1-2, KRTAP4-8, KRTAP8-2, KRT33, KRT34, KRT81, KRT85, KRT86. Se ha determinado también que el miRNA Let-7b promueve el crecimiento de fibra de alpaca mediante un represión transcripcional de TGF β R1 (Yan *et al.*, 2016).

1.2.5.3 Características con valor comercial de la fibra de alpaca

1.2.5.3.1 Diámetro de fibra

El diámetro de la fibra en alpaca es el índice de calidad más importante. De acuerdo al diámetro se han fijado distintas categorías de calidad (Tabla 4). En general mientras más delgada sea la fibra más fina se considera. Se ha demostrado que al igual que en ovinos la zona del costillar medio resulta representativa para la evaluación de la finura media del animal (Aylan-Parker y McGregor, 2002). Además del diámetro de fibra, se debe calcular el coeficiente de variación, el cual indica en porcentaje cuanto varían los diámetros de fibra para un mismo animal. Se espera que las alpacas tengan un bajo coeficiente de variación, de esta forma la fibra que producen es más homogénea en sus características. A partir del diámetro de fibra también se realiza el cálculo del índice de confort. Este

índice mide el grado de confort de la fibra, donde se espera que no más del 5% de las fibras tengan un diámetro mayor a 30 μm (Quispe *et al.*, 2013).

Tabla 4. Clasificación de la calidad de fibra de alpaca en el mercado. Según Norma Técnica Peruana (NTP 231.301:2004)

Categoría	Diámetro (μm)	longitud (cm)
Alpaca Baby	Hasta 23	65
Alpaca Fleece	23.2 a 26.5	70
Alpaca Medium Fleece	26.6 a 29	70
Alpaca Huarizo	29.1 a 31.5	70
Alpaca Gruesa	Mas de 31.5	70

Los registros de diámetro de fibra varían de acuerdo a las zonas donde se realizaron los estudios. En Perú los valores se encuentran entre 20 μm a 27 μm , mientras que en otros países productores de fibra estos oscilan entre 25 μm y 27 μm (Tabla 5). En cuanto al coeficiente de variación en alpacas peruanas este valor varía entre 18 % a 36 %, mientras que en otros países se encuentra entre 17 % y 24 % (Tabla 5). Se espera un valor de coeficiente de variación para el diámetro de fibra de 24% que representa el límite para rendimientos textiles. Por lo tanto los valores observados en alpacas peruanas se consideran un tanto elevados.

Tabla 5. Reportes de diámetro de fibra y coeficiente de variación para alpaca huacaya de distintas regiones del Perú y países en el mundo.

Region	Numero de alpacas	Diámetro (μm)	Coeficiente de variacion (%)	Fuente
Iscahuaca - Apurímac	405	20	21.2	Vasquez <i>et al.</i> , 2015
Pacomarca - Puno	2405	23	23.31	Cervantes <i>et al.</i> , 2010
Pacomarca - Puno	4718	22.82	23.12	Morante <i>et al.</i> , 2009
Lachoc - Huancavelica	185	23.42	18	Cordero <i>et al.</i> , 2011
Tocra - Arequipa	60	27.41	36.65	Valbonesi <i>et al.</i> , 2010
Parinacot - Chile	77	22.57	17.37	Crossley <i>et al.</i> , 2016
Estados Unidos	585	27	23.48	Lupton <i>et al.</i> , 2006
Australia	no definido	25.7	24.1	Ponzoni <i>et al.</i> , 1999

Existen diversos reportes para la heredabilidad en sentido estricto del diámetro de fibra. La heredabilidad mide la proporción de la varianza que se le atribuye al componente genético, este índice será explicado en mayor detalle más adelante en el texto. Estos reportes presentan valores muy diferentes que oscilan entre 0.18 y 0.73 (Tabla 6). En algunos casos estos valores sugieren una fuerte influencia genética sobre el diámetro de fibra.

Tabla 6. Cálculos de heredabilidad en sentido estricto (h^2) para el diámetro de fibra de diferentes regiones del Perú y países del mundo

Region	Heredabilidad	Referencia
Pacomarca - Puno	0.32	Cruz et al., 2015
Tocra - Arequipa	0.36 ± 0.20	Paredes y Peralta, 2011
San Simon - La Libertad	0.54	Cruz, 2011
no especificado	0.18	Leon Velarde and Guerrero 2001
Australia	0.67 ± 0.30	Ponzoni et al., 1999
Nueva Zelanda	0.73 ± 0.19	Wuliji e al., 2000

1.2.5.3.2 Peso del vellón

El peso del vellón también es una característica muy importante porque indica la cantidad de fibra que produce el animal anualmente. Un vellón más pesado vale más que uno más ligero. En Nueva Zelanda se han reportado datos de peso de vellón de 2.2 Kg (Wuliji, *et al.*, 2000), mientras que en Australia se han reportado entre 2 Kg a 3.3 Kg (McGregor y Butler, 2004). En el Perú en la comunidades campesinas el peso de vellón es bajo, se calcula una producción bianual de 2.3 Kg. Esta producción baja se debe en parte al déficit en tecnología que se utiliza para las esquilas (Quispe *et al.*, 2013).

La heredabilidad para esta característica también ha sido estudiada. Considerando solo la primera esquila en poblaciones de Perú se han determinado heredabilidades de 0.38 ± 0.34 (Mamani, 1991) y de 0.31 ± 0.17 (Ruiz de Castilla *et al.*, 1992). Teniendo en

consideración todas las edades los valores de heredabilidad reportados son 0.79 ± 0.36 (Ponzoni *et al.*, 1999) y 0.73 ± 0.19 .

1.2.5.4 Factores que afectan las características de valor comercial

Se han identificado varios factores que afectan al diámetro de fibra. Uno de estos factores es la densidad folicular y el ratio folículos secundarios: folículos primarios (S/P). En un estudio en una población de alpacas en Chile se encontró una correlación negativa ($r = -0.47$, $P < 0.05$) en cuanto a la densidad folicular y el diámetro en alpacas huacaya blancas (Crossley *et al.*, 2016). Moore y colaboradores también encontraron una asociación entre la densidad folicular y el diámetro de fibra como también entre el ratio S/P y el diámetro de fibra en alpacas. A mayor densidad folicular menor es el diámetro de fibra y mientras más alto es el ratio S/P menor es el diámetro de fibra (Moore *et al.*, 2015). En cuanto a la cantidad de vellón no se han encontrado estudios que encuentren asociación entre el peso y la densidad folicular en alpacas, sin embargo se espera que a mayor número de folículos, mayor peso del vellón.

Otro factor importante que afecta el diámetro de fibra es la edad. Se ha observado que conforme envejecen las alpacas, estas producen una fibra con un diámetro mayor (Quispe *et al.*, 2013). Lupton y colaboradores calcularon valores de diámetro de fibra de $24.3 \mu\text{m}$, $26.5 \mu\text{m}$ y $30.1 \mu\text{m}$ para alpacas de 1, 2 y 3 años o más respectivamente (Lupton *et al.*, 2006). Varios autores proponen que esta correlación con la edad podría deberse a que las esquilas tienden a aumentar el funcionamiento folicular (Roger, 2006). En cuanto al peso del vellón se ha observado que a mayor edad mayor es el peso del vellón. Esto se debe a que alpacas mayores presentan una mayor superficie corporal (Frank *et al.*, 2006).

El sexo es un factor donde existen ciertas discrepancias con respecto a su relación con el diámetro de fibra. Existen reportes que las alpacas macho producen fibra más fina, sin

embargo esto podría deberse a que estos son sometidos a procesos de selección más minuciosos (Morante *et al.*, 2009; Quispe *et al.*, 2009). Por otro lado existen reportes que las hembras producen fibra más delgada (Aylan – Parker y McGregor, 2002; Lupton *et al.*, 2006), esto podría estar asociado a una priorización del uso de aminoácidos para producción de leche (Adams y Cronje, 2003). También hay autores que indican que el sexo no tiene efecto sobre el diámetro de fibra (Wuliji *et al.*, 2000; McGregor and Butler *et al.*, 2004). En cuanto al peso del vellón el efecto del sexo es más claro, encontrándose que los machos producen un vellón más pesado. Esto se debe a que los machos tienen un tamaño mayor (Lupton *et al.*, 2006; Wujili *et al.*, 2000)

La alimentación también tiene un efecto importante sobre el diámetro de la fibra. Se han llevado a cabo estudios donde se suplementaron alpacas con dietas con un valor nutricional bajo y lo que se observó fue una disminución en el diámetro de la fibra (Franco *et al.*, 2009). Esta relación se ha observado también en ovinos. Por otro lado relacionado a la nutrición, se ha observado diferencias en el diámetro de la fibra de acuerdo a la locación, esto podría estar relacionado con el tipo de pasto de las distintas zonas (Quispe *et al.*, 2013). Con respecto al peso de vellón, se observó que el peso del vellón disminuye con una dieta de valor nutricional bajo (Franco *et al.*, 2009)

La sarna es considerada una de las enfermedades más importantes en alpaca (FAO, 2005). Sin embargo no se ha realizado ningún estudio sobre el efecto en el diámetro de fibra, pero existen evidencias que demuestran una tendencia a disminuir el diámetro (Quispe *et al.*, 2013). Sobre el peso de vellón la carga parasitaria tiene un efecto negativo, disminuyendo la producción (Quispe *et al.*, 2013).

1.2.6 Rasgos cuantitativos y mapeo de QTLs

Dentro de una población se observan individuos con diferencias en varios rasgos a nivel fenotípico. Estas diferencias pueden ser cualitativas o cuantitativas. En el caso de características cualitativas se observan fenotipos fáciles de distinguir y con una distribución discreta, por ejemplo el grupo sanguíneo en el humano. Estas características por lo general son controladas por un único o pocos genes. En el caso de las características cuantitativas los fenotipos no son fáciles de distinguir, generalmente tienen que medirse y presentan una distribución continua (Falconer y Mackay, 1996). Este tipo de características son muy importantes porque implican a la gran mayoría de rasgos como por ejemplo enfermedades complejas, conductas y rasgos de valor comercial para animales o plantas (Hamilton, 2009). Son varios genes los que gobiernan este tipo de características. Además del componente genético, estos rasgos son influenciados por el ambiente. Esto quiere decir que dos individuos con un mismo genotipo pueden presentar valores fenotípicos diferentes o que dos individuos con genotipos diferentes pueden presentar valores fenotípicos iguales (Falconer y Mackay, 1996).

La base genética de los rasgos cuantitativos es que los genes que los afectan poseen alelos que suman o restan a la magnitud del carácter. Estos genes son conocidos como QTLs por sus siglas en inglés (*quantitative trait loci*) (Falconer y Mackay, 1996). Por lo general un rasgo cuantitativo se caracteriza por estar influenciado por un gran número de genes con efectos pequeños (Mackay *et al.*, 2009). Además estos genes pueden presentar efectos dependientes del contexto, lo cual quiere decir que un mismo genotipo puede presentar efectos diferentes en individuos con diferentes edades, pesos u otras características (Mackay *et al.*, 2009). Por otro lado pleiotropía es otro fenómeno común entre estos genes, donde un gen influencia a más de un carácter (Mackay *et al.*, 2009).

Al trabajar con rasgos cuantitativos siempre es importante determinar el porcentaje de variación que se debe al efecto de los QTLs y al efecto ambiental. Para lograr esto la variación fenotípica se puede descomponer en las distintas fuentes de variación (genética y ambiental) (Figura 12). Como se observa la varianza genética se descompone en 3 distintas varianzas (aditiva, dominancia e interacción) (Falconer y Mackay, 1996). La aditiva corresponde a la variación debido a la suma o resta de los efectos de los diferentes alelos sobre la magnitud del rasgo. La varianza por dominancia hace referencia a la varianza debido a relaciones de dominancia entre los genes. Por último la varianza por interacción se refiere a la variación producida por epístasis (Figura 12) (Falconer and Mackay, 1996). Un parámetro importante en el estudio de estos rasgos es la heredabilidad, Existen dos tipos, heredabilidad en sentido amplio y heredabilidad en sentido estricto (Figura 12). La heredabilidad en sentido amplio es el porcentaje de varianza que se explica debido a factores genéticos en general. La heredabilidad en sentido estricto es la proporción de variabilidad explicada por la varianza genética aditiva. Esta última es muy importante para la selección de caracteres ya que la varianza aditiva es el componente de la varianza genética que explica el parecido entre parientes (Falconer y Mackay, 1996).

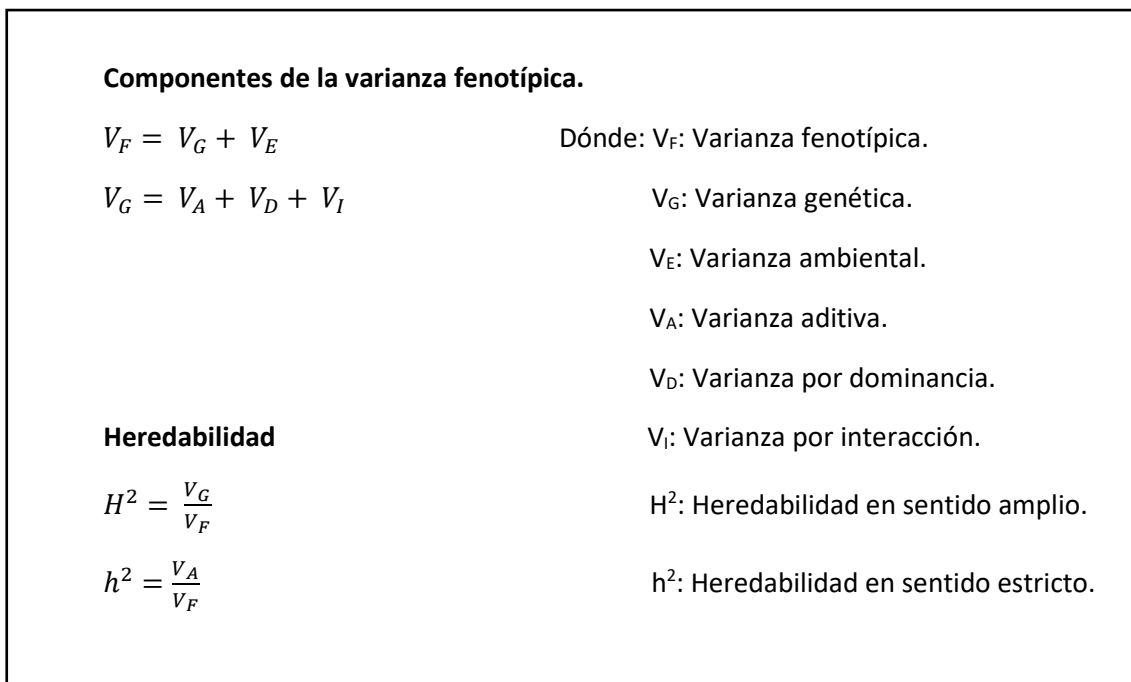


Figura 12. Descomposición de la varianza fenotípica y determinación de la heredabilidad en sentido amplio y estricto.

Otro aspecto importante del estudio de rasgos cuantitativos es el mapeo de QTLs. Este consiste en la premisa que los QTLs pueden ser localizados por el ligamiento genético con otros marcadores ya conocidos. Existen dos aproximaciones para el mapeo de QTLs, el de un test de ligamiento o un test de asociación (Mackay *et al*, 2009). El test de ligamiento consiste en el uso de una muestra con relaciones de parentesco conocidas, donde se busca un marcador que este próximo al loci causante. Un ejemplo de análisis por ligamiento es la prueba de *Quantitative Trait Linkage Disequilibrium Test* (QTDT) (Havil *et al.*, 2005). El test de asociación se basa en los mismos supuestos, sin embargo la muestra consta de individuos no relacionados y los eventos de recombinación que se buscan son eventos históricos ocurridos en la población (Balding, 2006). En ambos casos el fin es localizar una región en el genoma donde se encuentre o este próximo el gen o la mutación responsable del fenotipo observado.

En el caso de un análisis por asociación existen diversas aproximaciones que se pueden optar. Una muy utilizada hoy en día es el *Genome Wide Association Study* (GWAS), que consiste en un estudio de asociación analizando todo o gran parte del genoma utilizando un número grande de marcadores genéticos (Foulkes, 2009). Otra aproximación es la de gen candidato; en la que se seleccionan, previo al estudio de asociación, genes candidatos a estar relacionados a la característica de interés. Para la selección de estos genes se utilizan reportes previos sobre las funciones de estos genes. Una vez identificados los genes, se deben identificar marcadores genéticos dentro o cerca a estos (Foulkes, 2009). La evaluación de la asociación se puede realizar mediante varios métodos, siendo uno de los más comunes la regresión lineal. Sin embargo el método varía de acuerdo a la muestra (Foulkes, 2009).

1.2.7 Marcadores genéticos moleculares

Los marcadores genéticos moleculares son polimorfismos únicos de una secuencia de ADN que pueden ser usados para identificar la contribución del gen en determinado fenotipo, para identificar un gen responsable de un rasgo, localizar un gen en una región cromosómica, analizándolo en individuos, familias o poblaciones. En general, los marcadores genéticos moleculares pueden ser encontrados en cualquier región del genoma, ya sea nuclear o mitocondrial, deben ser polimórficos y el alelo menos representado debe tener una frecuencia de al menos 5%. Existen diversos marcadores moleculares de acuerdo a la forma en que se generan; además, los criterios de clasificación son muy variados (Van Tassel *et al.*, 2008).

Los marcadores pueden ser clasificados de acuerdo a su capacidad de poder diferenciar entre individuos heterocigotos y homocigotos. Se llaman co-dominantes aquellos que permiten esta diferenciación mientras que los que no, se denominan dominantes (Sunnucks, 2000). Otra clasificación va de acuerdo al tipo de secuencia que analizan, en

este caso encontramos del tipo 1 y del tipo 2. Los marcadores de tipo 1 se ubican en secuencias codantes mientras que los del tipo 2 en secuencias no codantes (Sunnucks, 2000). Por último, el tipo de herencia también es otro criterio de clasificación; en este caso, aquellos marcadores que son heredados únicamente por uno de los padres se clasifican como marcadores de herencia uniparental, mientras que aquellos que son heredados de ambos padres son de herencia biparental (Galtier *et al.*, 2009).

Los marcadores moleculares son muy usados en varios campos de la ciencia como por ejemplo genética de poblaciones, biología forense, medicina entre otros (Feral *et al.*, 2002).

1.2.7.1 Polimorfismos de nucleótido simple (SNP)

Este marcador consiste en una mutación puntual dentro de una secuencia única de ADN. Sin embargo, para que esta mutación pueda ser considerada como un marcador todas las variantes deben estar presentes a una frecuencia mayor al 5%. Es un marcador co-dominante que puede ser encontrado en cualquier parte del genoma (Van Tassel *et al.*, 2008; Vignal *et al.*, 2002). Es, por lo general, un marcador bialelico a pesar de que en una posición determinada de una secuencia se puedan presentar los 4 nucleótidos. Una de las razones para esto es la baja frecuencia con la que se da una sustitución nucleotídica. Se ha estimado que en el caso de una posición neutral en un mamífero, la frecuencia está entre 1×10^{-9} y 5×10^{-9} por nucleótido por año. Esto quiere decir que la probabilidad de que se den dos cambios independientes en una sola posición es muy baja (1×10^{-18} y 5×10^{-18}) (Martinez – Arias *et al.*, 2001). Otra razón es debido a un sesgo al momento de ocurrir la mutación. Existen dos tipos de sustituciones nucleotídicas; una son las transiciones, donde los cambios pueden ser: purina- purina (A<->G) o pirimidina – pirimidina (C<->T). El otro tipo se conoce como transversión, donde los cambios son: purina – pirimidina o pirimidina – purina (A<->T, A<->C, G<->T, G<->C) (Nelson y

Cox, 2009). Debido a que hay el doble de transversiones posibles con respecto a transiciones, uno esperaría que la tasa de transición/transversión sea menor a 0.5 en caso las mutaciones sean al azar. Sin embargo en la práctica no se observa esto. Un estudio que compara secuencias de ratones y humanos ha encontrado una tasa de transición/transversión de 1.4 (Colins y Jukes, 1994). Otro estudio de generación de SNPs en humanos a partir de EST encontró una tasa de transición/transversión de 1.7 (Picoult-Newberg *et al.*, 1999). En aves esta tasa es mayor con un valor de 2.3 (Smith *et al.*, 2001).

Existen varias estrategias para el descubrimiento de los SNPs. Las principales consisten en la comparación de secuencias específicas generadas en cromosomas diferentes. La forma más simple es la comparación de secuencias obtenidas a partir de productos de PCR. Sin embargo, esta estrategia tiene sus desventajas debido a que se necesita tener conocimientos previos de la secuencia y produce secuencias diploides. Al ser diploides, dos picos superpuestos pueden ser un SNP pero también pueden ser artefactos de secuenciamiento (Figura 13) (Vignal *et al.*, 2002). Para evitar el problema de las secuencias diploides se puede utilizar otra estrategia que consiste en la comparación de secuencias obtenidas por clonación de fragmentos, de esta forma al encontrarse dos picos superpuestos no existe la incertidumbre y es considerado como un artefacto de secuenciamiento. Otra estrategia es la comparación de EST sin embargo esta estrategia produce falsos positivos (Vignal *et al.*, 2002).

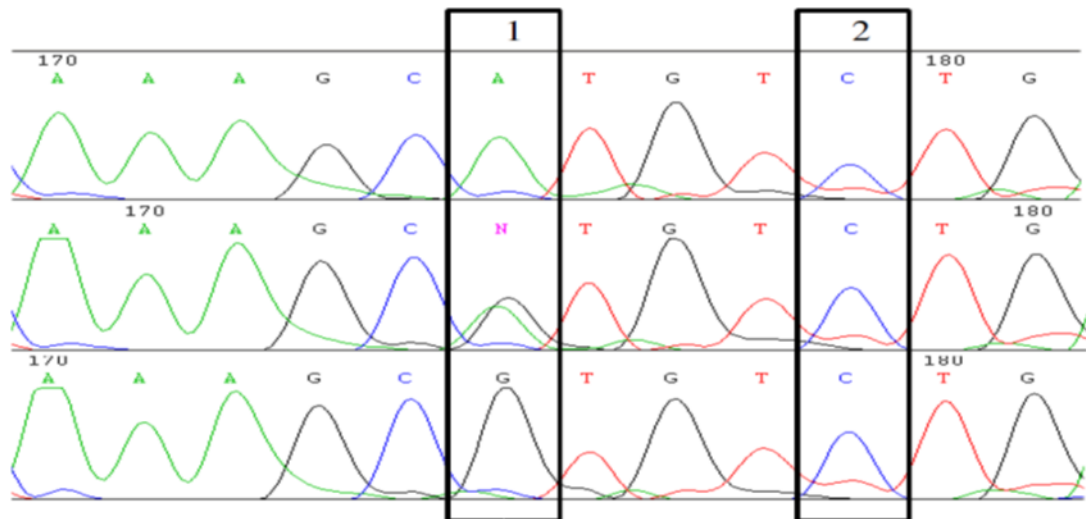


Figura 13. Identificación de SNPs por secuenciación directa de producto de PCR. En la caja 1 en la primera secuencia se observa un individuo homocigoto para el alelo A, en la segunda secuencia se observa un individuo heterocigoto AG y en la tercera secuencia se observa un individuo homocigoto para el alelo G. En la caja 2 se puede observar en las tres secuencias los errores de secuenciación a los que está sujeta esta técnica (Imagen reproducida de Vignal *et al* 2002).

1.2.8 Parámetros genéticos de los marcadores moleculares

1.2.8.1 Equilibrio de Hardy – Weinberg

El equilibrio de Hardy – Weinberg permite predecir frecuencias genotípicas a partir de frecuencias alélicas o viceversa. La relación es la siguiente (Hamilton, 2009):

Si las frecuencias alélicas para un gen con dos alelos son p y q, entonces la frecuencias genotípicas serán: p^2 para homocigotos para ese alelo, $2*p*q$ para heterocigotos y q^2 para homocigotos para ese alelo.

Es importante mencionar que para que una población se encuentre en equilibrio de Hardy-Weinberg esta debe ser de un organismo diploide, debe ser una población grande, no debe haber mutaciones, ni selección, ni migración. En caso se determine que esta relación no se cumple, se dice que la población no se encuentra en equilibrio de Hardy-Weinberg y probablemente se debe a que alguno de los supuestos no se cumple (Hamilton, 2009).

1.2.8.2 Heterocigosidad

La heterocigosidad (H) indica la proporción de individuos heterocigotos dentro de una población. El índice toma valores de 0 a 1, un valor alto que existe una alta diversidad alélica. Existen dos tipos de heterocigosidad: La heterocigosidad esperada (HE) y la heterocigosidad observada (Ho) (Shete *et al.*, 2009).

La heterocigosidad esperada (HE) se define como la probabilidad de que un individuo elegido al azar de una población sea heterocigoto para un locus. Asumiendo equilibrio de Hardy- Weinberg (Nei, 1973).

$$He = 1 - \sum_{i=1}^n p_i^2$$

Dónde: p_i es la frecuencia del alelo i y n es el número de alelos en un locus

La Heterocigosidad observada (Ho) indica la proporción de individuos heterocigotos observados para un locus en particular en una población (Nei, 1973).

$$Ho = \frac{\text{Individuos Heterocigotos}}{\text{Individuos totales}}$$

1.2.8.3 Contenido de información polimórfica (PIC)

El contenido de información polimórfica (PIC) indica que tan informativo es un marcador. Este está en función del número de alelos y de la frecuencia alélica (Shete *et al.*, 2009).

$$PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2 p_i^2 p_j^2$$

Dónde: p_i es la frecuencia de alelo I, p_j es la frecuencia del alelo J y n es el número de alelos en un locus.

1.2.8.4 Desequilibrio de ligamiento (LD)

Este parámetro indica el nivel de asociación entre alelos de diferentes loci. Se calcula a través de la determinante D (Hamilton, 2009).

$$D = (fg_{11})(fg_{22}) - (fg_{21})(fg_{12})$$

Donde: (fg_{11}) (fg_{22}) representa el producto de las frecuencias gaméticas en fase de acoplamiento y (fg_{12}) (fg_{21}) representa el producto de las frecuencias gaméticas en fase de repulsión.

Sin embargo si la fase gamética es desconocida se puede calcular el LD en base a un *Likelihood Ratio Test* (LRT). En este caso se asume primero que los marcadores se encuentran en equilibrio de ligamiento, en donde las frecuencias haplotípicas son iguales a la multiplicación de las frecuencias alélicas. Esto es comparado contra un modelo asumiendo desequilibrio de ligamiento, para lo que se calculan las frecuencias haplotípicas en base a la predicción de los haplotipos utilizando un algoritmo de máxima verosimilitud. La comparación asume una distribución de Chi- cuadrado (Excoffier y Slatkin, 1998).

II. Objetivos

2.1 Objetivo general

Identificación de marcadores del tipo SNP en genes que codifiquen para proteínas asociadas a queratinas (KRTAPs) en alpaca huacaya y evaluar la posible asociación de uno de estos genes al diámetro de fibra siguiendo una aproximación de gen candidato.

2.2 Objetivos específicos

- Identificar y caracterizar genes candidatos de la familia KRTAP en alpaca huacaya que podrían estar asociados a distintas características de valor comercial de fibra.
- Identificar variaciones en las secuencias de los genes candidatos seleccionados.
- Validar las variaciones identificadas en los genes candidatos como marcadores genéticos del tipo SNP.
- Seleccionar un gen candidato para evaluar la asociación entre este y el diámetro de fibra en alpaca huacaya.
- Evaluar si el gen candidato seleccionado está asociado al diámetro de fibra en alpaca huacaya.

III. Hipotesis

Hipotesis 1: Los genes KRTAP son buenos candidatos para la identificación de marcadores genéticos del tipo SNP.

Hipotesis 2: El gen KRTAP11-1 esta asociado al diámetro de fibra en alpaca.

IV. Componente I. Validación de marcadores en genes KRTAP en alpaca

4.1 Materiales y Métodos

4.1.1 Determinación de genes candidatos a estar relacionados a características de fibra

Entre las dos familias principales de proteínas relacionadas a características de fibra se seleccionó las proteínas asociadas a queratinas (KAPs); esta decisión se explicará en la sección de discusión. Para la selección de los genes se tuvieron en cuenta dos criterios importantes: Se buscaron genes que se estén expresando en fibra de alpaca y genes en donde se haya establecido alguna relación o asociación importante con alguna característica de fibra en otras especies. Para elegir genes que se estén expresando en fibra de alpaca se utilizó una biblioteca de cDNA generada a partir de piel de alpaca por la Unidad de Biotecnología Molecular (UBM). Se identificaron todos los genes KRTAPs utilizando la herramienta BLAST. Una vez identificados estos genes se realizó una búsqueda bibliográfica de genes KRTAPs en los que se hayan reportado polimorfismos (en otras especies productoras de fibra) y asociaciones a diferentes características de fibra. Teniendo en cuenta ambos criterios se seleccionaron 5 genes, uno de estos sin embargo no se encontraba en la biblioteca de cDNA.

4.1.2 Obtención de secuencias *in silico* a partir del genoma de alpaca (Vicugna-Pacos 2.0.1)

La obtención de las secuencias *in silico* de los genes seleccionados se hizo utilizando los fragmentos de los genes (Anexo 1) identificados utilizando la biblioteca de cDNA previamente descrita. Estas secuencias fueron comparadas con el genoma de referencia de alpaca (**Vicugna-Pacos 2.0.1**) en el Genbank utilizando la herramienta BLAST. De esta forma se ubicó una parte del gen. Cerca de esta región se buscaron marcos de lectura

abiertos (ORF) que pudieran codificar una proteína utilizando la herramienta ExPASy-translate tool (<http://web.expasy.org/translate/>). En el caso del gen que no se encontró en la biblioteca de cDNA (KRTAP6-1) se utilizó como secuencia de referencia, para la búsqueda en el genoma, el gen reportado en oveja (Acc. Nr. NP_001180328.1) (Anexo 1). Una vez identificada la posición de los ORFs para cada uno de los genes de interés se extrajeron 500pb *downstream* y *upstream* de esta región, para la identificación de secuencias regulatorias y el diseño de cebadores.

4.1.3 Caracterización in silico de genes candidatos KRTAP en alpaca

En primer lugar se realizó un alineamiento entre los ORF traducidos y proteínas ortólogas para confirmar que se estaba trabajando con el gen deseado, esperando un alto grado de identidad entre las secuencias. El alineamiento fue global y pareado utilizando el algoritmo de Needleman – Wunsch (http://www.ebi.ac.uk/Tools/psa/emboss_needle/). Para la comparación se utilizó proteínas de oveja y en los casos donde no se habían reportado para esta especie se utilizaron proteínas humanas.

Utilizando las secuencias extraídas del genoma de referencia se identificó la región codante y la secuencia de aminoácidos de la proteína. Además se identificaron elementos regulatorios de la región promotora (caja TATA). Por otro lado utilizando las secuencias de la biblioteca de cDNA se pudo identificar la región 3'UTR de algunos de los genes, también se identificó la secuencia señal de poliadenilación (AAUAAA).

4.1.4 Diseño de cebadores y estandarización de PCR

Para el diseño de cebadores se utilizaron las secuencias extraídas a partir del genoma. Los cebadores fueron diseñados utilizando la interfase de internet de Primer3web 3.0 (<http://primer3.sourceforge.net/webif.php>) teniendo en cuenta que las temperaturas de fusión (T_m) entre cada par de cebadores no sean mayores a 2°C y que el contenido de GC

se encuentre entre 40 y 60 %. Para asegurar que la PCR fuera eficiente se calculó la energía libre de Gibbs (ΔG) para las distintas estructuras que podían adoptar los cebadores utilizando *Oligo Analyzer tool* de IDT (<https://www.idtdna.com/calc/analyzer>). Se calculó el ΔG para la formación de homodímeros, heterodímeros y estructuras secundarias (*Hairpin*). Se tomó como valor mínimo de ΔG -9 Kcal/mole para la formación de cada una de las estructuras, de acuerdo a las recomendaciones de los creadores de la herramienta de análisis utilizada.

Para la estandarización de las PCRs se realizaron curvas de magnesio y de temperatura. En el caso de la curva de magnesio, en la mayoría de casos, se utilizaron 3 concentraciones: 1mM, 1.5mM y 2mM. Las temperaturas que se probaron variaron de acuerdo al T_m de los cebadores diseñados. Las curvas de temperatura y de magnesio fueron realizadas utilizando 3 muestras de ADN genómico de alpaca de la variedad huacaya. La PCR se realizó en un termociclador de modelo Veriti de la marca *Applied Biosystems* utilizando las concentraciones finales indicadas en la Tabla 7.

Tabla 7. Concentraciones finales de los reactivos utilizados en la estandarización de las PCRs.

Reactivo	Concentración final
ADN	20ng/ μ l
PCRBuffer	1x
dNTPs	0.2mM
MgCl ₂	1.0/1.5/2.0 (mM)
Primer F.	10pM
Primer R.	10pM
TAQ	0.5U

Una vez estandarizadas las PCRs (por gen) se enviaron tres productos de cada par de cebadores a la empresa Macrogen (Korea), donde fueron purificados y secuenciados. Se usó un ABI PRISM *Big Dye Terminator Cycle Sequencing Ready Reaction Kit* y un secuenciador automático ABI PRISM 3730 Genetic Analyser®. En caso que hubiera

problemas con los cromatogramas obtenidos se diseñaron nuevos cebadores y se estandarizó nuevamente la reacción de acuerdo a lo descrito anteriormente.

4.1.5 Muestras para la validación de marcadores

Para validar los marcadores en los diferentes genes seleccionados se amplificaron 96 muestras de ADN genómico de alpaca huacaya provenientes de 6 regiones distintas del Perú (Tabla 8) (Anexo 2). Estas 6 regiones son las que presentan el mayor número de alpacas en el país (>90%) (Tabla 1), el número de alpacas utilizado por región fue proporcional al número de alpacas presentes en la misma.

Tabla 8. Numero de alpacas por región utilizadas para la validación de los marcadores.

Región	Número de individuos
Puno	42
Huancavelica	17
Cuzco	17
Arequipa	5
Apurímac	8
Junín	7

Para determinar el número total de muestras a analizar, se utilizó el programa MINSAGE, el cual fue desarrollado en la Universidad de Lübeck basado en las publicaciones de Gregorius y Gillet (Gregorius, 1980; Gillet, 1999). Este programa permite el cálculo del número mínimo de individuos necesarios en un marcador bialelico para encontrar alelos con una frecuencia mínima del 5% con un nivel de confianza del 95%. De acuerdo a este programa una muestra de 60 individuos es suficiente sin asumir equilibrio de Hardy-Weinberg.

4.1.6 Análisis de secuencias

Para cada gen se alinearon 40 secuencias al azar (de las 96) y se buscó sitios polimórficos. Una vez identificados las posiciones, éstas fueron observadas en los cromatogramas para

descartar errores de secuenciamiento. Se determinó mediante la inspección de los cromatogramas los 96 genotipos para aquellas posiciones que no correspondían a errores de secuenciamiento. En base a estos genotipos y utilizando el programa Cervus se determinó las frecuencias alélicas así como los índices de variabilidad: Índice de contenido polimórfico (PIC), heterocigosidad esperada (He) y heterocigosidad observada (Ho). Utilizando el programa Genepop se determinó si los marcadores se encontraban en equilibrio de Hardy-Weinberg. Además, se determinó los haplotipos y sus frecuencias para cada gen utilizando el programa Arlequin.

Por último se determinó qué marcadores se encontraban en desequilibrio de ligamiento utilizando el programa Arlequin. Para esto primero se realizó la prueba entre los marcadores dentro de un mismo gen. Los parámetros para esta prueba fueron 16000 permutaciones y 10 posiciones de inicio de acuerdo a la recomendación del desarrollador del programa. También se realizó la misma evaluación pero entre todas las posiciones polimórficas de todos los genes, los parámetros para esta prueba fueron los mismos. Además, para esta evaluación, debido al gran número de pruebas (496) y de que se trata de un *Likelihood ratio test* (LRT) se realizó la corrección de Benjamini-Hochberg (Benjamini y Hochberg, 1995).

4.2 Resultados

4.2.1 Determinación de genes candidatos a estar relacionados a características de fibra

Primero se buscaron todos los genes de la familia KRTAP en la biblioteca de cDNA. En total se encontraron 51 secuencias que pertenecían a 9 genes distintos, cada uno presentó un tamaño y un número de secuencias variable (Tabla 9).

Tabla 9. Genes de la familia KRTAP identificados en biblioteca de cDNA

Gen	Tamaño de secuencia (pb)	Numero de secuencias
KRTAP1-2	439	22
KRTAP3-3	350	4
KRTAP4-8	367	1
KRTAP7-1	394	2
KRTAP8-2	250	3
KRTAP9-2	142	1
KRTAP11-1	393	7
KRTAP17-1	167	1
KRTAP13-1	241	10

Se realizó una revisión bibliográfica con el fin de buscar reportes de polimorfismos en estos genes en otras especies productoras de fibra. Cuatro de los genes que se identificaron en la biblioteca de cDNA presentaban reportes de polimorfismos asociados a características de fibra en otras especies (KRTAP1-2, KRTAP9-2, KRTAP11-1 y KRTAP13-1). Además de estos 4 genes, se encontraron reportes de asociación del gen KRTAP6-1 al diámetro de fibra, por lo cual se incluyó en el trabajo. Por lo tanto se eligieron como genes candidatos a estar relacionados a características de fibra en alpaca a los genes: KRTAP1-2, KRTAP6-1, KRTAP9-2, KRTAP11-1, KRTAP13-1. Se describirá en detalle estos estudios en la sección de discusión.

4.2.2 Obtención de secuencias *in silico* a partir del genoma de alpaca

Para la obtención de las secuencias *in silico* se utilizó la herramienta BLAST y el genoma de referencia en el Genebank. Como se mencionó anteriormente no se encontró una secuencia para el gen KRTAP6-1 en la biblioteca de cDNA, por lo tanto, para ubicar este gen se utilizó la secuencia del gen en oveja. En todos los casos se encontró una región homóloga o idéntica en el genoma de alpaca con *E- Values* adecuados (Tabla 10). Dos de los genes (KRTAP11-1 y KRTAP6-1) se ubicaron en el mismo *scaffold* (Tabla 10).

Tabla 10. Región del genoma de alpaca donde se ubicaron las secuencias pertenecientes a los distintos genes.

Gen	Secuencia inicial	E- Value (Blast)	Scaffold	Posiciones en Scaffold
KRTAP1-2	Biblioteca cDNA (anexo 1)	0	188	3451998 - 3451456
KRTAP6-1	Genebank (anexo 1)	3.00 E-66	234	28605 - 28279
KRTAP9-2	Biblioteca cDNA (anexo 1)	6.00 E-67	627	54997 - 54856
KRTAP11-1	Biblioteca cDNA (anexo 1)	0	234	286258 - 285828
KRTAP13-1	Biblioteca cDNA (anexo 1)	0	101	150879 - 151598

Luego de determinar la posición de las secuencias se buscó un ORF cercano. Para todos los genes, menos KRTAP9-2, se encontró un ORF completo (Tabla 11). Una vez identificados los ORFs se extrajeron del genoma las secuencias de los genes para el diseño de cebadores (Tabla 11). El gen KRTAP9-2 presenta un *gap* en la parte 5' de la región codante. Para este gen se extrajo la secuencia incluyendo el *gap* y parte de la secuencia *downstream* del *gap* con la finalidad de diseñar cebadores para obtener la secuencia faltante mediante secuenciamiento.

Tabla 11. Tamaño de los ORFs encontrados y tamaño y posición de las secuencias a partir de las cuales se diseñaron los cebadores.

Gen	Tamaño del ORF (pb)	Scaffold	Tamaño de secuencia extraída (pb) y posición en scaffold
KRTAP1-2	390	188	1600 (3451301 - 3452901)
KRTAP6-1	225	234	1600 (27661 - 29261)
KRTAP9-2	No determinado	627	1600 (54066 - 55636)
KRTAP11-1	489	234	1600 (285665 - 287264)
KRTAP13-1	504	101	1600 (150402 - 151991)

4.2.3 Caracterización *in silico* de los genes candidatos en alpaca

Primero se realizaron alineamientos entre las proteínas traducidas a partir de los ORFs y las proteínas ortólogas en oveja. Esto se realizó para confirmar la identidad de los genes con los que se están trabajando. En el caso de KRTAP9-2 y KRTAP13-1 no se encontró las secuencias en oveja por lo tanto se utilizó las proteínas reportadas en humanos. Además del alineamiento se analizó el gen *in silico* para identificar la región codante y elementos regulatorios de la región promotora, 3'UTR y 5'UTR. También se presenta la secuencia predicha de la proteína en alpaca.

4.2.3.1 KRTAP1-2

Se realizó un alineamiento entre el ORF traducido del gen KRTAP1-2 en alpaca y la secuencia de la proteína KAP1-2 reportada en oveja (Acc. Nr. AEK78078.1) (Figura 14). Se encontró un porcentaje de identidad de 84 %, donde la diferencia más importante se encuentra en una delección de 25 aminoácidos en la secuencia de alpaca con respecto a la secuencia de oveja.

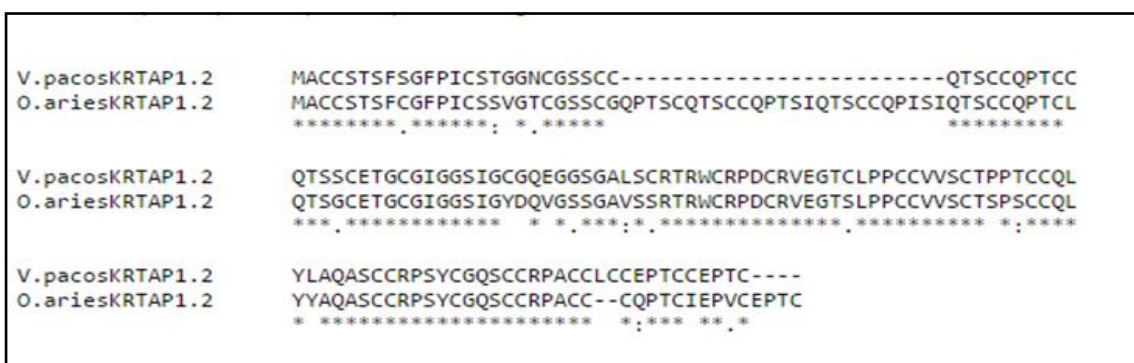


Figura 14. Alineamiento entre la proteína KRTAP1-2 reportada en oveja y la proteína predicha del gen KRTAP1-2 en alpaca.

Al analizar la secuencia del gen KRTAP1-2 en alpaca se identificaron elementos y regiones regulatorias. La región codante de este gen tiene un tamaño de 390 pb, se identificó la secuencia consenso de la caja TATA a 83 pb *upstream* del codón de inicio (Figura 15A). Utilizando la secuencia en la biblioteca de cDNA se identificó la región

3'UTR, la cual tiene un tamaño de 439 pb (Figura 15A). Además, dentro de esta región y previo al final de la secuencia de cDNA se pudo identificar la secuencia señal de poliadenilación AAUAAA a 426 pb del último codón (Figura 15A). De acuerdo a esto, tomando en cuenta la posición de la caja TATA como el inicio y la última porción de la secuencia de biblioteca de cDNA como el final, el gen KRTAP1-2 en alpaca tiene un tamaño aproximado de 912 pb (Figura 15A).

Además de identificar estos elementos en la secuencia de ADN se tradujo la región codante para obtener la secuencia de aminoácidos de la proteína, obteniéndose una secuencia de 130 aminoácidos que tiene un peso molecular de 13.2 kDa. (Figura 15B). La proteína presenta 26% de residuos de cisteína, los dos otros residuos más frecuentes son serina (12.3%) y glicina (11.5%) (Figura 15B). Debido a que la proteína presenta un porcentaje menor al 30% de residuos de cisteína, se le clasifica en el grupo HS (Tabla 3).



Figura 15A. Probable secuencia del gen KRTAP1-2 en alpaca. En letras negras se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, en verde se muestra la secuencia de la biblioteca de cDNA (región 3'UTR) y en rojo la secuencia señal de poliadenilación (AAUAAA). **15B.** Secuencia traducida de aminoácidos de la proteína KAP1-2 en alpaca obtenida a partir de la región codante identificada *in silico*.

4.2.3.2 KRTAP6-1

Se realizó un alineamiento entre el ORF traducido del gen KRTAP6-1 en alpaca y la secuencia de la proteína KAP6-1 reportada en oveja (Acc. Nr. NP_001180328.1) (Figura 16). Ambas proteínas comparten una identidad del 86.7%, diferenciándose en gran parte debido a una delección en la secuencia de alpaca de 8 residuos con respecto a la secuencia en oveja (Figura 16).

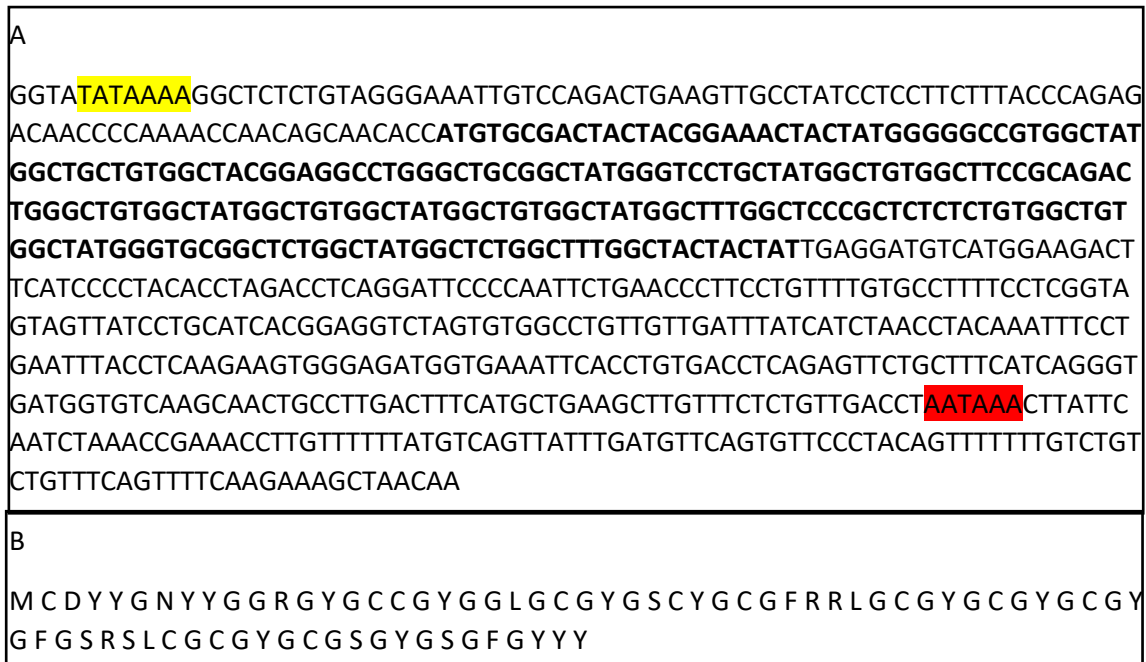


Figura 17A. Probable secuencia del gen KRTAP6-1 en alpaca. En letras negras se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, y en rojo la secuencia señal de poliadenilación (AAUAAA). **17B.** Secuencia traducida de aminoácidos de la proteína KAP6-1 en alpaca obtenida a partir de la región codante identificada *in silico*.

4.2.3.3 KRTAP9-2

En el caso de KRTAP9-2 como se mencionó previamente, el genoma de referencia presentaba un *gap*. Los resultados que se presentan a continuación corresponden luego de descifrar la secuencia de ese *gap*. La metodología que se utilizó para el establecimiento de esta secuencia de describirá más adelante. Se realizó un alineamiento entre el ORF traducido del gen KRTAP9-2 en alpaca y la secuencia de la proteína KAP9-2 reportada en humano (Acc. Nr. NP_114167.2) (Figura 18). Se observó una identidad entre ambas proteínas del 65%, considerablemente menor que en los casos anteriores.

```

V.pacosKAP9.2      MTHSCCSPCCQPTCCQPTCCPEPSCCQPCCPQTCYEATQTTCCSTTCRKPTCVTTCCQPTC
H.sapiensKAP9.2    -MTHCCSPCCQPTCCRTTCCRITTCWKPTTVTTC---SSTPCCQPACCVSSCCQPCCRPTC
                    *****; ** . :* :* ** :.* ** . :* :* **;***

V.pacosKAP9.2      CEPSSCGQTCQPTCVTSCCSPSCCQS----TCCSSSCGQNCGGSSCCQPASCAPVYC
H.sapiensKAP9.2    CQNTCCRTTCCQPTCVTSCCQPSCCSTPCCQPTCCGSSCCGQTS CGSSCGQSSSCAPVYC
                    *; :.* *****.***.; ** **.* ** . *;*****

V.pacosKAP9.2      HRTWYHPTCCCLPGCLAQSCGSSCCQPRCRPVCCQTTCCR----PSCVSSCCQPSC-
H.sapiensKAP9.2    RRTCYYPTTVCLPGCLNQSCGSNCCQPCCRPACCETTCCRTTCFQPTCVSSCCQPSCC
                    ;** *; ** ***** ***** .*** ** .;*;***** *;*****

```

Figura 18. Alineamiento entre la proteína KRTAP9-2 reportada en humano y la proteína predicha del gen KRTAP9-2 en alpaca.

La región codante de este gen tiene un tamaño de 504 pb, se identificó la secuencia consenso de la caja TATA a 84 pb upstream del codón de inicio (Figura 19A). A pesar que este gen si se encontró en la biblioteca de cDNA, solo se encontró una secuencia que pertenecía a la región codante y no al extremo 3'UTR. Por lo tanto tampoco se pudo determinar de esta forma la secuencia de esta región. La señal de poliadenilacion AAUAAA se identificó a 468 pb *downstream* del último codón. De acuerdo a esto, el gen KRTAP9-2 en alpaca tiene un tamaño aproximado de 1057 pb.

<p>A</p> <p>AGTGTTATAAAAGCCCCAGAACAGAAGGAGTGATCAGAATCTGAGAACTCACCACTCAACAGGAGTTC CAACTTCCACCACTGACACCATGACCCACTCATGCTGCTCCCCTTGCTGCCAGCCCACCTGCTGCCAGCC CACCTGCTGTGAGCCCAGCTGCTGCCAGCCTTGCTGCCCCAAACTTGCTATGAAGCTACTCAGACCAC CTGCTGCAGCACCACTGCCGCAAACCTACTTGTGTGACCACCTGCTGCCAGCCCACCTGCTGTGAGCC CAGCAGCTGTGGACAAACCTGCTGCCAGCCTACCTGTGTAACCAGCTGTTGCAGCCCTCCTGCTGCC AGTCCACTTGCTGTGAGTCTAGCAGCTGCGGACAAAAGTGCAGGTGGGTCCAGCTGCTGCCAGCCAGC TAGCTGTGCACCCGTGACTGCCACCGAACCTGGTACCACCCACGTGCTGCTGCCTGCCTGGGTGCCT GGCCAGAGCTGTGGATCCAGCTGCTGCCAGCCTCGCTGCCGCCCTGTCTGCTGTCAGACCACGTGCT GCCGCCCTAGCTGTGTGTCCAGCTGCTGCCAGCCCTCCTGCTGATCACCTACCAAGAGCCATCCCCTG CATCCAACACAATCTGTCAACTGAGTTGCCGTTTTGGGGGCAAATTCACCTCTTGGGTTTGCCAATTACC AGCATGCTCTCCAGAACTTCTGTTACTCATCTGCTTGTAAAGCTTGTGAATCAGCTTGTGGAGTGC AGAGGACTTCCCCAATTCTTCTCCTCTCGTGTAGGTGGATGATGAGCCAGCTTGTGCATCCTTGATAT GGAGTTTGTCTGGGCTTTGACTCTGAAGTCAGGTCCACGCAGTCCCTCTGAGTGACTAAGGAAAATAA TTCTTGTAAGTGTTCACAGTTTCTGCAACTGATCAATATTATTATAATCATATTTCTTTATCAATGTA CTCGTGGCTCTCACAGCTTTCTTGTCTTTATGATCACTTTGAGTTGTCTTCTGATACAGGAGTCTT CCCTATGTGTCTCCCAAATAAATTCCGTATCATAATTTAT</p>
<p>B</p> <p>MTHSCCSPCCQPTCCQPTCCEPSCCQPCCPQTCYEATQTTCCSTTCRKPTC VTTCCQPTCCEPSSCGQTCQPTCVTSCCSPSCCQSTCCESSSCGQNCGGS SCCQPASCAPVYCHRTWYHPTCCCLPGCLAQSCGSSCCQPRCRPVCCQTT CCRPSCVSSCCQPSC</p>

Figura 19A. Probable secuencia del gen KRTAP9-2 en alpaca. En letras negras se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, y en rojo la secuencia señal de poliadenilación (AAUAAA). 19B. Secuencia traducida de aminoácidos de la proteína KAP9-2 en alpaca obtenida a partir de la región codante identificada *in silico*.

Además de identificar estos elementos en la secuencia de ADN se tradujo la región codante para obtener la secuencia de aminoácidos de la proteína, se obtuvo una secuencia de 167 aminoácidos con un peso molecular de 7.5 kDa. Al determinar su composición se observó que cisteína es el residuo que aparece con mayor frecuencia 31.7%. Otros residuos frecuentes en esta proteína son serina (13.8%), treonina (12.6%) y prolina (11.4) (Figura 19B). El porcentaje de cisteína mayor al 30% es lo que se espera para esta proteína por pertenecer al grupo UHS (tabla 3).

4.2.3.4 KRTAP11-1

Se realizó un alineamiento entre el ORF traducido del gen KRTAP11-1 en alpaca y la secuencia de la proteína KAP11-1 reportada en oveja (Acc. Nr. AEB55022.1) (Figura 20). Se encontró un porcentaje de identidad de 83.4 % entre ambas proteínas, presentando la proteína en alpaca una pequeña inserción de 3 aminoácidos con respecto a la proteína de oveja (Figura 20).

V.pacosKAP11.1	MSYNCSTRNCSSRLIGGQYSVPVTPVVTASTQDADCLSGIHLPSFQGTGSWLLDHCQETC
O.ariesKAP11.1	MSYSCSTRNCSSRRIGGEYTPV---VTVSSPDADCLSGIYLPSSFQGTGSWLLDHCQETC
	.** ***,*:* **,*: *****:*****
V.pacosKAP11.1	CEPTVCQPTCYQQTSCVSSPVQVTCRQTNCVSNPCSTAYSRLPSFVSSGCRPLGGISTV
O.ariesKAP11.1	CEPTVCQSTCYQPTPCVSSPVRVT-SRQTTCVSSPCSTTCRPLTFISSGCQPLSGVSTV
	***** **** * *****:* ***,*:* ***,*:* ***,*:* ***,*:* ***,*:*
V.pacosKAP11.1	CQPVRSVSTVCRPVGGVSTICQPACGVSRTYQQSCVSSCRRTC
O.ariesKAP11.1	CKPVRISITVCQPVGGVSTICQPTCGVSRTYQQSCVSSCRRIC
	:

Figura 20. Alineamiento entre la proteína KRTAP11-1 reportada en oveja y la proteína predicha del gen KRTAP11-1 en alpaca.

La región codante de este gen tiene un tamaño de 439 pb, se identificó la secuencia consenso de la caja TATA a 83 pb *upstream* del codón de inicio (Figura 21A). Utilizando la secuencia en la biblioteca de cDNA se identificó la región 3'UTR, la cual tiene un tamaño de 393 pb (Figura 21A). Esta secuencia es muy probable que represente la región 3'UTR del gen debido que al final presento una región de poliA. Además dentro de esta región y previo al final de la secuencia de cDNA se pudo identificar una secuencia señal de poliadenilacion AAUAAA a 374 pb del último codón (Figura 21A). De acuerdo a esto, el gen KRTAP11-1 en alpaca tiene un tamaño aproximado de 965 pb.

Además de identificar estos elementos en la secuencia de ADN se tradujo la región codante para obtener la secuencia de aminoácidos de la proteína, se obtuvo una secuencia de 167 aminoácidos con un peso molecular de 17.3 kDa (Figura 21B). Los aminoácidos

más representados fueron serina (17.2%), cisteína (12.3%), valina (11%) y treonina (10.4%) (Figura 21B). Esta frecuencia de residuos de cisteína va de acuerdo a lo que se espera para una proteína del grupo HS (Tabla 3).

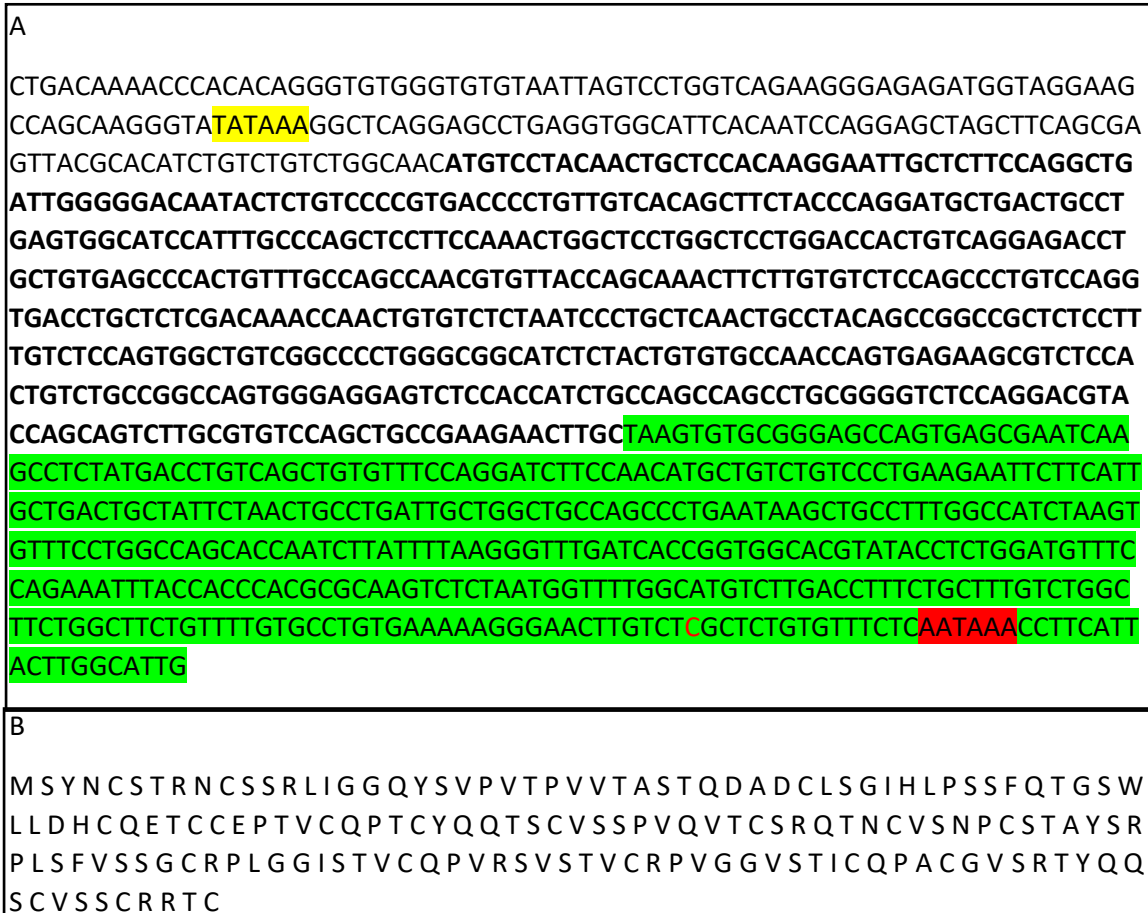


Figura 21A. Probable secuencia del gen KRTAP11-1 en alpaca. En letras negras se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, en verde se muestra la secuencia de la biblioteca de cDNA (región 3'UTR) y en rojo la secuencia señal de poliadenilación (AAUAAA). **21B.** Secuencia traducida de aminoácidos de la proteína KAP11-1 en alpaca obtenida a partir de la región codante identificada *in silico*.

4.2.3.5 KRTAP13-1

Se realizó un alineamiento entre el ORF traducido del gen KRTAP13-1 en alpaca y la secuencia de la proteína KAP13-1 reportada en humano (Acc. Nr. EAX09899.1) (Figura 22). Se encontró un porcentaje de identidad de 61.9 % (Figura 22), este valor al igual que el obtenido para KRTAP6-1 fue más bajo que al realizar las comparaciones entre oveja y alpaca.

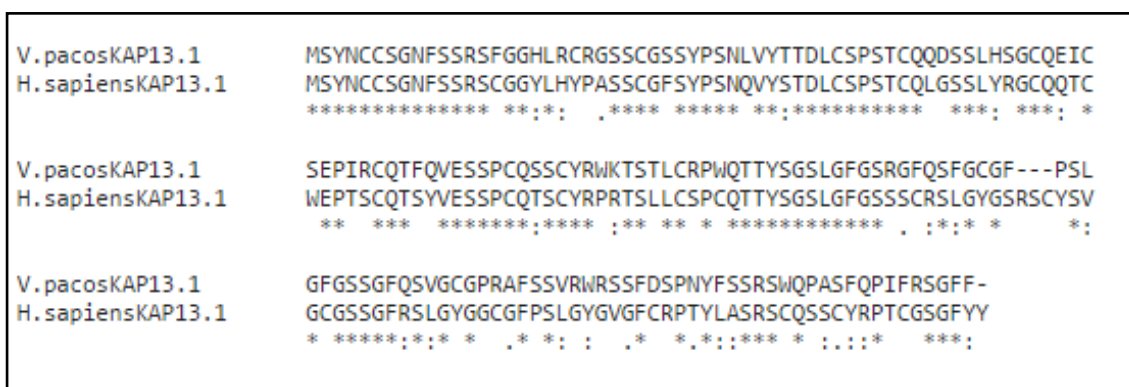


Figura 22. Alineamiento entre la proteína KRTAP13-1 reportada en humano y la proteína predicha del gen KRTAP13-1 en alpaca.

La región codante de este gen tiene un tamaño de 504 pb, se identificó la secuencia consenso de la caja TATA a 83 pb *upstream* del codón de inicio (Figura 23A). La secuencia de este gen en la biblioteca de cDNA presento 720 pb de las cuales 210 corresponden a la región 3'UTR, el resto a región codante. Por lo tanto utilizando esta secuencia se identificó parte de la región 3'UTR, sin embargo dentro de esta secuencia no se encuentra la señal de poliadenilación (AAUAAA) (Figura 23A). Esta se pudo identificar 118 pb *downstream* del último nucleótido de la secuencia proveniente de la biblioteca de cDNA. De acuerdo a esto, el gen KRTAP13-1 en alpaca tiene un tamaño aproximado de 915 pb (Figura 23A).

Al traducir la región codante se obtuvo la secuencia de una proteína de 168 aminoácidos con un peso molecular de 18.4 kDa (Figura 23B). El residuo más frecuente es serina que

aparece en un porcentaje de 23.2%. El residuo Cisteína está presente en un 8%, otro aminoácido bastante frecuente es Glicina con 10.7% (Figura 23B). La frecuencia de cisteína es baja en comparación al resto de proteínas pero aun así está dentro del rango del grupo HS (Tabla 3).



Figura 23A. Probable secuencia del gen KRTAP13-1 en alpaca. En letras negritas se muestra la región codante, en amarillo se muestra la secuencia consenso de la caja TATA, en verde se muestra la secuencia de la biblioteca de cDNA parte de la región 3'UTR y en rojo la secuencia señal de poliadenilación (AAUAAA). **23B.** Secuencia traducida de aminoácidos de la proteína KAP11-1 en alpaca obtenida a partir de la región codante identificada *in silico*.

4.2.4 Diseño y prueba de cebadores

Para el diseño de cebadores se utilizó el programa Primer3. Al diseñar los cebadores se tuvo en cuenta varios parámetros como, por ejemplo, el T_m y el contenido de GC. Además se calculó el ΔG para la formación de estructuras secundarias, donde se aceptaron únicamente valores mayores a -9 kcal/mol para garantizar una PCR efectiva. Si bien todos los cebadores diseñados amplificaron las regiones deseadas, en algunos casos hubo la necesidad de diseñar más de un par de cebadores. En el caso de KRTAP6-1 y KRTAP1-2 se tuvo que diseñar dos pares de cebadores debido a la presencia de una inserción/delección en la región amplificada por el primer par. Los individuos heterocigotos para la inserción/delección mostraban cromatogramas no informativos, debido a una superposición de picos. El segundo par se diseñó para amplificar una región que no incluya la inserción/delección. Para amplificar el gen KRTAP9-2 se tuvieron que diseñar 3 pares distintos de cebadores, el primero fue para determinar la región faltante del gen en el genoma de referencia. El segundo par amplificó una región que contenía nuevamente una inserción/delección, por lo que fue necesario el diseño de un tercer par que no incluyera esa región. Para los genes KRTAP11-1 y KRTAP13-1 solo fue necesario el diseño de un par de cebadores debido a la ausencia de inserciones/delecciones. A continuación se explicará en detalle cada uno de los casos. Para las pruebas de los cebadores se realizó una curva de temperatura y una curva de magnesio utilizando 3 muestras de ADN genómico de alpaca huacaya. Todas las reacciones de secuenciamiento se realizaron en ambas direcciones (*Forward* y *Reverse*)

4.2.4.1 KRTAP1-2

4.2.4.1.1 Diseño y prueba del primer par de cebadores (1F y 1R)

El primer par de cebadores (1F y 1R) que se diseñó para amplificar el gen KRTAP1-2 en alpaca, flanqueaba una región de 765 pb que contenía la región codante del gen (Tabla 12) (Figura 27). Para el diseño se contempló el ΔG para la formación de estructuras secundarias entre los cebadores, ninguno de los valores fue menor a -9 kcal/mole (Tabla 12).

Tabla 12. Información de los cebadores diseñados para amplificar el gen KRTAP1-2.

Cebador	Secuencia	T _m (°C)	Tamaño amplicon	ΔG Hairpin (Kcal/mole)	ΔG Homodimero (Kcal/mole)	ΔG Heterpdimero (Kcal/mole)
1F	ACCAGTGAACGCTAATGACT	58	765 pb	0.27	-3.61	-4.41
1R	CCAGGGGTGATTGAGGAGA	60		0.7	-3.07	
1NF	TGGCATTGGTGGTAGCATTG	60	894 pb	-0.52	-3.14	-6.21
1NR	AACTGGAGGCCTGTGTAAC	58		-0.29	-3.07	

Posterior al diseño se realizó una curva de temperatura y de magnesio utilizando tres muestras de ADN genómico de alpaca. Las temperaturas que se evaluaron fueron: 56°C, 58°C y 60°C. Las concentraciones de MgCl₂ que se probaron fueron: 1mM, 1.5mM y 2mM (Figura 24). A la concentración de 1 mM de magnesio no se observó producto de PCR a ninguna temperatura, mientras que si hubo producto para el resto de concentraciones. Se observó producto de PCR para todas las temperaturas que se probaron. Todos los productos presentaron un mismo tamaño que va de acuerdo a lo esperado al diseñar los cebadores (Figura 24).

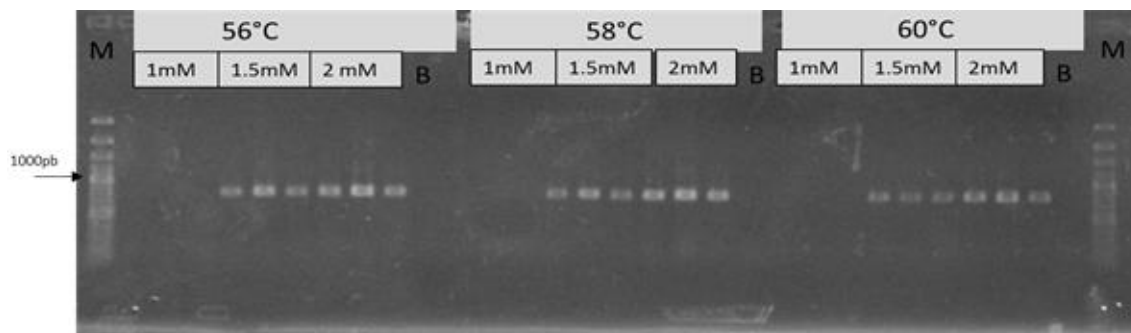


Figura 24. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 1F y 1R a diferentes concentraciones de magnesio y temperatura, todos los productos se encuentran en un tamaño cercano a 800 pb. El pozo con la M hace referencia al marcador de peso molecular (100bp), El pozo con la B hace referencia a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 1000 pb.

Teniendo en cuenta como criterio la intensidad de banda se estandarizó la reacción a una concentración de 2 mM de magnesio y 58 °C (Tabla 13).

Tabla 13. Condiciones a la cuales se estandarizaron las PCR para los cebadores diseñados en el gen KRTAP1-2.

Par de cebadores	Concentraciones		Condiciones Termicas		
	Reactivos	Concentracion final	Temperatura	Tiempo	Numero de ciclos
1F y 1R	ADN	20ng	94°C	4 minutos	1
	Buffer PCR	1X			
	dNTPs	0.2mM	94°C	30 segundos	30
	MgCl2	2 mM	62°C	45 segundos	
	Primer Forward	10 picomoles	72°C	50 segundos	
	Primer Reverse	10 picomoles	72°C	5 minutos	1
	TAQ polimerasa	0.5 U			
1NF y 1NR	ADN	20ng	94°C	4 minutos	1
	Buffer PCR	1X			
	dNTPs	0.2mM	94°C	30 segundos	30
	MgCl2	2 mM	62°C	45 segundos	
	Primer Forward	10 picomoles	72°C	1 minuto	
	Primer Reverse	10 picomoles	72°C	5 minutos	1
TAQ polimerasa	0.5 U				

Se amplificaron 3 muestras utilizando las concentraciones a las que se estandarizó la reacción y se enviaron a secuenciar los productos de PCR a la empresa MacroGen Inc. Al observar los cromatogramas se encontró que uno de ellos presentaba un patrón anómalo (Figura 25). Este patrón presentaba picos bien definidos y de buena calidad hasta cierta

región de la secuencia, donde se observaba una constante superposición de picos. Esto se debe a la presencia de una inserción/delección en la región codante. Este patrón anómalo aparece solo en aquellas muestras que son heterocigotos para la inserción/delección. Se identificó que la inserción/delección es de 30 nucleótidos gracias a la presencia de un individuo homocigoto para esta. Este individuo con respecto al genoma de referencia presenta una inserción de 30 nt en la posición 80 de la región codante del gen (Figura 27). Para resolver este problema se diseñaron nuevos cebadores que excluyeran la inserción/delección.

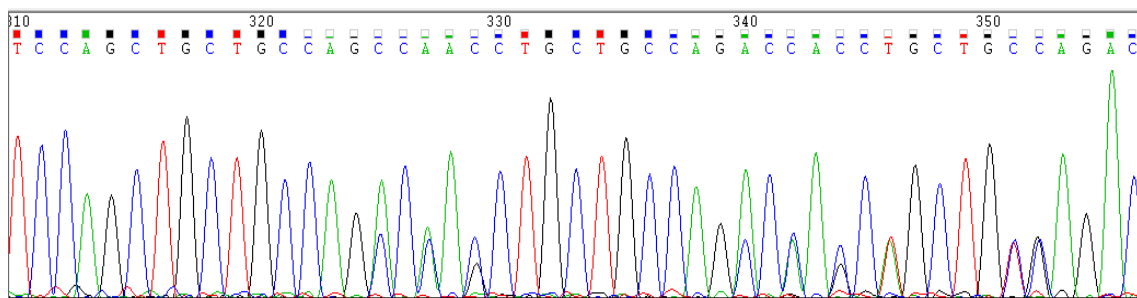


Figura 25. Cromatograma con patrón anómalo de la secuencia de KRTAP1-2 obtenido con los cebadores 1F y 1R. Se observa un gran número de picos superpuestos, esto se debe que el cromatograma corresponde a un individuo heterocigoto para la inserción/delección.

4.2.4.1.2 Diseño y prueba del segundo par de cebadores (1NF y 1NR)

El segundo par de cebadores se diseñó con el propósito de amplificar parte de la región codante y la región 3'UTR sin incluir la inserción/delección, el tamaño del amplicon es de 894 pb (Tabla 12) (Figura 27). Se evaluó el ΔG para la formación de estructuras secundarias, ninguno de los valores fue menor a -9 kcal/mole (Tabla 12).

Posterior al diseño de los cebadores se realizó una curva de temperatura y de magnesio utilizando tres muestras de ADN genómico de alpaca. Las temperaturas que se evaluaron

fueron: 60°C, 62°C y 64°C. Las concentraciones de MgCl₂ que se probaron fueron: 1.5mM y 2mM (Figura 26). Se observó producto de PCR para todas las temperaturas y concentraciones de magnesio que se probaron. Todos los productos presentaron un mismo tamaño que va de acuerdo a lo esperado (Figura 26).

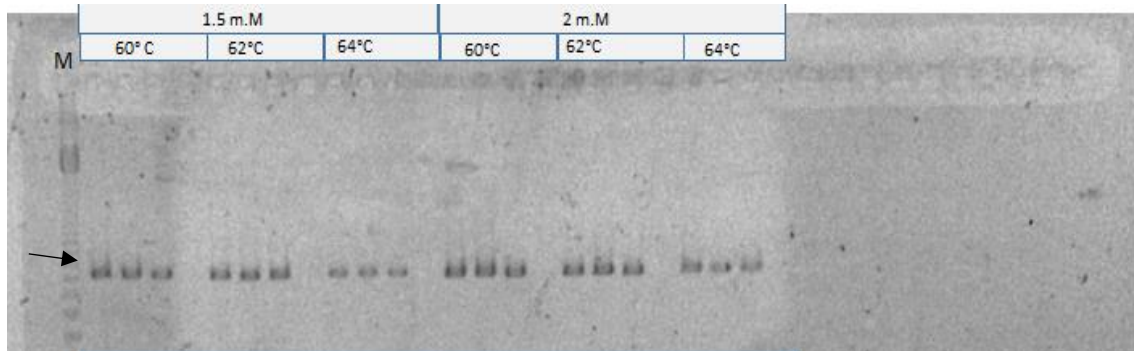


Figura 26. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 1NF y 1NR a diferentes concentraciones de magnesio y temperatura, las bandas se encuentran entre 984 y 861 pb. El pozo con la M hace referencia al marcador de peso molecular (123bp), el pozo en blanco entre muestras corresponde a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 984 pb.

De acuerdo a la intensidad de banda se estandarizó la reacción de PCR a la temperatura de *annealing* de 64°C y a una concentración de MgCl₂ de 2 mM (Tabla 13). Se enviaron 3 muestras a secuenciar en ambas direcciones para observar la calidad de los cromatogramas. En todos los casos se obtuvieron cromatogramas con picos bien definidos y sin presentar el patrón anómalo anterior.

Posiciones de los cebadores con respecto a la secuencia

```
CCTGTCATGTGAGACCCAAGAATTCAATTACAAGGGTCAACACTGAGTCATATTGACACTCCTTAAAGG
TGAGGCTTCAGTTAATTATTCTTGGGGCAATGAGCAAATTTAACAATATGTGGAATTAGGGGTTATA
TAACCAGTGAACGCTAATGACTACTTGACCATGGTGAAGAATCCTAATTGAAGGCCAAATAAAATTAAC
TTTTAAAAGCTTTCAGAGGTTGAACAAGTAACTCAGGGATAGGCCAATGACCATTGAGAGACAAGTTAT
TTTTAAAACAAACAAGGAAGGGGCTTTAAGAATTATGTAAATAATAGCTGGTCAAGGCTTATAAAAGGC
CAGATGTAGAAGGTGAAGCCAAAACACTCAGAACTCCTCTTAACAAGCCAGCTCTCAGCCCAACTCTTGA
TACCATGGCCTGCTGTTCCACCAGCTTCAGCGGATTTCCCATCTGTTCTACTGGTGGGAACTGTGGCTC
CAGCTGCTGCCAGAC*CAGCTGCTGCCAGCCAACCTGCTGCCAGACCAGCAGCTGTGAGACTGGCTGT
GGCATTGGTGGTAGCATGGCTGTGGCCAGGAGGGTGGCAGCGGAGCTCTGAGCTGCCGCACCAGG
TGGTGCCGCCCTGACTGCCCGGTGGAGGGCACCTGCCTGCCTCCCTGCTGCGTGGTGAGCTGCACCCC
ACCAACCTGCTGCCAGCTGTACCTCGCCAGGCCTCCTGCTGCCGCCATCCTACTGTGGACAGTCTTG
CTGCCGCCAGCCTGCTGCCTCTGCTGTGAGCCACCTGCTGTGAGCCACCTGCTGAGAGCCACCTG
CTAAAAGTCAGTTTTCTGATTTCAACTTGAAATTTCAACTTCCAGCTCAGTCCATGAACTAATAATCTCC
TCAATCACCCCTGGGCCACTAACAAGTTCTTGAATTTTATTGACTCTTTTTTGAAGGTTTACCAAATAT
TTGGGCCTCACAGACCTGTCTGATTCCATTCAATGCTAAAAGATCTTGAGCGCTCCACTGAAGCTGCC
AAGATATAAGTTTGACCTAAGAAATGTGAAAACCAATTTGCCTCAGGATCAGCCTTCAGCAAGCATCTT
CTCCACGGCCAGTGCTTCCAAGGCTGTGGCAGAACCATCTGTCCTTCTCAGAAACATTCATCTCCTAAAC
TCCACCATCTTGCAAACCTGCATTTCTATGATAGGGGAACAATATACCGAGGTTTAATAAACTCTATCTT
TGGTACAGCAGACATGTCTCTTTTCTTTGAGTCTGCTCATGATGATATTTTGGAGGAAGTCACAC
AGTAAGATTCTAATCTGGTTTTGATATCCCACTGTGATTCTTATACCGTGTAGGAAACATGATCAAA
TCCAGGGGAGGTTAATCCCCATAAAAAGATACAGCAGAGGTGGTTACACAGGCCTCCAGTGGCATGG
```

Figura 27. Secuencia a partir de la cual se diseñaron ambos pares de cebadores y las posiciones de estos dentro de la misma. En amarillo el primer par de cebadores (1F y 1R), en rojo el segundo par de cebadores (1NF y 1NR), en negrita la región codante y un asterisco indica la posición de la inserción/delección.

4.2.4.2 KRTAP6-1

4.2.4.2.1 Diseño y prueba del primer par de cebadores (6F y 6R)

El primer par de cebadores (6F y 6R) que se diseñó para amplificar el gen KRTAP6-1 en alpaca, flanqueaba una región de 749 pb que contenía la región codante del gen (Tabla 14) (Figura 36). Para el diseño se contempló el ΔG para la formación de estructuras secundarias entre los cebadores, ninguno de los valores fue menor a -9 kcal/mole (Tabla 14).

Tabla 14. Información de los cebadores diseñados para amplificar el gen KRTAP6-1.

Cebador	Secuencia	Tm (°C)	Tamaño amplicon	ΔG Hairpin (Kcal/mole)	ΔG Homodimero (Kcal/mole)	ΔG Heterpdimero (Kcal/mole)
6F	TCAGATGTGACTCAACATTAGC	62	749 pb	-0.72	-4.77	-4.87
6R	TTAGGTCAACAGAGAAACAAG	58		0.79	-1.94	
6NF	TGGGGTGCAGGTGTATTTATG	62	958 pb	0.54	-7.05	-8.09
6NR	CAGGACCCATAGCCGCAG	60		0.95	-3.61	

Posterior al diseño se realizó una curva de temperatura y de magnesio utilizando tres muestras de ADN genómico de alpaca. Las temperaturas que se evaluaron fueron: 56°C, 58°C y 60°C. Las concentraciones de MgCl₂ que se probaron fueron: 1mM, 1.5mM y 2mM (Figura 28). A la concentración de 1 mM de magnesio no se observó producto de PCR a ninguna temperatura, mientras que si hubo producto para el resto de concentraciones. Se observó producto de PCR para todas las temperaturas que se probaron. Todos los productos presentaron un mismo tamaño que va de acuerdo a lo esperado al diseñar los cebadores (Figura 28).

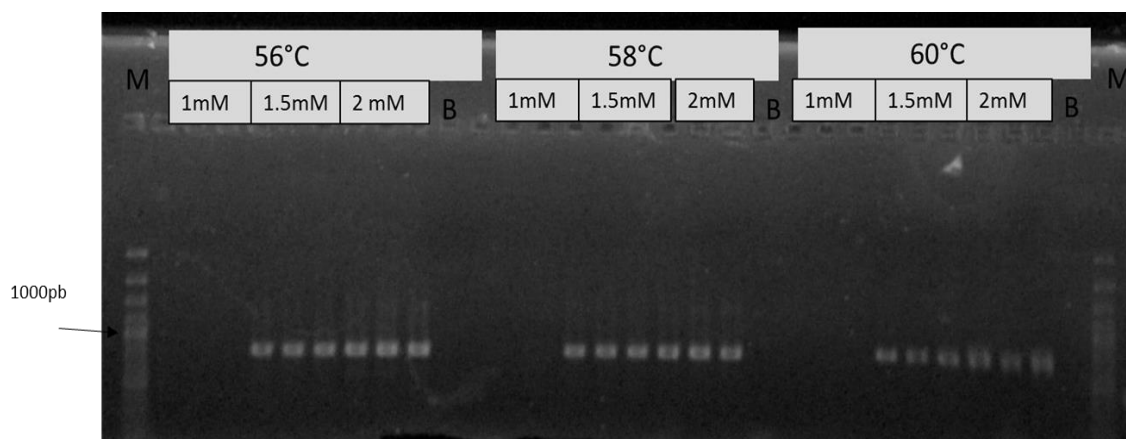


Figura 28. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 6F y 6R (KRTAP6-1) a diferentes concentraciones de magnesio y temperatura, se observa una banda cerca a las 800 pb. El pozo con la M hace referencia al marcador de peso molecular (100bp), El pozo con la B hace referencia a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 1000 pb.

Teniendo en cuenta como criterio la intensidad de banda se estandarizo la reacción a una concentración de 2 mM de magnesio y 58 °C (Tabla 15).

Tabla 15. Condiciones a las cuales se estandarizaron las reacciones de PCR para los cebadores en el gen KRTAP6-1.

Par de cebadores	Concentraciones		Condiciones Termicas		
	Reactivos	Concentracion final	Temperatura	Tiempo	Numero de ciclos
6F y 6R	ADN	20ng	94°C	4 minutos	1
	Buffer PCR	1X			
	dNTPs	0.2mM	94°C	30 segundos	30
	MgCl2	2 mM	62°C	45 segundos	
	Primer Forward	10 picomoles	72°C	50 segundos	
	Primer Reverse	10 picomoles	72°C	5 minutos	1
	TAQ polimerasa	0.5 U			
6NF y 6NR	ADN	20ng	94°C	4 minutos	1
	Buffer PCR	1X			
	dNTPs	0.2mM	94°C	30 segundos	30
	MgCl2	2 mM	62°C	45 segundos	
	Primer Forward	10 picomoles	72°C	1 minuto	
	Primer Reverse	10 picomoles	72°C	5 minutos	1
	TAQ polimerasa	0.5 U			

Se amplificaron 3 muestras utilizando las concentraciones a las que se estandarizo la reacción y se enviaron a secuenciar los productos de PCR a la empresa MacroGen Inc. Al observar los cromatogramas se encontró que una muestra presentaba el patrón anómalo mencionado anteriormente para el gen KRTAP1-2 (Figura 29). Este patrón se debía a una inserción/delección la cual se identificó gracias a la presencia de un individuo homocigoto para esta. Este individuo, con respecto al genoma de referencia, presenta una delección de 12 nt en la posición 144 de la región codante del gen (Figura 31). Para resolver este problema se diseñaron nuevos cebadores que excluyeran la inserción/delección.

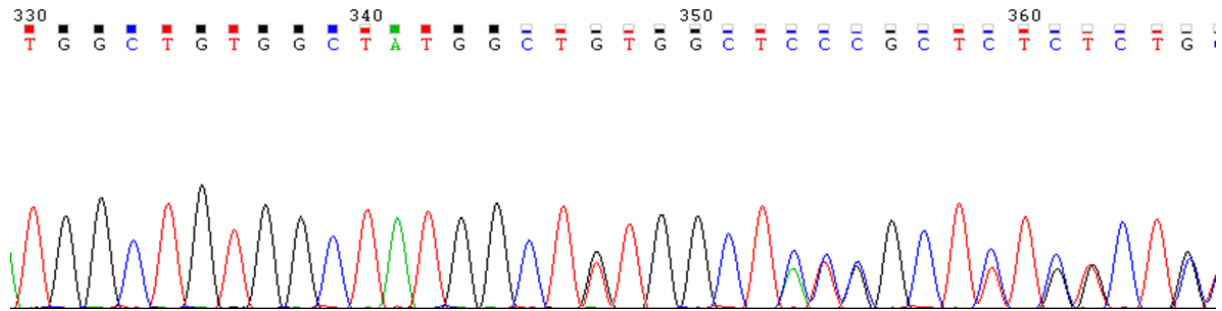


Figura 29. Cromatograma con patrón anómalo de la secuencia de KRTAP6-1 obtenido con los cebadores 6F y 6R. Se observa un gran número de picos superpuestos, esto se debe que el cromatograma corresponde a un individuo heterocigoto para la inserción/delección.

4.2.4.2.2 Diseño y prueba del segundo par de cebadores (6NF y 6NR)

El segundo par de cebadores (6NF y 6NR) se diseñó con el propósito de amplificar parte de la región codante y la región 5'UTR sin incluir la inserción/delección, el tamaño del amplicon es de 958 pb (Tabla 14) (Figura 31). Se evaluó el ΔG para la formación de estructuras secundarias ninguno de los valores fue menor a -9 Kcal/mole (Tabla 14).

Posterior al diseño de los cebadores se realizó una curva de temperatura y de magnesio utilizando tres muestras de ADN genómico de alpaca. Las temperaturas que se evaluaron fueron: 58°C, 60°C y 62°C. Las concentraciones de $MgCl_2$ que se probaron fueron: 1.5mM y 2mM (Figura 30). Se observó producto de PCR para todas las temperaturas y concentraciones de magnesio que se probaron. Todos los productos presentaron un mismo tamaño que va de acuerdo a lo esperado al diseñar los cebadores (Figura 30).

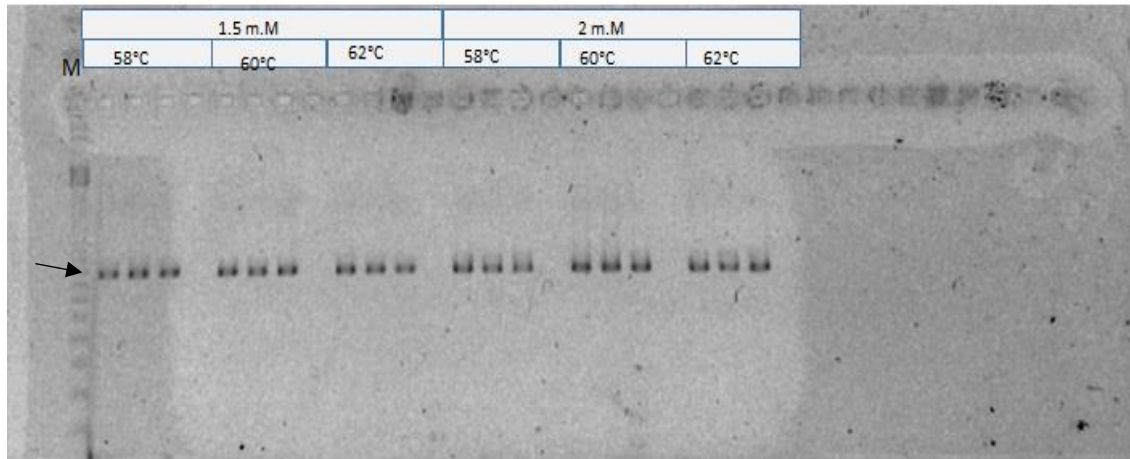


Figura 30. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 6NF y 6NR a diferentes concentraciones de magnesio y temperatura, las bandas se encuentran entre 861 y 984 pb. El pozo con la M hace referencia al marcador de peso molecular (123bp), el pozo en blanco entre muestras corresponde a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 984 pb.

De acuerdo a la intensidad de banda se estandarizó la reacción de PCR a la temperatura de *annealing* de 62°C y a una concentración de MgCl₂ de 2 mM (Tabla 15). Se mandaron 3 muestras a secuenciar en ambas direcciones para observar la calidad de los cromatogramas. En todos los casos se obtuvieron cromatogramas con picos bien definidos y sin presentar el patrón anómalo anterior.

Posiciones de los cebadores con respecto a la secuencia

```
AACACTGGGGTGCAGGTGTATTTATGAAGTAGGGTTCCTTCTGGATATATGCCCAGGAGCGGGATTCC
TGGGTCATATGGTTAGTCTATTCCTAGTCTTTTGAGGAATCTCCATCCTGTTTTCCACAGTGGCTGTACC
AAACTGCATTCCCACCAGCAGTGTAGGAGGATTCCCCTTTCTCCACAGCCTCTCCAGCATTTGTTTTAAA
AACAGAATCATAAAAGGAAGTATTAATGTAAGTAAGAAATAGAAGTCAATTTTTCTCAGGCCTGGTGT
CCAGAATTTCAAGTTATTGGGTTTTGTCAATCATCTCATGAAGCTGAGTGGTACTTTTTGGCTTTTCATA
GTAAACCATCTCTCCTGTGAAAATAAGGGGATCCTTAAAGGAAACAAAACAAACATGAGTTTGAA
CAGCAGAGAAGAGATAGAAAAGCTCAGTGTAAAGGTAGGTTATCATAATTTGCTGAATGTCAGCCAAA
ATAACATATTATCATTGGTTGTTTTGGGTAGTTTGAATGTCTACTAAAAATGCTATCTTGCCATAATGG
TTGATTAGAAATGGATAATTTGAAGATCTTATATAAATAGAGTTTTGAATATGAGACAGTAGATTCTTAC
TAATGACTCAAAAATGATTCAGATGTGACTCAACATTAGCAACTGTCCTCAATACCCCTGAGGCCATGA
CTTAAAACCAGGAGTCAACAAATTACTTAGATTGGTCTGCAAATTAGCCAATGGTATGACTTCTGTGA
TGGTTCCCATTCACACTGAGGGTATATAAAAGGCTCTCTGTAGGGAAATTGTCCAGACTGAAGTTGCC
ATCCTCCTTCTTACCCAGAGACAACCCCAAAACCAACAGCAACACCATGTGCGACTACTACGGAAACT
ACTATGGGGGGCCGTGGCTATGGCTGCTGTGGCTACGGAGGCCTGGGCTGCGGCTATGGGTCTGCTA
TGGCTGTGGCTTCCGAGACTGGGCTGTGGCTATGGCTGTGGCTATGGCTGTGGC*TATGGCTTTGGC
TCCCGCTCTCTGTGGCTGTGGCTATGGGTGCGGCTCTGGCTATGGCTCTGGCTTTGGCTACTACTAT
TGAGGATGTCATGGAAGACTTCATCCCCTACACCTAGACCTCAGGATTCCCAATTCTGAACCCTTCTG
TTTTGTGCCTTTTCTCGGTAGTAGTTATCCTGCATCACGGAGGTCTAGTGTGGCCTGTTGTTGATTATC
ATCTAACCTACAAATTTCTGAATTTACCTCAAGAAGTGGGAGATGGTGAAATTCACCTGTGACCTCAG
AGTTCTGCTTTCATCAGGGTGTGGTGTCAAGCAACTGCCTTGACTTTCATGCTGAAGCTTGTTTCTCTG
TTGACCTAATAAATTATTCAATCTAAACCGAAACCTGTTTTTATGTCAGTTATTTGATGTTGAGTGT
CCCTACAGTTTTTTGTCTGTCTGTTTCAGTTTTCAAGAAAGCTAACAATTATGTTCTTTGTGAGATATTC
TCTCTTACTATTTTTCGGCTGGTTTCTTACCAATTCAAATATCAATTGTCTTAGGAAAGTATCCTTTTTT
TTTTTTTTTTT
```

Figura 31. Secuencia a partir de la cual se diseñaron ambos pares de cebadores y las posiciones de estos dentro de la misma. En amarillo el primer par de cebadores (6F y 6R), en rojo el segundo par de cebadores (6NF y 6NR), en negrita se muestra la región codante y un asterisco indica la posición de la inserción/delección.

4.2.4.3 KRTAP9-2

4.2.4.3.1 Diseño y prueba del primer par de cebadores (9F y 9R)

Como se mencionó anteriormente en el caso de KRTAP9.2 el genoma de referencia presentaba un gap en el extremo 5' de la región codante (Tabla 16) (Figura 34). Se desconocía el tamaño de este gap, aun así se diseñó un primer par de cebadores (9F y 9R) para amplificar y secuenciar esta porción. Se determinaron los valores de ΔG para la formación de estructuras secundarias, ninguno de los valores fue menor a -9 Kcal/mole (Tabla 16).

Tabla 16. Información de los cebadores diseñados para amplificar el gen KRTAP9-2.

Cebador	Secuencia	T _m (°C)	Tamaño amplicon	ΔG Hairpin (Kcal/mole)	ΔG Homodimero (Kcal/mole)	ΔG Heterpdimero (Kcal/mole)
9F	CGGAAATGTTGCTGGTGATG	59°C	?	-1.16	-3.61	-4.74
9R	GGTGGTCTGAGTAGCTTCATAG	59°C		0.55	-6.34	
9NF	AGTCCTGCTTATTGGTCCCA	60	707 pb	0.42	-5.02	-6.97
9NR	ACCCAAGAAGTGAATTTGCC	62		-0.1	-5.36	
9NNF	TGAAGCTACTCAGACCACCT	62	836 pb	0.15	-6.34	-4.67
9NNR	CCAGGAAGACAACCTCAAAGTGA	62		0.03	-3.53	

Para la amplificación de estos cebadores se utilizaron 3 muestras de ADN genómico de alpaca, una concentración de 2 mM de magnesio y se probaron 3 temperaturas distintas: 56°C, 58°C y 60°C. A 56° C una de las muestras no presento producto de amplificación, en el resto de casos se observó la presencia de productos. Se obtuvo un tamaño de fragmento entre 400 y 500 pb (Figura 32).

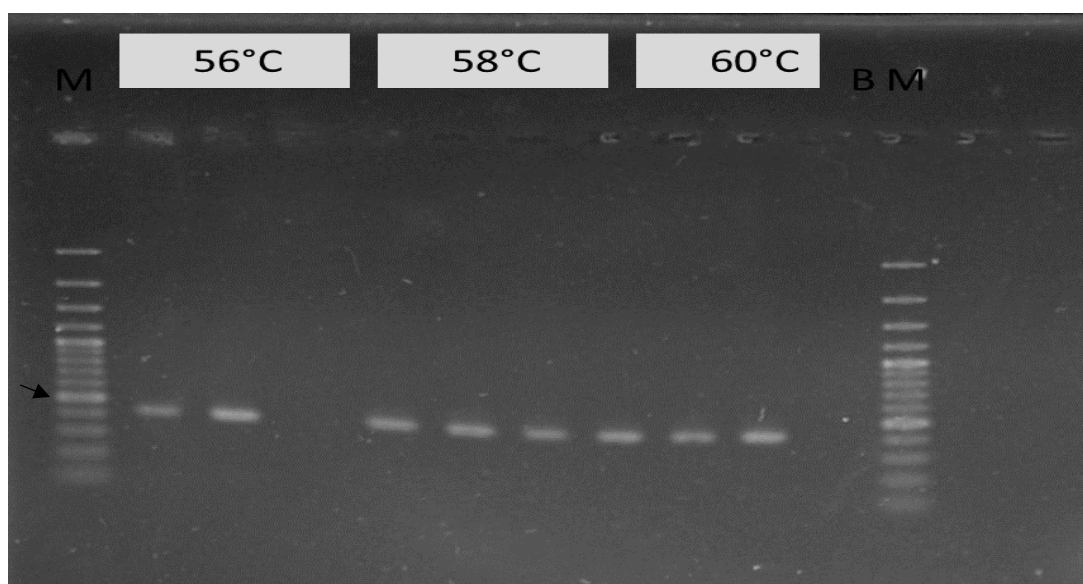


Figura 32. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 9F y 9R (KRTAP9-2) a diferentes temperaturas, se observan bandas entre 400 y 500 pb. El pozo con la M hace referencia al marcador de peso molecular (100bp), El último pozo contiene la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 500 pb. Se observa una distorsión en la corrida del gel a causa de un alto voltaje.

Se enviaron a secuenciar a Macrogen Inc. 3 productos de PCR a las condiciones a la cual se estandarizó la reacción. Se obtuvo una secuencia de 427 pb, el gap tenía una extensión de 291 pb que correspondían la parte faltante de la región codante como la región 5'UTR.

4.2.4.3.2 Diseño y prueba del segundo par de cebadores (9NF y 9NR)

El segundo par de cebadores (9NF y 9NR) que se diseñó para amplificar el gen KRTAP9-2 en alpaca, flanqueaba una región de 707 pb que contenía la región codante del gen (Tabla 16) (Figura 36). Para el diseño se contempló el ΔG para la formación de estructuras secundarias entre los cebadores, para ninguna de estas el valor fue menor a -9 kcal/mole (Tabla 16).

Posterior al diseño se realizó una curva de temperatura utilizando tres muestras de ADN genómico de Alpaca, con el fin de estandarizar la reacción utilizando estos cebadores. Las temperaturas que se evaluaron fueron: 60°C, 62°C y 64°C. Se utilizó una sola concentración de magnesio (2 mM). En todas las temperaturas probadas se observó la presencia de producto de reacción. Todos los productos presentaron un mismo tamaño que va de acuerdo a lo esperado al diseñar los cebadores (Figura 33).

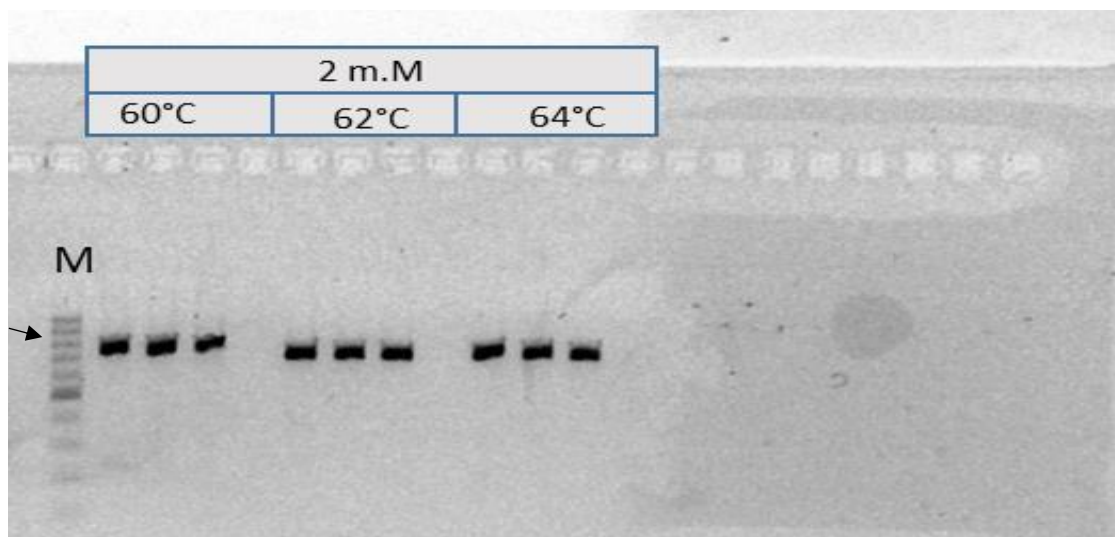


Figura 33. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 9NF y 9NR (KRTAP9-2) a diferentes temperaturas. El pozo con la M hace referencia al marcador de peso molecular (100bp), los espacios en blanco entre las muestras corresponden a las reacciones en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 800 pb.

Teniendo en cuenta como criterio la intensidad de banda se estandarizo la reacción a una concentración de 2 mM de magnesio y 64 °C (Tabla. 17).

Tabla 17. Condiciones a las cuales se estandarizaron las PCR para los cebadores diseñados en el gen KRTAP9-2.

Par de cebadores	Concentraciones		Condiciones Termicas		
	Reactivos	Concentracion final	Temperatura	Tiempo	Numero de ciclos
9NF y 9NR	ADN	20ng	94°C	4 minutos	1
	Buffer PCR	1X			
	dNTPs	0.2mM	94°C	30 segundos	30
	MgCl2	2 mM	62°C	45 segundos	
	Primer Forward	10 picomoles	72°C	50 segundos	
	Primer Reverse	10 picomoles	72°C	5 minutos	1
	TAQ polimerasa	0.5 U			
9NMF y 9NMR	ADN	20ng	94°C	4 minutos	1
	Buffer PCR	1X			
	dNTPs	0.2mM	94°C	30 segundos	30
	MgCl2	2 mM	62°C	45 segundos	
	Primer Forward	10 picomoles	72°C	55 segundos	
	Primer Reverse	10 picomoles	72°C	5 minutos	1
	TAQ polimerasa	0.5 U			

Se amplificaron 3 muestras utilizando las concentraciones a las que se estandarizo la reacción y se mandaron a secuenciar a la empresa Macrogen Inc. Al observar los cromatogramas se encontró uno que presentaba el patrón anómalo mencionado anteriormente (Figura 34). Se identificó que la inserción/delección es de 15 nucleótidos gracias a la presencia de un individuo homocigoto para esta. Este individuo con respecto al genoma de referencia presenta una delección de 15 nt en la posición 36 de la región codante del gen (Figura 36). Para resolver este problema se diseñaron nuevos cebadores que excluyeran la inserción/delección.

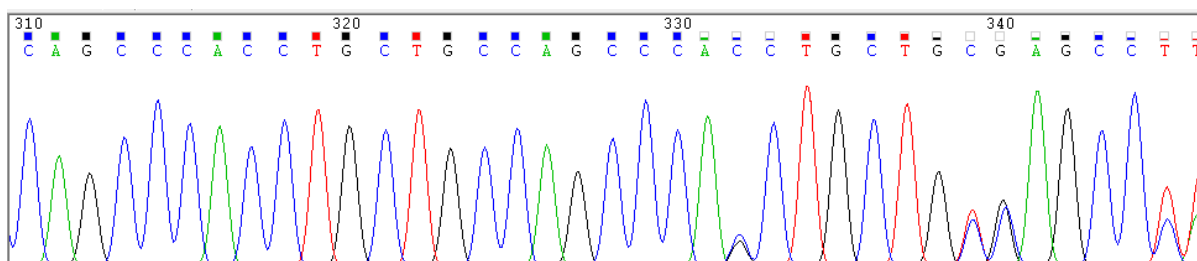


Figura 34. Cromatograma con patrón anómalo de la secuencia de KRTAP9-2 obtenido con los cebadores 9NF y 9NR. Se observa un gran número de picos superpuestos, esto se debe que el cromatograma corresponde a un individuo heterocigoto para la inserción/delección.

4.2.4.3.3 Diseño y prueba del tercer par de cebadores (9NNF y 9NNR)

El tercer par de cebadores (9NNF y 9NNR) se diseñó con el propósito de amplificar parte de la región codante y la región 3'UTR sin incluir la inserción/delección, el tamaño del amplicon es de 836 pb (tabla 23) (Figura 36). Se evaluó el ΔG para la formación de estructuras secundarias ninguno de los valores fue menor a -9 Kcal/mole (tabla 16).

Posterior al diseño de los cebadores se realizó una curva de temperatura y de magnesio utilizando tres muestras de ADN genómico de alpaca. Las temperaturas que se evaluaron fueron: 60°C, 62°C y 64°C. Las concentraciones de MgCl₂ que se probaron fueron: 1.5mM y 2mM (Figura 35). Se observó producto de PCR para todas las temperaturas y

concentraciones de magnesio que se probaron. Todos los productos presentaron un mismo tamaño que va de acuerdo a lo esperado al diseñar los cebadores (Figura 35).

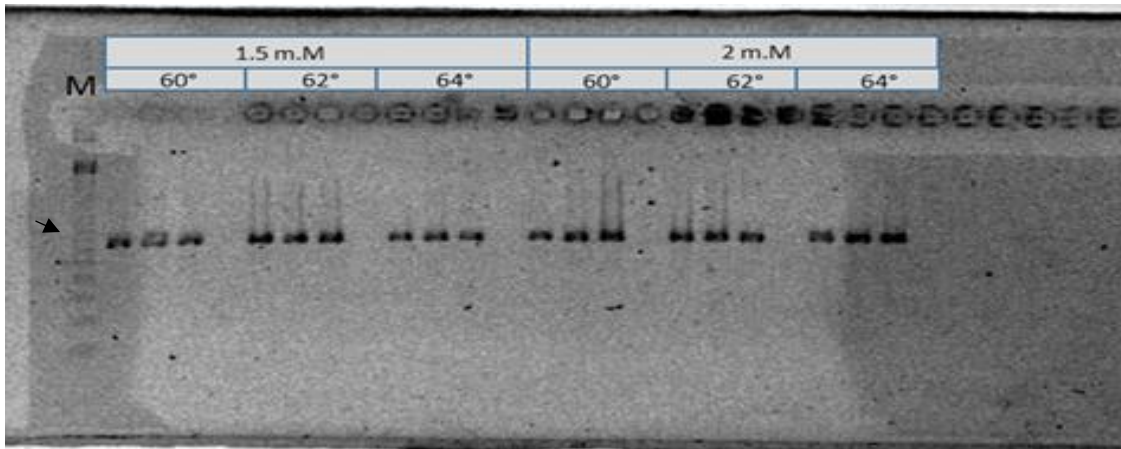


Figura 35. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 9NNF y 9NNR a diferentes concentraciones de magnesio y temperatura. El pozo con la M hace referencia al marcador de peso molecular (123bp), el pozo en blanco entre muestras corresponde a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 861 pb.

De acuerdo a la intensidad de banda se estandarizó la reacción de PCR a la temperatura de *annealing* de 64°C y a una concentración de MgCl₂ de 2 mM (Tabla 17). Se enviaron 3 muestras a secuenciar en ambas direcciones para observar la calidad de los cromatogramas. En todos los casos se obtuvieron cromatogramas con picos bien definidos y sin presentar el patrón anómalo anterior.

Posiciones de los cebadores con respecto a la secuencia

CGGAAATGTTGCTGGTGATGATTCTGACTGTTCCATTTCTGGAAGAGGTTCTTAAGTTGTAGTCACAA
 TTAAGAAAATGATCAAACATTAATTTACAGAGGTATCTCTATAAAATGCCAAGAGAACAACCTCCTCCAT
 AACAAATTACCTAATGACTCCAAAAACAACGTAACAACAACCAGGAAAGTCTGCTTATTGGTCCCAAC
 ACTGGGGGTCCAGTGTATAAAAGCCCCAGAACAGAAGGAGTGATCAGAATCTGAGAACTCACCACCT
 AACAGGAGTTCCAACCTCCACCCTGACACCATGACCCACTCATGCTGCTCCCTTGCTGCCAGCCCACC
 TGCTGCCAGCCCACCTGCTGTGAGCCCAGCTGCTGCCAGCCTTGCTGCCCCAAACTTGCTATGAAGCT
 ACTCAGACCCTGCTGCGAGCACCACTGCCGCAACCTACTTGTGTGACCACCTGCTGCCAGCCCACC
 TGCTGTGAGCCCAGCAGCTGTGGACAAACCTGCTGCCAGCCTACCTGTGTAACCAGCTGTTGCAGCCC
 TTCTGCTGCCAGTCCACTTGCTGTGAGTCTAGCAGCTGCGGACAAAACCTGCGGTGGGTCCAGCTGCT
 GCCAGCCAGCTAGCTGTGCACCCGTGTACTGCCACCGAACCTGGTACCACCCACGTGCTGCTGCCTG
 CCTGGGTGCCTGGCCAGAGCTGTGGATCCAGCTGCTGCCAGCCTCGCTGCCGCCCTGTCTGCTGTCA
 GACCACGTGCTGCCGCCCTAGCTGTGTGTCCAGCTGCTGCCAGCCTCCTGCTGATCACCTACCAAGA
 GCCATCCCCTGCATCCAACACAATCTGTCAACTGAGTTGCCGTTTTGGGGCAAATTCACCTTCTTGGGT
 TGCCAATTACCAGCATGCTCTCCAGAACTTCTGTTACTCATCTGCTTGTTAAAGCTTGTGAATCAGCTT
 GTTGGAGTGAGAGGACTTCCCCAATTCTTCTCTCTCGTGTAGGTGGATGATGAGCCAGCTTTGTGC
 ATCCTTGATATGGAGTTTGTCTGGGCTTGACTCTGAAGTCAGGTCCACGCAGTCCCTCTGAGTGACTA
 AGGAAAATAATTCTGTAAGTGTTCACAGTTTCTGCAACTGATCAATATTATTCATAATCATATTTTC
 TTTATCAATGTAAGTCTGCTGCTCTCACAGCTTTCTTGTCTTTTATGATCACTTTGAGTTGCTTCTCTGG
 ATACAGGAGTCTTCCCTATGTGTCTCCCAATAAATCCGTATCATAATTTATCCCATATTGTGTTTGATTAT
 TGCACAGC

Figura 36. Secuencia a partir de la cual se diseñaron los tres pares de cebadores y las posiciones de estos dentro de la misma. En verde el primer par de cebadores (9F y 9R), en amarillo el segundo par de cebadores (9NF y 9NR), en rojo el tercer par de cebadores (9NNF y 9NNR) y un asterisco indica la posición de la inserción/delección.

4.2.4.4 KRTAP11-1

Diseño de cebadores

Para la amplificación del gen KRTAP11-1 se diseñó un par de cebadores que flanquean una secuencia de 1145 pb (Tabla 18) (Figura 38). Este fragmento contiene la región codante y parte de las regiones 5'UTR y 3'UTR. Al diseñar los cebadores ninguno presento valores de ΔG para la formación de estructuras secundarias menores a -9 Kcal/mole (Tabla 18).

Tabla 18. Información del par de cebadores (11F y 11R) diseñados para amplificar el gen KRTAP11-1.

Cebador	Secuencia	Tm (°C)	Tamaño amplicon	ΔG Hairpin (Kcal/mole)	ΔG Homodimero (Kcal/mole)	ΔG Heterpdimero (Kcal/mole)
11F	TTGAAGCTGTCAACCAACCTT	61°C	1145 pb	1.25	-6.34	-4.89
11R	CACAGAGCAAGACAAGTTCCC	62°C		1.25	-3.14	

Posterior al diseño se realizó una curva de temperatura y de magnesio utilizando tres muestras de ADN genómico de alpaca. Las temperaturas que se evaluaron fueron: 58°C, 60°C y 62°C. Las concentraciones de MgCl₂ que se probaron fueron: 1mM, 1.5mM y 2mM (Figura 37). A la concentración de 1 mM de magnesio no se observó producto de PCR a ninguna temperatura, mientras que si hubo producto para el resto de concentraciones. Se observó producto de PCR para todas las temperaturas que se probaron. Todos los productos presentaron un mismo tamaño que va de acuerdo a lo esperado al diseñar los cebadores (Figura 37).

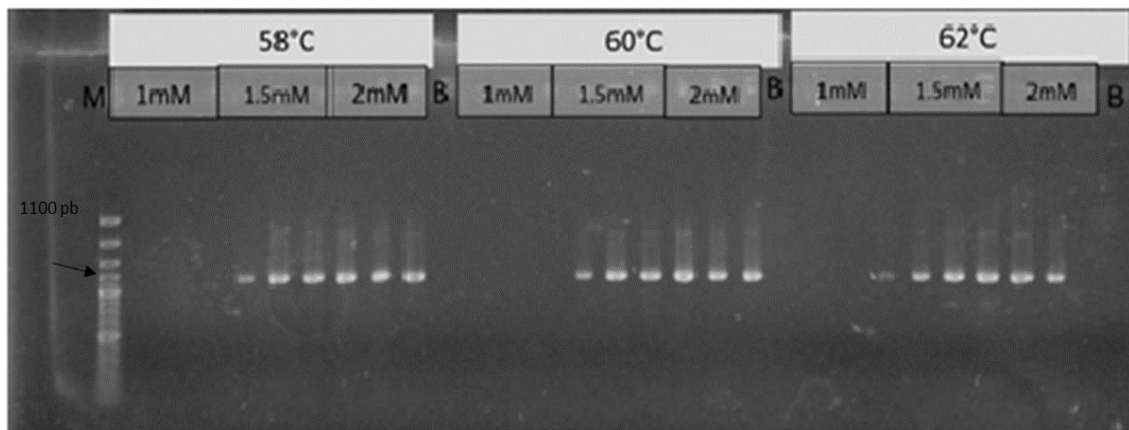


Figura 37. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 11F y 11R (KRTAP11-1) a diferentes concentraciones de magnesio y temperatura. El pozo con la M hace referencia al marcador de peso molecular (100bp plus), El pozo con la B hace referencia a la reacción en blanco y la flecha negra indica la banda del marcador de peso molecular que corresponde a 1100 pb.

De acuerdo a la intensidad de banda se estandarizó la reacción de PCR a la temperatura de *annealing* de 60°C y a una concentración de MgCl₂ de 2 mM (Tabla 19). Se enviaron 3 muestras a secuenciar en ambas direcciones para observar la calidad de los cromatogramas. En todos los casos se obtuvieron cromatogramas con picos bien definidos.

Tabla 19. Condiciones a la que se estandarizó la PCR para los cebadores 11F y 11R.

Par de cebadores	Concentraciones		Condiciones Termicas		
	Reactivos	Concentracion final	Temperatura	Tiempo	Numero de ciclos
11F y 11R	ADN	20ng	94°C	4 minutos	1
	Buffer PCR	1X			
	dNTPs	0.2mM	94°C	30 segundos	30
	MgCl2	2 mM	62°C	45 segundos	
	Primer Forward	10 picomoles	72°C	1: 15 minutos	
	Primer Reverse	10 picomoles	72°C	5 minutos	1
	TAQ polimerasa	0.5 U			

Posiciones de los cebadores con respecto a la secuencia

```

CTGACAAAACCCACACAGGGTGTGGGTGTGTAATTAGTCCTGGTCAGAAGGGAGAGATGGTAGGAAG
GGGGCATGTTTGGCGTTACTGTCAGATCGACACTCATTGTATCCAAGCAAGATGAATGAGAGATGAAT
TGTGGGCTGGGATGACCAATAGCTTTGAAGCTGTCAACCAACCTTCTTCATCTGGGACTAAGTACAGT
ATGGAAGCCTTATGATTTAGTTACCAAACACTAAACTTCCTTTATGAGACTGAAGGCTTCTCAAGAGGC
AATTAACCTGGAGAGTCACAAGATGTTCTAAGGCAAGCTGGTCTGAGCCAAAGAAGCCAGGAGGACC
AGGGTGGGAGCCCCACCCAGCAAGGGTATATAAAGGCTCAGGAGCCTGAGGTGGCATTACAAT
CCAGGAGCTAGCTTACGCGAGTTACGCACATCTGTCTGTCTGGCAACATGTCCTACAACCTGCTCCACAA
GGAATTGCTCTCCAGGCTGATTGGGGGACAATACTCTGTCCCCGTGACCCCTGTTGTCACAGCTTCTA
CCCAGGATGCTGACTGCCTGAGTGGCATCCATTTGCCAGCTCCTTCCAAACTGGCTCCTGGCTCCTGG
ACCACTGTCAGGAGACCTGCTGTGAGCCACTGTTTGCCAGCCAACGTGTTACCAGCAAACCTTCTTGT
GTCTCCAGCCCTGTCCAGGTGACCTGCTCTCGACAAACCAACTGTGTCTCTAATCCCTGCTCAACTGCC
TACAGCCGGCCGCTCTCCTTTGTCTCCAGTGGCTGTGGCCCTGGGCGGCATCTCTACTGTGTGCCAA
CCAGTGAGAAGCGTCTCACTGTCTGCCGGCCAGTGGGAGGAGTCTCCACCATCTGCCAGCCAGCCT
GCGGGGTCTCCAGGACGTACCAGCAGTCTTGCGTGTCCAGCTGCCGAAGAAGTCTAAGTGTGCGG
GAGCCAGTGAGCGAATCAAGCCTCTATGACCTGTCAGCTGTGTTTCCAGGATCTTCCAACATGCTGTCT
GTCCCTGAAGAATTCTTATTGCTGACTGCTATTCTAACTGCCTGATTGCTGGCTGCCAGCCCTGAATAA
GCTGCCTTTGGCCATCTAAGTGTTCCTGGCCAGCACAATCTTATTTAAGGGTTTGTACCCGGTGGC
ACGTATACCTCTGGATGTTTCCAGAAATTTACCACCCACGCGCAAGTCTCTAATGGTTTTGGCATGTCTT
GACCTTTCTGCTTTGTCTGGCTTCTGGCTTCTGTTTTGTGCCTGTGAAAAAGGGAACCTGTCTCGCTCTGT
GTTTCTCAATAAACCTTCACTTGGCATTGCAATGTGTGTCTCAACAGAGTCTTTATTTGCAAGAAT
TTTTGCTATCTATAAATCATCTTTGCTGATCAGGGTCTGATATT
    
```

Figura 38. Secuencia a partir de la cual se diseñó el par de cebadores y las posiciones de estos dentro de la misma. En amarillo la posición de los cebadores 11F y 11R.

4.2.4.5 KRTAP13-1

Diseño de cebadores

Para la amplificación del gen KRTAP13-1 se diseñó un par de cebadores que flanquean una secuencia de 862 pb (Tabla 20) (Figura 40). Este fragmento contiene la región

codante y parte de las regiones 5'UTR y 3'UTR. Al diseñar los cebadores ninguno presento valores de ΔG para la formación de estructuras secundarias menores a -9 Kcal/mole (Tabla 20).

Tabla 20. Información del par de cebadores (13F y 13R) diseñados para amplificar el gen KRTAP13-1.

Cebador	Secuencia	Tm (°C)	Tamaño amplicon	ΔG Hairpin (Kcal/mole)	ΔG Homodimero (Kcal/mole)	ΔG Heterpdimero (Kcal/mole)
13F	TGAGTCACGGGGATGACATC	62°C	862 pb	-2.22	-5	6.82
13R	TGTGCTTCAGTGTCACATGTC	61°C		-0.91	-7.07	

Posterior al diseño se realizó una curva de temperatura y de magnesio utilizando tres muestras de ADN genómico de alpaca. Las temperaturas que se evaluaron fueron: 58°C, 60°C y 62°C. Las concentraciones de MgCl₂ fueron: 1mM, 1.5mM y 2mM (Figura 39). A la concentración de 1 mM de magnesio no se observó producto de PCR a ninguna temperatura, mientras que si hubo producto para el resto de concentraciones. Se observó producto de PCR para todas las temperaturas que se probaron. Todos los productos presentaron un mismo tamaño que va de acuerdo a lo esperado al diseñar los cebadores (Figura 39).

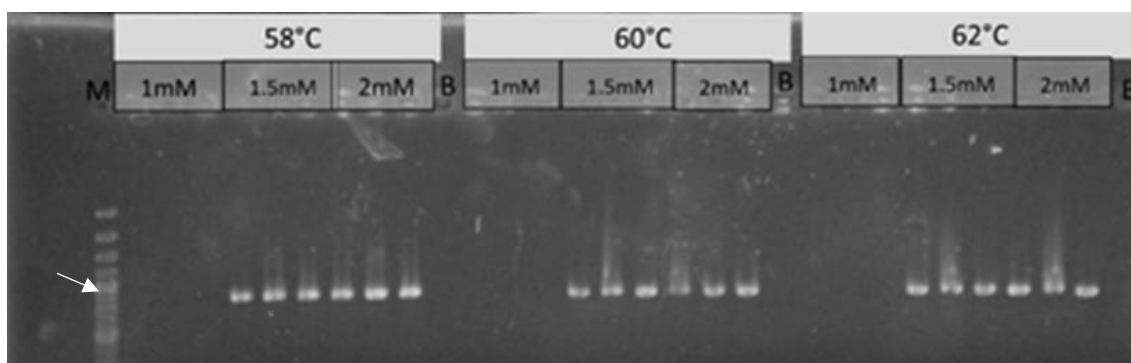


Figura 39. Gel de agarosa al 1%. Se muestran los productos de PCR utilizando los cebadores 13F y 13R (KRTAP13-1) a diferentes concentraciones de magnesio y temperatura. El pozo con la M hace referencia al marcador de peso molecular (100bp plus), El pozo con la B hace referencia a la reacción en blanco y la flecha blanca indica la banda del marcador de peso molecular que corresponde a 1000 pb.

De acuerdo a la intensidad de banda se estandarizo la reacción de PCR a la temperatura de *annealing* de 62°C y a una concentración de MgCl₂ de 2 mM (Tabla 21). Se enviaron 3 muestras a secuenciar en ambas direcciones para observar la calidad de los cromatogramas. En todos los casos se obtuvieron cromatogramas con picos bien definidos.

Tabla 21. Condiciones a la que se estandarizó la PCR para los cebadores 13F y 13R.

Par de cebadores	Concentraciones		Condiciones Termicas		
	Reactivos	Concentracion final	Temperatura	Tiempo	Numero de ciclos
13F y 13R	ADN	20ng	94°C	4 minutos	1
	Buffer PCR	1X			
	dNTPs	0.2mM	94°C	30 segundos	30
	MgCl ₂	2 mM	62°C	45 segundos	
	Primer Forward	10 picomoles	72°C	55 segundos	
	Primer Reverse	10 picomoles	72°C	5 minutos	1
	TAQ polimerasa	0.5 U			

Posiciones de los cebadores con respecto a la secuencia

```

AGTGAGTTAGCAACTTTGCTATCCAGGTGGCAATAATTTTGAGACAGATTGTGAAAGAGATATCAATTA
TTTTAATTAAGCCAGCTTCCCTCAGCCCAGACATTGTCTACCTCTAATGTAGTTCTCTGAGTGATCTGCT
CGGTGACCTCTGGTAGAGAAAAGATGACTAACAGCTTTGAAGATTCAAACTATTACCTGAATTCCAAGG
CACAAACATGACTGCACCTTCCAGTTATAAAACCACAGAACTAAGTATGCATAACCATGCTGCTGTCTT
CAATTCAATTAAGGAAGATAACTTGTTTTTATGAGTCACGGGGATGACATCCCCAGTGAAAGA
GTATATAAACGTCCCAGTCCAAGAAGTGACATTCAAACTCAGATCCTTCTCACAGTAACTCAGCTGAACT
CACACCTCCTGTACCATGTCTACAACCTGCTGCTCTGGAACTTCTCCTCTCGCTCCTTTGGGGCCAC
CTGCGTGCCGAGGCTCCTCCTGTGGCTTCTCTACCCAGCAACCTGGTCTACACCACGGACCTCTGC
TCTCCAGCACCTGCCAGCAGGACTCCTCTCTCCACAGCGGCTGTCAGGAGATCTGCTCTGAGCCTATC
AGGTGCCAGACGTTCCAGGTGGAGTCCAGTCCCTGCCAGTCATCCTGCTACCGCTGGAAGACCTCCAC
GCTCTGCCGTCCCTGGCAGACGACTTACTCTGGGTCTCTGGGCTTTGGCTCCAGAGTTTTCAATCTTT
TGGTTGTGGCTTCCCGTCTCTGGGCTTTGGATCCAGTGGTTTCCAATCAGTGGGTTGTGGCCCAGGG
CTTTCTCATCTGTAAGATGGAGATCCAGCTTTGACAGTCCAACTACTTCTTCTTAGGAGCTGGCAGC
CTGCTTCTTTCCAACCAATCTTTAGATCTGGCTTCTTCTAATCTTATGCTTGAAAACAACACCGACTGTAT
CTAGAAATTTGATGCAATATCTCCTATATCAAGATCTATACTATCTATTGGTCTCCTGTCTACTGTTAGC
TTACAAATTTGACTTCTAGTGAATCTATTATAGATTGGGAATAATTTGAATACCAGTAACTGGAATGC
AATCAGTAAAGAAAGATGTCAAACTTCTTCTTGTACTTACACA GACATGTGACACTGAAGCACA TTT
ATTTCTATTTCTTCTCCTACTTATTTCCCTAATGTCTCTATTTCAAATTTTACTAGTAGAAAAAATGT
AAGTTTAAATAAAAATATGTCATATGATATCCACATAATTGCTTAAATATCATTTTGTCTGTGTCTT
CAAAATGACTTCAATTATCTGGCTAGATTTCCATATCTTGTATTCACTGGAAAAAATTTATTTAA
AAAGTAAAATCTTCAGTATCATTGAGGTTTGGGAAGAAATCATTTTCATCTTTCAT

```

Figura 40. Secuencia a partir de la cual se diseñó el par de cebadores y las posiciones de estos dentro de la misma. En amarillo la posición de los cebadores 13F y 13R.

4.2.5 Amplificación y secuenciamiento de las muestras para validación de marcadores

Para la validación de los marcadores se utilizaron 96 muestras de ADN genómico de alpaca huacaya provenientes de distintas regiones del Perú (Tabla 8). Se amplificaron las 96 muestras para cada uno de los 5 genes seleccionados (KRTAP1-2, KRTAP6-1, KRTAP9-2, KRTAP11-1 y KRTAP13-1). Por gen solo se utilizó uno de los pares de cebadores diseñados a las condiciones estandarizadas previamente (Tabla 22). Una vez amplificadas las muestras se comprobó en geles de agarosa al 1% la presencia de producto para posteriormente mandarlas a purificar y secuenciar a la empresa MacroGen Inc.

Tabla 22. Selección de cebadores para la validación de los marcadores en los 5 genes.

Gen	Par de cebadores utilizados	Condiciones
KRTAP1.2	1NF	Tabla 13
	1NR	
KRTAP6.1	6NF	Tabla 15
	6NR	
KRTAP9.2	9NNF	Tabla 17
	9NNR	
KRTAP11.1	11F	Tabla 19
	11R	
KRTAP13.1	13F	Tabla 21
	13R	

4.2.6 Identificación de sitios polimórficos y validación de marcadores

Una vez obtenidas las secuencias para cada uno de los genes, estas fueron alineadas utilizando el Software Mega v.7 para identificar los sitios polimórficos. De las 96 muestras se alinearon 40 muestras seleccionadas al azar. Una vez identificadas las posiciones polimórficas se inspeccionaron los cromatogramas en estos sitios de manera manual para determinar los genotipos en cada una de las posiciones. Los sitios polimórficos fueron nombrados de acuerdo a su posición dentro del fragmento amplificado. A partir de este listado se calcularon las frecuencias alélicas, los posibles

haplotipos por gen, índices de variabilidad, si los marcadores se encuentran en equilibrio de Hardy-Weinberg (EHW) y que pares se encuentran en desequilibrio de ligamiento (LD). En total se identificaron 32 posiciones polimórficas (Anexo 4), a continuación se detalla esta información para cada uno de los genes.

4.2.6.1 KRTAP1-2

Al realizar el alineamiento de las 40 secuencias (Forward y Reverse) del fragmento de KRTAP1-2 se identificaron 10 sitios polimórficos. Por lo tanto mediante inspección de los cromatogramas se determinó el genotipo en estos 10 sitios para las 96 muestras. Una de las secuencias no se secuenció correctamente y fue descartada del análisis. En cada una de las 10 posiciones se encontraron dos alelos y en cada caso se pudo encontrar los tres genotipos posibles (homocigotos para ambos alelos y heterocigotos) (Tabla 23). Cinco de las posiciones polimórficas se originaron por una mutación de transición (158, 370, 639, 705 y 814) mientras que las 5 restantes se originaron por mutaciones de transversion (Tabla 23). Al determinar las frecuencias alélicas en ninguno de los casos estas fueron menores al 5% (Tabla 23). Por lo tanto los 10 sitios presentan frecuencias alélicas adecuadas para ser marcadores genéticos.

Tabla 23. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 10 sitios polimórficos identificados en el fragmento del gen KRTAP1-2.

SNP	Alelos		Conteo de individuos			Frecuencias alélicas	
	Alelo1	Alelo2	Homocigotos1	Homocigotos2	Heterocigotos	Alelo1	Alelo2
158	C	T	80	1	14	0.916	0.084
224	C	A	20	26	49	0.468	0.532
246	G	T	54	4	37	0.763	0.237
262	A	C	54	4	37	0.763	0.237
370	A	G	4	54	37	0.237	0.763
639	A	G	60	5	30	0.789	0.211
687	A	T	4	54	37	0.237	0.763
705	A	G	4	56	35	0.226	0.774
721	A	T	1	69	25	0.142	0.858
814	A	G	18	22	55	0.479	0.521

A partir del listado de genotipos se determinó los posibles haplotipos para el fragmento amplificado tomando en consideración a los 10 loci. En total se encontraron 8 haplotipos, los cuales presentan distintas frecuencias haplotípicas (Tabla 24). Los más frecuentes son los haplotipos 3, 5, 7 y 8; que juntos tienen un frecuencia de 82%.

Tabla 24. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP1-2.

Haplotipos	1	2	3	4	5	6	7	8	9	10	Frecuencias
1	C	A	G	A	G	A	T	G	T	A	0.032
2	C	A	G	A	G	A	T	G	T	G	0.053
3	C	A	G	A	G	G	T	G	A	A	0.142
4	C	A	G	A	G	G	T	G	T	A	0.069
5	C	A	T	C	A	A	A	A	T	A	0.142
6	C	A	T	C	A	A	A	G	T	A	0.011
7	C	C	G	A	G	A	T	G	T	G	0.468
8	T	A	T	C	A	A	A	A	T	A	0.084

Designación de las posiciones del 1 al 10 de acuerdo al orden en que aparecen en el fragmento amplificado.

Se determinó cuáles de estos polimorfismos se ubican en la región codante y cuales en la región reguladora 3´UTR. De los diez polimorfismos 3 (158, 224 y 246) se ubican en la región codante mientras que los otros 7 se encuentran dispersos en la región 3´UTR (Tabla 25). Se evaluó que tipo de cambio producían en la proteína aquellas posiciones que se ubican en la región codante, observándose que en los tres casos se dan cambios no sinónimos (Tabla 25). De los tres cambios que se dan, el que probablemente tenga un impacto más importante en la función de la proteína es el de la posición 246 que corresponde a un cambio de una cisteína a una fenilalanina (Tabla 25).

Tabla 25. Ubicación de los sitios polimórficos dentro del gen KRTAP1-2 y los cambios en la secuencia de nucleótidos y de aminoácidos.

SNP	Region	Cambio en codon	Cambio en proteína
158	Codante	CTC <--> TTC	L <--> F
224	Codante	CTC <--> ATC	L <--> I
246	Codante	TGT <--> TTT	C <--> F
262	3'UTR	-	-
370	3'UTR	-	-
639	3'UTR	-	-
687	3'UTR	-	-
705	3'UTR	-	-
721	3'UTR	-	-
814	3'UTR	-	-

En cuanto a los valores de heterocigosidad con excepción de la posición 158, todas presentaron valores intermedios a altos (0.3 a 0.5) (Tabla 26). El marcador 158 presentó valores por debajo de 0.2 (Tabla 26). Estos valores dependen exclusivamente del número de heterocigotos presentes por marcador. En cuanto al PIC, los valores se encuentran entre 0.15 y 0.35, son bajos en comparación a otro tipo de marcadores como por ejemplo microsatelites. Sin embargo para marcadores SNP estos valores son altos. Por otro lado todos los marcadores se encuentran en equilibrio de Hardy-Weinberg (Tabla 26). Por lo tanto se pudo validar 10 marcadores del tipo SNP en el fragmento del gen KRTAP1-2 en alpaca.

Tabla 26. Índices de variabilidad genética para cada loci en el gen KRTAP1-2 así como la evaluación del equilibrio de Hardy – Weinberg.

SNP	Hobs	Hesp	PIC	HW p-Value (EH)	HW p- Value (DH)
158	0.147	0.155	0.142	0.879	0.498
224	0.516	0.501	0.374	0.464	0.693
246	0.389	0.363	0.296	0.346	0.84
262	0.389	0.363	0.296	0.346	0.84
370	0.389	0.363	0.296	0.346	0.84
639	0.316	0.334	0.277	0.806	0.395
687	0.389	0.363	0.296	0.346	0.84
705	0.368	0.352	0.289	0.447	0.77
721	0.263	0.245	0.214	0.41	0.884
814	0.579	0.502	0.375	0.096	0.957

Hobs: Heterocigosidad observada

Hesp: Heterocigosidad esperada

HW (EH): Equilibrio de Hardy- Weinberg asumiendo exceso de heterocigotos

HW (DH): Equilibrio de Hardy- Weinberg asumiendo déficit de heterocigotos

También se quiso determinar las variantes a nivel proteico tanto a causa de las sustituciones nucleotídicas como también a causa de la inserción/delección mencionada anteriormente. En cuanto a las variantes a causa de la sustitución nucleotídica se trabajó con los haplotipos identificados en este fragmento (Tabla 24). Cada haplotipo fue traducido a una proteína donde se obtuvieron las distintas variantes y sus frecuencias (Tabla 27). Se identificaron 4 variantes distintas. Las variantes que presenta un residuo de cisteína en la posición 126 tiene una frecuencia aproximada de 0.75 mayor que la frecuencia de las variante con una fenilalanina en la misma posición (Tabla 27). En cuanto a la inserción/delección la variante que presenta 30 nucleótidos más, produce una proteína 10 residuos más grande sin que se corra el marco de lectura. Tres de estos 10 residuos son de cisteína, lo cual podría ser un cambio importante.

Tabla 27. Variantes a nivel proteico de acuerdo a los haplotipos identificados para la proteína KAP1-2.

Variante	haplotipo en gen	Residuo 97	Residuo 119	Residuo 126	Frecuencia
1	1, 2, 3, 4	L	I	C	0.296
2	5, 6	L	I	F	0.152
3	7	L	L	C	0.468
4	8	F	I	F	0.084

Por último se determinó si los marcadores encontrados en el fragmento del gen KRTAP1-2 se encontraban en desequilibrio de ligamiento. Todos los marcadores validados en este fragmento se encuentran en desequilibrio de ligamiento entre ellos ($p < 0.0001$) (Tabla 28).

Tabla 28. Determinación de desequilibrio de ligamiento entre los 10 marcadores del fragmento del gen KRTAP1-2.

Posicion	158	224	246	262	370	639	687	705	721	814
158	-	0	0	0	0	0.0001	0	0	0	0
224	*	-	0	0	0	0	0	0	0	0
246	*	*	-	0	0	0	0	0	0	0
262	*	*	*	-	0	0	0	0	0	0
370	*	*	*	*	-	0	0	0	0	0
639	*	*	*	*	*	-	0	0	0	0
687	*	*	*	*	*	*	-	0	0	0
705	*	*	*	*	*	*	*	-	0	0
721	*	*	*	*	*	*	*	*	-	0
814	*	*	*	*	*	*	*	*	*	-

Encima de la diagonal se muestran los p-values
 Debajo de la diagonal si es significativo (*)

4.2.6.2 KRTAP6-1

Al realizar el alineamiento de las 40 secuencias seleccionadas al azar se encontraron 9 sitios polimórficos. Estos 9 sitios polimórficos fueron inspeccionados en los 96 cromatogramas para identificar los genotipos en cada uno de estos sitios. Este gen fue el que mostró mayor cantidad de errores en el secuenciamiento donde las posiciones 77 y 110 del amplicon solo pudieron ser genotificadas para 88 individuos, mientras que el

resto tan solo 93 veces (Tabla 29). En cada una de las 9 posiciones se encontraron dos alelos, sin embargo no en todo los casos se observaron los tres genotipos posibles. Para la posición 77 no se encontraron homocigotos AA como tampoco se observaron homocigotos CC en la posición 506. En el resto de posiciones si se observaron homocigotos para ambos alelos y heterocigotos. Cinco de las posiciones polimórficas se originaron por una mutación de transición (110, 502, 557, 612 y 697) mientras que las 4 restantes se originaron por mutaciones de transversion (Tabla 29). La posición 77 es la única donde uno de los alelos presenta una frecuencia menor al 5% (3.4%), esto quiere decir que solo 8 posiciones en este gen tienen frecuencias alélicas adecuadas para ser marcadores genéticos (Tabla 29).

Tabla 29. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 9 sitios polimórficos identificados en el fragmento del gen KRTAP6-1.

SNP	Alelos		Conteo de individuos			Frecuencias alelicas	
	Alelo1	Alelo2	Homocigotos1	Homocigotos2	Heterocigotos	Alelo1	Alelo2
77	A	T	0	82	6	0.034	0.966
110	C	T	71	14	3	0.824	0.176
212	A	C	1	69	22	0.130	0.870
369	A	C	58	6	29	0.780	0.220
502	C	T	20	51	22	0.333	0.667
506	C	G	0	82	11	0.059	0.941
557	A	G	20	49	24	0.344	0.656
612	A	G	77	2	14	0.903	0.097
697	C	T	82	1	10	0.935	0.065

A partir del listado de genotipos se determinó los posibles haplotipos para el fragmento amplificado tomando en consideración a los 9 loci. En total se encontraron 12 haplotipos (Tabla 30) con frecuencias variables, sin embargo los más representados fueron el haplotipo 7 seguido por los haplotipos 12, 5 y 4 que sumados tienen una frecuencia de aproximadamente 80% (Tabla 30). Es importante mencionar que este fragmento es el que presentó un mayor número de haplotipos con frecuencias muy bajas como por ejemplo los haplotipos 3 y 8 que aparecen en un solo individuo cada uno (Tabla 30). Esto se debe

a que gran parte del fragmento amplificado se encuentra *upstream* de la región 5'UTR del gen (Figura 31).

Tabla 30. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP6-1.

Haplotipos	1	2	3	4	5	6	7	8	9	Frecuencias
1	A	C	C	A	T	G	A	G	C	0.01
2	A	C	C	A	T	G	G	G	C	0.035
3	A	C	C	C	C	G	A	G	T	0.007
4	T	C	A	C	C	G	A	A	C	0.123
5	T	C	C	A	C	G	A	A	C	0.113
6	T	C	C	A	T	C	G	G	C	0.056
7	T	C	C	A	T	G	G	A	C	0.37
8	T	C	C	C	C	C	A	G	T	0.008
9	T	C	C	C	C	G	A	A	C	0.021
10	T	C	C	C	C	G	A	A	T	0.059
11	T	C	C	C	T	G	G	A	C	0.026
12	T	T	C	A	T	G	G	A	C	0.172

Designación de las posiciones del 1 al 9 de acuerdo al orden en que aparecen en el fragmento amplificado.

Al determinar la posición de los marcadores dentro del gen, la mayoría de estos se ubican probablemente *upstream* de la región promotora (tomando la caja TATA como el inicio). Ninguno de los marcadores se ubicó *downstream* de la caja TATA ni en la región codante del gen. Las posiciones polimórficas dentro de este gen presentan valores bajos a intermedios de heterocigosidad observada (0.03 - 0.312) (Tabla 31). La posición que tiene el valor mas alto es la 369 (0.312) mientras que la que presenta el valor más pequeño es la posición 110 (0.03) (Tabla 31). En cuanto a los valores de PIC, estos fueron bajos en comparación a los observados en otros genes (Tabla 31). Tres de las 8 posiciones (110, 502 y 557) no se encuentran en equilibrio de Hardy – Weinberg, en todos los casos se debe a un déficit de heterocigotos (Tabla 31). Teniendo esto en consideración además del numero alto de errores en el secuenciamiento es posible que en estas posiciones existan alelos nulos, por lo tanto estas posiciones no fueron consideradas como marcadores

genéticos. Se validaron 5 marcadores del tipo SNP en el fragmento del gen KRTAP6-1 en alpaca.

Tabla 31. Índices de variabilidad genética para cada loci en el gen KRTAP6-1 así como la evaluación del equilibrio de Hardy – Weinberg.

SNP	Hobs	Hesp	PIC	HW p-Value (EH)	HW p- Value (DH)
77	0.068	0.066	0.064	0.916	0.084
110	0.034	0.292	0.248	1	0
212	0.239	0.228	0.201	0.534	0.822
369	0.312	0.346	0.285	0.893	0.252
502	0.237	0.447	0.346	1	0
506	0.118	0.112	0.105	0.731	0.269
557	0.258	0.454	0.35	1	0
612	0.151	0.176	0.16	0.966	0.192
697	0.108	0.121	0.113	0.961	0.315

Hobs: Heterocigosidad observada

Hesp: Heterocigosidad esperada

HW (EH): Equilibrio de Hardy- Weinberg asumiendo exceso de heterocigotos

HW (DH): Equilibrio de Hardy- Weinberg asumiendo déficit de heterocigotos

En cuanto a la proteína si bien no se identificaron polimorfismos en la región codante, se pudo encontrar la presencia de una inserción/delección de 12 nucleótidos en la región codante. Todos los marcadores validados en este fragmento se encuentran en desequilibrio de ligamiento entre si ($p < 0.0001$) (Tabla 32).

Tabla 32. Determinación de desequilibrio de ligamiento entre los 9 posiciones polimórficas del fragmento del gen KRTAP6-1.

Posicion	77	110	212	369	502	506	557	612	697
77	-	0	0	0	0	0	0	0	0
110	*	-	0	0	0	0	0	0	0
212	*	*	-	0	0	0	0	0	0
369	*	*	*	-	0	0	0	0	0
502	*	*	*	*	-	0	0	0	0
506	*	*	*	*	*	-	0	0	0
557	*	*	*	*	*	*	-	0	0
612	*	*	*	*	*	*	*	-	0
697	*	*	*	*	*	*	*	*	-

Encima de la diagonal se muestran los p-values

Debajo de la diagonal si es significativo (*)

4.2.6.3 KRTAP9-2

Al realizar el alineamiento de las 40 secuencias del fragmento de KRTAP9-2 se identificaron 3 sitios polimórficos. Mediante inspección de los cromatogramas se determinó el genotipo en estos 3 sitios para las 96 muestras (Tabla 33). En cada una de las 3 posiciones se encontraron dos alelos, solo para la posición 598 no se encontró individuos homocigotos AA (Tabla 33). La posición polimórfica 598 se originó por una mutación de transición mientras que las 2 restantes se originaron por mutaciones de transversion (Tabla 33). Al determinar las frecuencias alélicas, la posición 598 presento uno de sus alelos con frecuencia menor al 5% (4.2%) (Tabla 33). Por lo tanto solo 2 sitios presentan frecuencias alélicas adecuadas para ser considerados marcadores genéticos.

Tabla 33. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 3 sitios polimórficos identificados en el fragmento del gen KRTAP9-2.

SNP	Alelos		Conteo de individuos			Frecuencias alelicas	
	Alelo1	Alelo2	Homocigotos1	Homocigotos2	Heterocigotos	Alelo1	Alelo2
86	C	T	63	1	31	0.827	0.173
291	A	T	1	78	15	0.090	0.91
598	A	G	0	87	8	0.042	0.958

A partir del listado de genotipos se determinó los posibles haplotipos para el fragmento amplificado tomando en consideración los 3 loci. En total se encontraron 6 haplotipos, los cuales presentan distintas frecuencias haplotípicas (Tabla 34). Los más comunes son los haplotipos 3 y 6, los cuales sumados tienen una frecuencia 86% (Tabla 34).

Tabla 34. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP9-2.

Haplotipos	1	2	3	Frecuencias
1	C	A	G	0.04
2	C	T	A	0.04
3	C	T	G	0.746
4	T	A	A	0.003
5	T	A	G	0.057
6	T	T	G	0.114

Designación de las posiciones del 1 al 3 de acuerdo al orden en que aparecen en el fragmento amplificado.

Se determinó en que región del gen se ubican estas posiciones. Las posiciones 86 y 291 se encuentran en la región codante mientras que la posición 598 se ubica en la región 3'UTR (Tabla 35). En cuanto a los polimorfismos en la región codante, ambos corresponden a cambios no sinónimos. La sustitución nucleotídica en la posición 86 produce un cambio en el residuo 62 de la proteína, donde se puede encontrar una Prolina o una Serina (Tabla 35). El sitio 291 produce un cambio en el residuo 131 de la proteína, donde los posibles aminoácidos son Leucina y Glutamina (Tabla 35).

Tabla 35. Ubicación de los sitios polimórficos dentro del gen KRTAP9-2 y los cambios en la secuencia de nucleótidos y de aminoácidos.

SNP	Region	Cambio en codon	Cambio en proteina
86	Codante	CCC <--> TCC	P <--> S
291	Codante	CTG <--> CAG	L <--> Q
598	3'UTR	-	-

Como se observa la posición 86 presenta un valor de heterocigosidad intermedio alrededor de 0.3, mientras que las otras dos posiciones muestran valores bajos, en especial la posición 598 (Tabla 36). Esto también se ve reflejado en los valores de PIC. Todas las posiciones se encuentran en EHW (Tabla 36). Por lo tanto se pudo validar 2 marcadores del tipo SNP en el fragmento del gen KRTAP9-2 en alpaca.

Tabla 36. Índices de variabilidad genética para cada loci en el gen KRTAP9-2 así como la evaluación del equilibrio de Hardy – Weinberg.

SNP	Hobs	Hesp	PIC	HW p-Value (EH)	HW p- Value (DH)
86	0.326	0.289	0.246	0.178	0.966
291	0.16	0.165	0.151	0.846	0.548
598	0.084	0.081	0.077	0.916	1

Hobs: Heterocigosidad observada

Hesp: Heterocigosidad esperada

HW (EH): Equilibrio de Hardy- Weinberg asumiendo exceso de heterocigotos

HW (DH): Equilibrio de Hardy- Weinberg asumiendo déficit de heterocigotos

También se quiso determinar las variantes a nivel proteico tanto a causa de las sustituciones nucleotidicas como también a causa de la inserción/delección. En cuanto a las variantes a causa de la sustitución nucleotidica se trabajó con los haplotipos encontrados en este fragmento (Tabla 34). Cada haplotipo fue traducido a una proteína donde se obtuvieron las distintas variantes y sus frecuencias (Tabla 37). Se identificaron 4 variantes distintas. Las variantes con un residuo de prolina en la posición 62 y un residuo de leucina en la posición 130 fueron las más frecuentes (0.78) (Tabla 37). El resto de variantes presento frecuencias considerablemente menores. En cuanto a la inserción/delección la variante que presenta 15 nucleótidos menos produce una proteína 5 residuos más pequeña sin correr el marco de lectura.

Tabla 37. Variantes a nivel proteico de acuerdo a los haplotipos identificados para la proteína KAP9-2.

Variante	haplotipo en gen	Residuo 62	Residuo 130	Frecuencia
1	1	P	Q	0.04
2	2,3	P	L	0.78
3	4,5	S	Q	0.06
4	6	S	L	0.12

Por último se determinó si los marcadores encontrados en el fragmento del gen KRTAP9-2 se encontraban en desequilibrio de ligamiento. Todos los marcadores validados en este fragmento se encuentran en desequilibrio de ligamiento entre si ($p < 0.0001$) (Tabla 38).

Tabla 38. Determinación de desequilibrio de ligamiento entre las 3 posiciones polimórficas del fragmento del gen KRTAP9-2.

Posicion	86	291	598
86	-	0	0
291	*	-	0
598	*	*	-

Encima de la diagonal se muestran los p-values
 Debajo de la diagonal si es significativo (*)

4.2.6.4 KRTAP11-1

Al realizar el alineamiento de las 40 secuencias del fragmento de KRTAP11-1 se identificaron 6 sitios polimórficos. Mediante inspección de los cromatogramas se determinó el genotipo en estos 6 sitios para las 96 muestras (Tabla 39). En cada una de las 6 posiciones se encontraron dos alelos. Para todas las posiciones se encontraron los tres genotipos posibles, solo en el caso de la posición 956 no se encontraron individuos homocigotos AA (Tabla 39). Dos de las posiciones polimórficas se originaron por una mutación de transición (761 y 795) mientras que las 4 restantes se originaron por mutaciones de transversion (Tabla 39). Todas las posiciones presentaron frecuencias alélicas mayores al 5% (Tabla 39). Por lo tanto los 6 sitios presentan frecuencias alélicas adecuadas para ser considerados marcadores genéticos.

Tabla 39. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 6 sitios polimórficos identificados en el fragmento del gen KRTAP11-1.

SNP	Alelos		Cuento de individuos			Frecuencias alélicas	
	Alelo1	Alelo2	Homocigotos1	Homocigotos2	Heterocigotos	Alelo1	Alelo2
330	T	G	25	30	41	0.474	0.526
583	T	G	30	25	41	0.526	0.474
595	A	C	27	30	41	0.474	0.526
761	C	G	25	30	41	0.474	0.526
795	A	G	30	25	41	0.474	0.526
956	A	G	0	72	23	0.121	0.879

A partir del listado de genotipos se determinó los posibles haplotipos para el fragmento amplificado tomando en consideración los 6 loci. En total se encontraron 3 haplotipos, los cuales presentan distintas frecuencias haplotípicas (Tabla 40). Los más comunes son los haplotipos 1 y 3, los cuales sumados tienen una frecuencia 88% (Tabla 40).

Tabla 40. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP11-1.

Haplotipos	1	2	3	4	5	6	Frecuencias
1	G	T	C	G	A	G	0.526
2	T	G	A	C	G	A	0.121
3	T	G	A	C	G	G	0.353

Designación de las posiciones del 1 al 6 de acuerdo al orden en que aparecen en el fragmento amplificado.

Se determinó también la ubicación de estos sitios dentro del gen. Cuatro posiciones (330, 583, 595 y 761) se encuentran en la región codante, mientras que las dos restantes se ubican en la región 3'UTR (Tabla 41). Todas las posiciones presentan sustituciones que producen cambios no sinónimos a nivel de la proteína (Tabla 41). La sustitución en la posición 330 produce un cambio en el residuo 12 de la proteína, donde las variantes pueden presentar una Serina o una Alanina (Tabla 41). La sustitución en la posición 583 produce un cambio en el residuo 96 de la proteína, este cambio corresponde a una Cisteína o una Fenilalanina (Tabla 41). En el caso de la sustitución en la posición 595, en la proteína el cambio se da en el residuo 100 y los posibles aminoácidos son Tirosina y Serina (Tabla 4). La sustitución en la posición 761 produce un cambio en el residuo 153 donde los posibles aminoácidos son una Cisteína o un Triptófano (Tabla 41).

Tabla 41. Ubicación de los sitios polimórficos dentro del gen KRTAP11-1 y los cambios en la secuencia de nucleótidos y de aminoácidos.

SNP	Region	Cambio en codon	Cambio en proteína
330	Codante	TCC <--> GCC	S <--> A
583	Codante	TGC <--> TTC	C <--> F
595	Codante	TAC <--> TCC	Y <--> S
761	Codante	TGC <--> TGG	C <--> W
795	3'UTR	-	-
956	3'UTR	-	-

Las 5 primeras posiciones presentan valores relativamente altos de heterocigosidad aproximadamente 0.4, mientras que la posición 956 presenta valores más pequeños (alrededor de 0.2) (Tabla 42). La misma tendencia se puede observar para los valores de PIC (Tabla 42). Todas las posiciones se encuentran en EHW (Tabla 42). Por lo tanto se pudo validar 6 marcadores del tipo SNP en el fragmento del gen KRTAP11-1 en alpaca.

Tabla 42. Índices de variabilidad genética para cada loci en el gen KRTAP11-1 así como la evaluación del equilibrio de Hardy – Weinberg.

SNP	Hobs	Hesp	PIC	HW p-Value (EH)	HW p- Value (DH)
330	0.396	0.502	0.375	0.952	0.105
583	0.396	0.502	0.375	0.952	0.105
595	0.396	0.502	0.375	0.952	0.105
761	0.4	0.501	0.374	0.952	0.105
795	0.4	0.501	0.374	0.952	0.105
956	0.242	0.214	0.19	0.22	1

Hobs: Heterocigosidad observada

Hesp: Heterocigosidad esperada

HW (EH): Equilibrio de Hardy- Weinberg asumiendo exceso de heterocigotos

HW (DH): Equilibrio de Hardy- Weinberg asumiendo déficit de heterocigotos

También se quiso determinar las variantes a nivel proteico a causa de las sustituciones nucleotídicas, se trabajó con los haplotipos encontrados en este fragmento (Tabla 40). Cada haplotipo fue traducido a una proteína donde se obtuvieron las distintas variantes y sus frecuencias (Tabla 43). Se identificaron 2 variantes, una de estas posee 2 residuos más de cisteína (variante 2). Ambas variantes presentan frecuencias similares (Tabla 43).

Tabla 43. Variantes a nivel proteico de acuerdo a los haplotipos identificados para la proteína KAP11-1.

Variante	haplotipo en gen	Residuo 12	Residuo 96	Residuo 100	Residuo 153	Frecuencia
1	1	A	F	S	W	0.53
2	2,3	S	C	Y	C	0.47

Por último se determinó si los marcadores encontrados en el fragmento del gen KRTAP11-1 se encontraban en desequilibrio de ligamiento. Todos los marcadores validados en este fragmento se encuentran en desequilibrio de ligamiento ($p < 0.0001$) (Tabla 44).

Tabla 44. Determinación de desequilibrio de ligamiento entre las 6 posiciones polimórficas del fragmento del gen KRTAP11-1.

Posicion	330	583	595	761	795	956
330	-	0	0	0	0	0
583	*	-	0	0	0	0
595	*	*	-	0	0	0
761	*	*	*	-	0	0
795	*	*	*	*	-	0
956	*	*	*	*	*	-

Encima de la diagonal se muestran los p-values
Debajo de la diagonal si es significativo (*)

4.2.6.5 KRTAP13-1

Al alinear 40 secuencias seleccionadas al azar que correspondían al fragmento del gen KRTAP13-1 se encontraron 4 sitios polimórficos. Se determinaron los genotipos en cada una de estas posiciones para las 96 muestras mediante una inspección manual de los cromatogramas. Cada uno de estos sitios presento 2 alelos distintos, en todas las posiciones se encontraron individuos con los 3 genotipos posibles (Tabla 45). El sitio 392 fue el único que se originó por una sustitución del tipo transición, mientras que el resto de sustituciones son del tipo transversion (Tabla 45). Todas las posiciones presentaron

frecuencias alélicas mayores al 5% (Tabla 45). Por lo tanto los 4 sitios presentan frecuencias alélicas adecuadas para ser marcadores genéticos.

Tabla 45. Alelos, frecuencia de los genotipos y frecuencias alélicas para los 6 sitios polimórficos identificados en el fragmento del gen KRTAP13-1.

SNP	Alelos		Conteo de individuos			Frecuencias alélicas	
	Alelo1	Alelo2	Homocigotos1	Homocigotos2	Heterocigotos	Alelo1	Alelo2
392	G	C	33	18	42	0.581	0.419
534	G	T	33	18	43	0.58	0.420
620	T	A	33	19	42	0.574	0.426
647	G	T	33	19	42	0.574	0.426

A partir del listado de genotipos se determinó los posibles haplotipos para el fragmento amplificado tomando en consideración los 4 loci. En total se encontraron 3 haplotipos, los cuales presentan distintas frecuencias haplotípicas (Tabla 46). Los más comunes son los haplotipos 1 y 3, los cuales sumados tienen una frecuencia 99.4%. El haplotipo 2 se encontró en solo un individuo.

Tabla 46. Haplotipos y sus frecuencias determinados para el fragmento del gen KRTAP13-1.

Haplotipos	1	2	3	4	Frecuencias
1	C	T	A	T	0.420
2	G	G	A	T	0.006
3	G	G	T	G	0.574

Designación de las posiciones del 1 al 4 de acuerdo al orden en que aparecen en el fragmento amplificado.

Al ubicar estos marcadores dentro del gen, se encontró que tres posiciones (392, 534 y 620) se ubican en la región codante, mientras que la posición 647 se ubica en la región 3'UTR (Tabla 47). En el caso de las posiciones dentro de la región codante, las tres sustituciones producen cambios no sinónimos en cuanto a la secuencia de la proteína (Tabla 47). La sustitución en la posición 392 produce un cambio en la posición 92 de la proteína presentándose los aminoácidos Triptófano o Cisteína, este mismo cambio se

observa en la sustitución en la posición 534 que produce un cambio en el residuo 139 (Tabla 47). La sustitución en la posición 620 produce un cambio en el residuo número 168, pudiéndose encontrar en esta posición una Fenilalanina o una Tirosina (Tabla 47).

Tabla 47. Ubicación de los sitios polimórficos dentro del gen KRTAP13-1 y los cambios en la secuencia de nucleótidos y de aminoácidos.

SNP	Region	Cambio en codon	Cambio en proteína
392	Codante	TGG <--> TGC	W <--> C
534	Codante	TGG <--> TGT	W <--> C
620	Codante	TTC <--> TAC	F <--> Y
647	3'UTR	-	-

En cuanto a los valores de Heterocigosidad, todos los marcadores presentan valores altos (mayor a 0.4) (Tabla 48). Lo mismo se observa en el caso de PIC (Tabla 48). Por ultimo todos los marcadores se encuentran en EHW (Tabla 48). Por lo tanto se pudo validar 4 marcadores del tipo SNP en el fragmento del gen KRTAP13-1 en alpaca.

Tabla 48. Índices de variabilidad genética para cada loci en el gen KRTAP13-1 así como la evaluación del equilibrio de Hardy – Weinberg.

SNP	Hobs	Hesp	PIC	HW p-Value (EH)	HW p- Value (DH)
392	0.452	0.49	0.368	0.834	0.294
534	0.457	0.49	0.369	0.805	0.331
620	0.447	0.492	0.369	0.865	0.249
647	0.447	0.492	0.369	0.865	0.249

Hobs: Heterocigosidad observada

Hesp: Heterocigosidad esperada

HW (EH): Equilibrio de Hardy- Weinberg asumiendo exceso de heterocigotos

HW (DH): Equilibrio de Hardy- Weinberg asumiendo déficit de heterocigotos

También se quiso determinar las variantes a nivel proteico a causa de las sustituciones nucleotidicas, se trabajó con los haplotipos encontrados en este fragmento (Tabla 46). Cada haplotipo fue traducido a una proteína donde se obtuvieron las distintas variantes y sus frecuencias (Tabla 49). Se identificaron 3 variantes, una de estas posee 2 residuos más

de cisteína (variante 1). Las variantes 1 y 3 tienen frecuencias altas y parecidas mientras que la variante 2 solo se encontró en un individuo.

Tabla 49. Variantes a nivel proteico de acuerdo a los haplotipos identificados para la proteína KAP13-1.

Variante	haplotipo en gen	Residuo 92	Residuo 139	Residuo 168	Frecuencia
1	1	C	C	Y	0.42
2	2	W	W	Y	0.006
3	3	W	W	F	0.574

Por último se determinó si los marcadores encontrados en el fragmento del gen KRTAP13-1 se encontraban en desequilibrio de ligamiento. Todos los marcadores validados en este fragmento se encuentran en desequilibrio de ligamiento ($p < 0.0001$) (Tabla 50).

Tabla 50. Determinación de desequilibrio de ligamiento entre las 4 posiciones polimórficas del fragmento del gen KRTAP13-1.

Posicion	392	534	620	647
392	-	0	0	0
534	*	-	0	0
620	*	*	-	0
647	*	*	*	-

Encima de la diagonal se muestran los p-values
Debajo de la diagonal si es significativo (*)

4.2.7 Desequilibrio de ligamiento entre marcadores de diferentes genes

Se evaluó, considerando las 32 posiciones polimórficas, que pares se encuentran en desequilibrio de ligamiento (Tabla 51). Se encontró lo observado para cada gen, que todas las posiciones dentro de un mismo fragmento se encuentran en desequilibrio de ligamiento (Tabla 51). Por otro lado 28 posiciones que corresponden a genes diferentes también se encuentran en desequilibrio de ligamiento. Se observa que son entre posiciones del gen KRTAP1-2 y KRTAP9-2, como también posiciones entre el gen KRTAP11-1 y KRTAP6-1 (Tabla 51). La posición 86 del gen KRTAP9-2 está en desequilibrio de ligamiento con 9 posiciones (224, 246, 262, 370, 639, 687, 705, 721, 814) del gen KRTAP1-2 (Tabla 51), la posición 291 del gen KRTAP9-2 se encuentra en desequilibrio de ligamiento con 7 posiciones (224, 246, 262, 370, 687, 705, 814) del gen KRTAP1-2 (Tabla 51) y la posición 598 se encuentra en desequilibrio de ligamiento con 2 posiciones (639, 814) del gen KRTAP1-2 (Tabla 51).

Por otro lado la posición 506 del gen KRTAP6-1 se encuentra en desequilibrio de ligamiento con 5 posiciones (330, 583, 595, 761, 795) del gen KRTAP11-1 (Tabla 51), la posición 557 del gen KRTAP6-1 se encuentra en desequilibrio de ligamiento con dos posiciones (330, 583) del gen KRTAP11-1 (Tabla 51) y la posición 612 del gen KRTAP6-1 se encuentra en desequilibrio de ligamiento con 3 posiciones (330, 583, 595) del gen KRTAP11-1 (Tabla 51).

V. **Componente II. Evaluación de asociación del gen KRTAP11-1 al diámetro de fibra en alpaca huacaya**

5.1 Materiales y Métodos

5.1.1 Selección de gen candidato para evaluar la asociación al diámetro de fibra en alpaca.

Para la selección del gen candidato se utilizaron varios criterios que permitieron una selección objetiva. Los criterios de selección fueron los siguientes:

- Presencia de marcadores en región codante, asumiendo que si existe una asociación debe estar relacionada a la función de la proteína.
- Presencia de cambios no sinónimos a nivel de la proteína.
- Sustituciones de aminoácidos significativas, por ejemplo el cambio de un aminoácido a una cisteína en el caso de proteínas HS y UHS.
- Niveles altos de heterocigosidad (alrededor de 0.3).
- Número de marcadores en región codante.
- Revisión bibliográfica de la importancia del gen en el desarrollo de la fibra.

5.1.2 Muestra para el estudio de asociación

Se trabajó con una muestra de alpacas blancas de la variedad Huacaya, todas pertenecientes a la región de Puno y al criadero Rural Alianza EPS. En total se contaba con una base de datos de 199 individuos de diferentes edades y de ambos sexos. De cada individuo se tenía una muestra de ADN genómico, la edad del animal expresada de acuerdo al número de dientes, el sexo del animal, el diámetro de la fibra en micrómetros, una ficha técnica (peso del animal, incidencia de parásitos, entre otros) y el genotipo para 10 microsatélites que conforman una prueba de paternidad desarrollada en la UBM.

5.1.3 Análisis de la muestra

5.1.3.1 Cuantificación de las muestras de ADN

Se cuantifico las 199 muestras de ADN genómico de alpaca utilizando un Nanodrop 2000c. Además de la cuantificación se evaluó la pureza de las muestras para determinar la presencia de contaminantes como, por ejemplo el fenol. La pureza se evaluó calculando los ratios de absorbancia 260/280 y 260/230. Para el ratio 260/280 se espera un valor de 1.8, mientras que para el ratio 260/230 se espera un valor entre 1.8 y 2.2. Una vez cuantificadas las muestras, se realizaron diluciones a 10 ng/ul.

5.1.3.2 Parentesco entre las muestras

El nivel de parentesco entre los individuos seleccionados para la evaluación es muy importante, debido a que de esto depende el modelo mediante el cual se va a evaluar la asociación. En este caso se busca obtener una muestra de individuos no relacionados de ser posible. Para lograr esto se realizó una prueba de parentesco. Al no contar con registros de empadre, se realizó una prueba “todos contra todos”. Esta consistía en enlistar a todos los individuos como crías y a los machos como padres y las hembras como madres. La prueba se realizó utilizando el Software Cervus y el genotipo de los diez microsatelites. Se buscaron individuos que presentaran un grado de parentesco madre – cría o padre – cría a un nivel de confianza del 95%. Al encontrar estas relaciones, se eliminaba uno de los individuos. De esta forma el grupo que restaba podía ser considerado como individuos no relacionados.

5.1.3.3 Estructuración poblacional

La presencia de estructuración poblacional es otra característica importante a tomar en cuenta para la selección del modelo mediante el cual se evaluará la asociación. Para evaluar esto se utilizó las 152 muestras no relacionadas (luego de aplicar la prueba de parentesco) y los genotipos de los 10 microsatélites. Se determinó la presencia/ausencia de estructuración poblacional utilizando el Software STRUCTURE y determinando los valores de ΔK (Evanno et al., 2005). Se evaluó hasta 7 *clusters*, cada uno con 20 repeticiones.

Parámetros para el análisis en Structure:

Burnin Period: 250000
Number of MCMC after burnin: 500000
Admixture model
Allele frequencies correlated

5.1.3.4 Distribución normal del diámetro de fibra en la muestra

Se evaluó que el diámetro de fibra tuviera una distribución normal en la muestra utilizando el paquete estadístico STATA y la prueba de Shapiro-Wilk.

5.1.3.5 Poder estadístico de la prueba

El poder de la prueba depende del número de la muestra, la varianza de la variable dependiente, las frecuencias alélicas y el efecto que tiene el gen sobre la característica de interés. Para evaluar el poder de la prueba se realizó una simulación en R. Se simuló distintas muestras considerando 152 individuos. Estas muestras presentaban efectos diferentes del gen sobre la media del diámetro ($a = 0.2, 0.4, 0.6, 0.8, 1$) y frecuencias alélicas diferentes ($p = 0.1, 0.2, 0.3, 0.4, 0.5$). Para la combinación de cada uno de estos parámetros se realizaron 10000 repeticiones. En

cada una de estas repeticiones se evaluó si una regresión lineal simple era capaz de detectar la asociación existente. Se calculó la proporción de resultados positivos sobre el total de resultados para evaluar el poder de la prueba. El resultado se presentó en forma de grafica donde el eje y representa el poder de la prueba y el eje x el efecto del gen expresado en función de la heredabilidad en sentido estricto (h^2). El *script* fue diseñado por el alumno de maestría Diego Veliz Otani.

5.1.4 Modelo de asociación

Para evaluar si existe asociación entre el gen candidato y el diámetro de fibra en alpaca se realizó una regresión lineal múltiple siguiendo un modelo aditivo usando el paquete estadístico de STATA. El modelo se eligió luego de asegurar que la muestra cumplía con todos los supuestos. Los datos se ajustaron al siguiente modelo:

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + \beta_3 j_i + \beta_4 k + e_i$$

Dónde: Y_i : Diámetro de fibra de la alpaca i .
 β_0 : Intercepto.
 β_1 : Coeficiente de regresión para la variable genotipo SNP330 (KRTAP11-1).
 X_i : Genotipo de la alpaca i del SNP330.
 β_2 : Coeficiente de regresión para la variable genotipo SNP956 (KRTAP11-1).
 z_i : Genotipo de la alpaca i del SNP956.
 β_3 : Coeficiente de regresión para la variable sexo.
 J_i : Sexo de la alpaca i .
 β_4 : Coeficiente de regresión para la variable edad.
 K_i : Edad de la alpaca i
 e_i : Error residual del modelo.

Para evaluar el efecto del gen se utilizó el genotipo del SNP330 (KRTAP11-1) y el genotipo del SNP956 (KRTAP11-1). El genotipo del SNP330 representa también el genotipo de los SNPs (583, 595, 761, 795) debido a que todos se encuentran en desequilibrio de ligamiento completo. Sin embargo el genotipo del SNP330 no predice el genotipo del SNP956 por lo tanto se incluyó el genotipo de este marcador en el modelo. Se asume un modelo de herencia aditivo, por lo tanto en los datos de entrada del genotipo del SNP330 para la regresión el valor 0 corresponde a la ausencia del alelo T (Homocigoto GG), 1 a la presencia de 1 alelo T (heterocigoto) y el valor 2 a la presencia de dos alelos T (homocigoto TT). Para el genotipo del SNP956 el valor 0 corresponde a la ausencia del alelo A (Homocigoto GG), el valor 1 corresponde a la presencia de un alelo A (Heterocigoto) y el valor 2 corresponde a la presencia de 2 alelos A (Homocigoto AA). La edad se presenta como una variable cualitativa, esta expresada en base a la cantidad de dientes que presenta el animal. El valor 1 se le asigna cuando el animal presenta dientes de leche (DL), el valor 2 se le asigna al animal cuando presenta 2 dientes, el valor 3 se le asigna al animal cuando presenta 4 dientes y el valor 4 se le asigna al animal cuando presenta la dentadura completa. En cuanto al sexo, el valor 1 corresponde a las hembras y el valor 2 corresponde a los machos (Anexo 6). La significancia estadística se evaluó mediante una prueba t- student.

5.1.5 Amplificación y secuenciamiento de las muestras

Para la amplificación de las muestras se utilizaron los cebadores 11F y 11R (Tabla 18) y siguiendo las condiciones con las que se estandarizó la PCR (Tabla 19). Una vez amplificadas las muestras se determinó la presencia de producto en un gel de agarosa al 1%. Aquellas muestras que presentaron producto fueron enviadas a la empresa Macrogen Inc. para la purificación y el secuenciamiento. Al obtener las secuencias se determinaron los genotipos para los 152 individuos en los 6 loci identificados para el gen KRTAP11-1.

Esta data fue utilizada para evaluar la asociación entre el gen KRTAP11-1 y el diámetro de la fibra, ajustando los datos al modelo previamente descrito (Anexo 6).

5.2 Resultados

5.2.1 Selección de KRTAP11-1 como gen candidato para evaluar la asociación al diámetro de fibra en alpaca

Se determinó al gen KRTAP11-1 como el gen candidato para evaluar la asociación luego de revisar cada uno de los criterios propuestos en la sección de materiales y métodos. En cuanto a los criterios presencia y número de marcadores en región codante, los genes KRTAP1-2, KRTAP9-2, KRTAP11-1 y KRTAP13-1 presentaron marcadores en región codante (Tabla 32; 42; 48; 54). De estos 4 genes KRTAP11-1 fue el que presentó un mayor número de marcadores en región codante (4). Con respecto a los cambios no sinónimos, los marcadores en región codante de los cuatro genes correspondían a cambios no sinónimos en la proteína (Tablas 25; 35; 41; 47). Sin embargo, al observar qué tipos de cambios eran los que ocurrían, en tres de los cuatro genes (KRTAP1-2, KRTAP11-1 y KRTAP13-1) las sustituciones producían variaciones en el número de cisteínas (Tablas 25; 41; 47). Este cambio es de importancia para este grupo de proteínas debido a la frecuente formación de puentes disulfuro para estabilizar las microfibrillas. Los genes KRTAP11-1 y KRTAP13-1 presentaron 2 sustituciones de cisteína a diferencia del gen KRTAP1-2 que tan solo presentó uno. En cuanto a la informatividad de los marcadores, los marcadores validados en los genes KRTAP11-1 y KRTAP13-1 fueron los más informativos de acuerdo a los índices de heterocigosidad y PIC (Tabla 26; 31; 36; 42; 48). Además es importante mencionar que el gen KRTAP11-1, ha sido descrito como componente crucial para la formación de la fibra (Fujimoto

et al., 2014). No se tomó en cuenta la presencia de inserciones/deleciones presentes en los diferentes genes debido a que no pudieron ser validadas como marcadores en este estudio, sin embargo es probable que sean de mucha importancia.

5.2.2 Análisis de la muestra

5.2.2.1 Parentesco

Se tomaron en cuenta solo aquellas relaciones estadísticamente significativas a un nivel de confianza de 95% donde el LOD score era alto y solo existía un mismatch en cuanto a los genotipos. En total se encontraron 73 relaciones de parentesco que seguían los criterios descritos. Sin embargo en 35 casos se repetían individuos, por lo tanto se eliminaron tan solo 47 individuos de la muestra original. Los 152 individuos restantes pueden ser considerados como individuos no relacionados.

5.2.2.2 Estructuración poblacional

Se evaluó la presencia de estructuración poblacional en la muestra constituida por las 152 alpacas no relacionadas. Se evaluó la misma población asumiendo distintos números de *clusters* ($K = 1, 2, 3, 4, 5, 6, 7$) (Anexo 5). Para la determinación del número de clusters K , en el que mejor se divide la muestra, se calcularon los valores de ΔK de acuerdo a Evanno y colaboradores (Anexo 5) (Evanno et al., 2005). Al observar los resultados se observa que no existe estructuración poblacional en la muestra (Figura 41; Anexo 5).



Figura 41 Diagrama de barras obtenido de Structure considerando los 10 microsateles y los 152 individuos, asumiendo 7 clusters ($K = 7$). En el eje X se muestran los 152 individuos analizados, cada color corresponde a uno de los 7 clusters. Se puede observar que a lo largo del eje x no se observa ninguna diferencia en las proporciones de cada cluster, esto indica la ausencia de estructuración poblacional

5.2.2.3 Distribución normal del diámetro de fibra en la muestra seleccionada para la evaluación de asociación

Utilizando la prueba de Shapiro – Wilk se determinó que el diámetro de la fibra en la muestra sigue una distribución normal ($p = 0.054$).

5.2.2.4 Poder estadístico de la prueba

Se realizaron 10000 simulaciones de muestras de 152 individuos utilizando los parámetros descritos en la sección de Materiales y Métodos. Se determinó que utilizando una regresión lineal simple es posible detectar un efecto del gen de 8% con un poder estadístico de 0.8. Este poder se alcanza siempre que se utilice un marcador donde el alelo menos representado tenga una frecuencia igual o mayor al 40% (Figura 42). Los marcadores en el gen KRTAP11-1 presentan frecuencias alélicas de alrededor de 0.4.

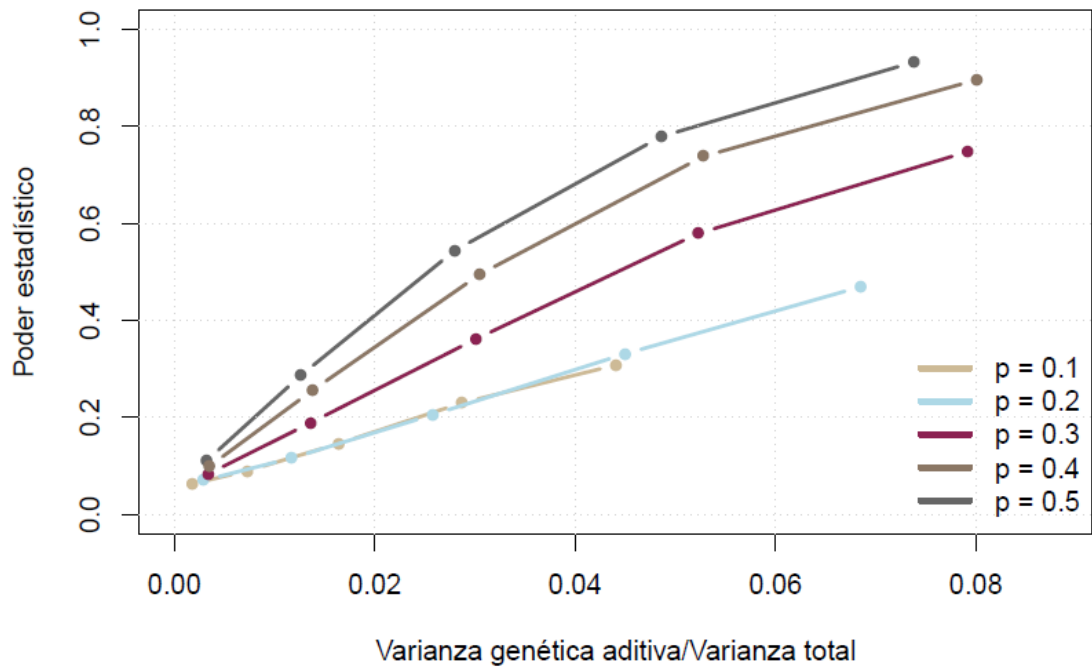


Figura 42. Se muestra la relación entre el poder de la prueba y el efecto del gen en función de las frecuencias alélicas del marcador seleccionado. En el eje Y se observa la proporción entre resultados positivo y el total de resultados (Poder estadístico). En el eje X se muestra el efecto del gen expresado como la heredabilidad en sentido estricto (h^2). (Fuente: Diego Veliz Otani, no publicado)

5.2.3 Amplificación de las muestras

Las 152 muestras fueron amplificadas utilizando los cebadores 11F y 11R. Se determinó la presencia de los 152 productos de PCR mediante electroforesis en un gel de agarosa al 1%. Al obtener las secuencias se determinó el genotipo en los 6 marcadores para los 152 individuos.

5.2.4 Evaluación de la asociación del gen KRTAP11-1 y el diámetro de fibra

Al ajustar la data al modelo propuesto en la sección de Materiales y Métodos utilizando una regresión lineal múltiple se obtuvo que los dos marcadores analizados en KRTAP11-1 no tiene un efecto significativo sobre el diámetro de fibra en alpaca. Se obtuvo un coeficiente de regresión de 0.101 y un p-Value de 0.8 para el genotipo en el SNP330 y un coeficiente de regresión de 0.215 y un p-Value de 0.723 para el genotipo en el SNP956 (Tabla 52). De acuerdo al modelo, la edad sí tiene un efecto significativo sobre el diámetro de fibra, con un coeficiente de regresión de 0.563 y un P- value de 0.01 (Tabla 52). El sexo no tiene un efecto significativo sobre el diámetro de la fibra a pesar de presentar un coeficiente de regresión alto. Esto podría deberse al hecho que el número de machos es muy pequeño con respecto al número de hembras (Tabla 52). El modelo sin embargo no explica el diámetro de fibra en alpaca debido a que R^2 fue de 0.05.

Tabla 52. Determinación de los coeficientes de regresión y su significancia estadística (P-value).

Variable	Coeficiente de regresion	ES	Valor de t	P value
SNP330	0.1017	0.3999	0.25	0.8
SNP956	0.2148	0.604	0.36	0.723
Edad	0.5667	0.2153	2.63	0.009
Sexo	1.015	0.8284	1.23	0.222
Constante	18.3964	1.183	15.55	0

VI. Discusión

El Perú es el mayor productor de fibra de alpaca a nivel mundial (Quispe *et al.*, 2009a), por lo tanto es necesario aprovechar este recurso al máximo garantizando una producción de primera calidad. Sin embargo se ha observado todo lo contrario, de la producción de fibra peruana solo el 30% corresponde a fibra categorizada como fina (De los Ríos, 2006). Varios factores influyen en este hecho, por un lado la falta de tecnología disminuye la producción, ya que las esquilas se realizan con elementos muy rudimentarios (Quispe, 2005; observación del autor). Por otro lado un mal manejo por parte de los productores, la falta de registros y los métodos de selección de individuos reproductores ha llevado a la pérdida de calidad de la fibra que se ve reflejada en un engrosamiento del diámetro promedio de las alpacas peruanas (Quispe *et al.*, 2013; De los Ríos, 2006; Wheeler, 2012). Para mejorar la calidad de la fibra de alpaca es necesario implementar planes de mejoramiento genético, orientados a una disminución del diámetro de fibra (Quispe *et al.*, 2013). Este estudio contribuye con este fin a través de la generación de marcadores genéticos en genes que codifican proteínas estructurales de la fibra, estos marcadores pueden servir como herramientas para la implementación de programas de mejoramiento genético. Además se evaluó a uno de estos genes, para determinar está asociado al diámetro de fibra.

Este estudio se enfocó en la familia de proteínas asociadas a queratina (KAPs), debido a que son buenos candidatos a estar asociados a diferentes características de interés comercial de la fibra en alpaca (Ian y Ian, 2005). Khan y colaboradores analizaron estos genes en los genomas de 30 mamíferos distintos y encontraron que algunas especies que presentan modificaciones a nivel del pelo presentan

diferencias importantes en el repertorio de estos genes (Khan *et al.*, 2014). Por otro lado en varios estudios con respecto a la expresión de estos genes, se ha observado un control estricto, donde la expresión de estos depende de la zona del folículo como también de la fase del ciclo en la que se encuentra (Rogers *et al.*, 2006; Plowman *et al.*, 2015). También se ha determinado que desempeñan un rol crítico en la formación del pelo, sobre todo durante el ensamblaje de los filamentos intermedios de queratina (Fujimoto *et al.*, 2014). Además, estos genes son considerados altamente variables a diferencia de la otra gran familia de proteínas estructurales; Las queratinas. Esta variabilidad podría explicar las adaptaciones tan distintas del pelo entre las distintas especies de mamíferos, como lo es la lana en alpaca (Khan *et al.*, 2014; Rogers *et al.*, 2006). Por otro lado diversos estudios en especies productoras de fibra (oveja y cabra) han reportado polimorfismos en esta familia de genes asociados a diferentes características comerciales de la fibra (Rogers y Schweizer, 2005; Gong *et al.*, 2011a, Gong *et al.*, 2011b; Itenge – Mweza *et al.*, 2007; Shimomura *et al.*, 2002; Zhou *et al.* 2015; Mckenzie *et al.*, 2010; Zhang *et al.*, 2011; Gong *et al.*, 2015; Yu *et al.*, 2008). Sin embargo es importante mencionar que existen otros grupos de genes que podrían clasificarse como buenos candidatos, como por ejemplo genes regulatorios durante el establecimiento de folículos o durante la formación de la fibra (Ian y Ian, 2005), siendo necesario también generar marcadores en este tipo de genes.

Los genes con los que se trabajaron en este estudio fueron seleccionados considerando dos criterios. En primer lugar: Que se expresen en el folículo de alpaca, lo cual fue determinado gracias a la disponibilidad de una biblioteca de cDNA de piel de alpaca previamente elaborada por la Unidad de Biotecnología

Molecular (UPCH). En segundo lugar se buscaron genes con reportes previos de asociación a diferentes características de fibra en otras especies. Se combinó ambas informaciones, para trabajar con genes que se estén expresando en alpaca y que cuenten con reportes en otras especies.

El uso de una biblioteca de cDNA permitió identificar 9 fragmentos de genes KRTAP mediante la herramienta BLAST (Tabla 9). La presencia de estos genes en la biblioteca de cDNA indica que están siendo expresados en el folículo de alpaca, debido a que esta familia solo se expresa en este tejido (Plowman *et al.*, 2015). Adicionalmente, se buscaron estudios en otras especies donde se reporten polimorfismos en estos genes y/o estudios de asociación a distintas características de la fibra o lana. Se encontraron estudios para los genes KRTAP1-2, KRTAP9-2, KRTAP11-1 y KRTAP13-1. Además de estos genes se decidió incluir al gen KRTAP6-1, que a pesar de no haberse encontrado en la biblioteca de cDNA, presenta varios reportes a estar asociado a diversas características de la fibra de otras especies.

El gen KRTAP1-2 ha sido estudiado en oveja y en humano, en ambas especies se han identificado polimorfismos en región codante. En humano se reportaron variaciones en la longitud de las secuencias y una diferencia alélica considerable entre individuos japoneses e individuos caucásicos (Shimomura *et al.*, 2002). En oveja se reportaron 10 alelos en este gen, además se evaluó el perfil de expresión encontrándose que se expresa en el córtex de la fibra (Gong *et al.*, 2011b). Por último en un estudio de asociación, se encontraron variantes del gen asociadas a distintas características de la fibra como por ejemplo al coeficiente de variación del diámetro de la fibra (Gong *et al.*, 2015).

El gen KRTAP6-1 ha sido estudiado únicamente en oveja. Los primeros estudios identificaron marcadores del tipo RFLP, en total 6 alelos diferentes. Estos presentaron asociación al diámetro de fibra sin embargo solo en la progenie de uno de seis machos (Parsons *et al.*, 1994). Por otro lado en un estudio de patrones de expresión se determinó que este gen se expresa en folículo de oveja (Yu *et al.*, 2009). Por último se generaron marcadores mediante SSCP también en oveja, en este estudio se encontró una inserción/delección de 57 pb que estaba asociada a un diámetro de fibra mayor (Zhou *et al.*, 2015).

El gen KRTAP9-2 ha sido muy estudiado en cabra productora de fibra cashemere. Se ha identificado una variación del tipo inserción/delección de 30 nucleótidos en la región codante del gen (Yu *et al.*, 2008). También se ha reportado un SNP C/T en este gen y se observó que la variante CC tenía una expresión mucho menor que CT y TT (Wang *et al.*, 2012). Por último se hizo un perfil de expresión del gen en dos fases distintas del ciclo del pelo. Este estudio encontró que existía una diferencia significativa en la expresión de este gen entre cabras con alta producción y baja producción de fibra cashemere (Wang *et al.*, 2015).

El gen KRTAP11-1 ha sido estudiado en oveja y en ratón. En oveja se identificaron 6 alelos diferentes, todos debido a mutaciones puntuales (Gong *et al.*, 2011a). En este gen, a diferencia de la mayoría de KRTAPs, no se identificó una inserción/delección. Por otro lado se ha observado en ratones que esta proteína es crucial para el ensamblaje correcto de los filamentos intermedios de queratina (KIF) y que una expresión anormal de esta proteína es uno de los factores responsables para un fenotipo aberrante (Fujimoto *et al.*, 2014).

El gen KRTAP13-1 ha sido estudiado en oveja y en cabra. En cabra se encontró una mutación G/T, donde el genotipo GG fue asociado a un diámetro de fibra más delgado (Fang *et al.*, 2010). Por otro lado hay un estudio que se realizó en oveja, donde un grupo de ovejas fueron alimentadas con una dieta restringida (pocos nutrientes) mientras que otro grupo sirvió de control. Lo que se observó fue que hubo una disminución en el diámetro de fibra en el grupo de la dieta restringida con respecto al grupo control. Luego se hizo un análisis proteómico y se encontró mayor presencia de KRTAP13-1 asociado a esta diferencia en el diámetro (Almeida *et al.*, 2014).

Para la obtención de las secuencias de los genes en alpaca se trabajó a partir de la información de un genoma de referencia de alpaca disponible en el Genbank (**Vicugna-Pacos 2.0.1**) y utilizando los fragmentos de la biblioteca de cDNA. Para determinar si las secuencias obtenidas corresponden a los genes deseados se compararon mediante alineamiento, las proteínas ortólogas de oveja y humano con los ORFs traducidos de alpaca. Para los genes KRTAP1-2, KRTAP6-1 y KRTAP11-1 se utilizaron las proteínas reportadas en oveja (Figuras 14; 16; 20). Mientras que para los genes KRTAP9-2 y KRTAP13-1 se utilizaron las proteínas en humano debido a que estos genes no han sido reportados en oveja (Figuras 18; 22). Para los tres genes donde se utilizaron proteínas de oveja la identidad fue aproximadamente 80% mientras que en los casos donde se utilizó la proteína en humano la identidad fue aproximadamente 60%. Se espera que haya una mayor identidad entre las proteínas de la alpaca y la oveja que entre las proteínas de la alpaca y el humano, debido a que ambas pertenecen al mismo orden (Artiodactyla). Para determinar si la diferencia en el porcentaje de identidad se debe a la cercanía

evolutiva entre la alpaca y la oveja y no a una incorrecta asignación del gen, se alineo el ORF traducido en alpaca del gen KRTAP6-1 con la proteína en humano. Se encontró una identidad menor que al compararla con la proteína en oveja y similar al obtenido a las otras proteínas que fueron comparadas con proteínas humanas (67%) (Anexo 3). Sin embargo al hacer lo mismo con la proteína del gen KRTAP11-1 se encontró una identidad similar a la observada con oveja de 80% (Anexo 3), este resultado podría sugerir que KRTAP11-1 es una proteína muy conservada entre las distintas especies de mamíferos. Se ha visto que este gen tiene uno de los perfiles de expresión más constantes durante la formación de la fibra o pelo, tanto en humano como en oveja (Plowman *et al.*, 2015; Rogers *et al.*, 2006). Estos resultados permiten determinar que las secuencias identificadas en alpaca en este estudio corresponden a los genes deseados.

Al analizar las secuencias de los genes se observa que los elementos regulatorios (caja TATA y señal de poliadenilacion AAUAAA) se encuentran distribuidos dentro de estos de manera similar. Para los genes KRTAP1-2, KRTAP9-2, KRTAP11-1 y KRTAP13-1 la secuencia consenso de la caja TATA se encuentra entre 83 -84 pb *upstream* del codón de inicio (Figuras 15A; 19A; 21A; 23A), mientras que para el gen KRTAP6-1 se encuentra a 92 pb *upstream* del codón de inicio (Figura 17A). Los 4 genes que tiene la caja TATA más cerca al codón de inicio (83 pb) pertenecen a los grupos UHS/HS (Tabla 3) mientras que el único gen que la presenta a 92 pb del codón de inicio pertenece al grupo HGT (Tabla 3). Se ha visto que estos grupos presentan diferencias en los patrones de expresión (Roger, 2006), la diferencia en la distancia de la caja TATA al codón de inicio podría estar asociado a esta diferencia. Sería interesante determinar más secuencias de genes

pertenecientes al del grupo HGT para observar si este patrón se repite. En cuanto a la región 3'UTR, esta tiene un tamaño promedio de 400 pb, con excepción del gen KRTAP6-1 donde el tamaño es menor. Esta diferencia en la región 3'UTR va de acuerdo a lo observado en humanos (Rogers *et al.*, 2006). Todos los genes presentan la secuencia señal de poliadenilación AAUAAA en esta región a una distancia similar del último codón (Figuras 15B; 17B; 19B; 21B; 23B).

Los tamaños de las proteínas en alpaca como el contenido de aminoácidos, guardan relación con lo reportado en proteínas humanas. La proteína KAP1-2 en alpaca tiene un tamaño de 13.2 kDa (Figura 15B); lo que se ha reportado para esta subfamilia en humano es un tamaño entre 12.3 – 18.2 kDa (Rogers *et al.*, 2006). En cuanto al contenido de aminoácidos, la proteína en alpaca tiene 26 % de residuos de cisteína, 12.3 % de serina y 11.5% de glicina (Figura 15B). En humano se ha reportado entre 24.1–26.5% de cisteína, entre 12.9–17.5 % de serina y entre 8.6–11.4% de glicina para esta subfamilia (Rogers *et al.*, 2006). Estas características colocan esta proteína en alpaca dentro del grupo HS. La proteína KAP6-1 en alpaca tiene un tamaño de 7.7 kDa (Figura 17B); lo que se ha reportado para esta subfamilia en humano es un tamaño entre 7.3 – 11.1 kDa (Rogers *et al.*, 2006). En cuanto al contenido de aminoácidos, la proteína en alpaca tiene 38.7% de residuos de glicina, 21.3 % de tirosina y 16% de cisteína (Figura 17B). En humano se ha reportado entre 32.3–38% de glicina, entre 18.2–24.2 % de tirosina y una alto contenido de cisteína (Rogers *et al.*, 2006). Estas características colocan esta proteína en alpaca dentro del grupo HGT. La proteína KAP9-2 en alpaca tiene un tamaño de 17.5 kDa (Figura 19B), lo que se ha reportado para esta subfamilia en humano es un tamaño entre 16.4 – 26.3 kDa (Rogers *et al.*, 2006). En cuanto al contenido de aminoácidos, la

proteína en alpaca tiene 31.7 % de residuos de cisteína, 13.8 % de serina y 12.6% de treonina (Figura 19B). En humano se ha reportado entre 31.3–32.7% de cisteína, además de altos contenidos de treonina, serina y prolina (Rogers *et al.*, 2006). Estas características colocan esta proteína en alpaca dentro del grupo UHS. La proteína KAP11-1 en alpaca tiene un tamaño de 17.3 kDa (Figura 21B), lo que se ha reportado para este gen en humano es un tamaño de 17.1 kDa (Rogers *et al.*, 2006). En cuanto al contenido de aminoácidos, la proteína en alpaca tiene 12.3 % de residuos de cisteína, 17.2 % de serina y 11% de valina (Figura 21B). En humano se ha reportado 12.8% de cisteína, 14.7% de serina y contenido alto en treonina (Rogers *et al.*, 2006). Estas características colocan esta proteína en alpaca dentro del grupo HS. La proteína KAP13-1 en alpaca tiene un tamaño de 18.4 kDa (Figura 23B), lo que se ha reportado para esta subfamilia en humano es un tamaño entre 17.8 – 18.2 kDa (Rogers *et al.*, 2006). En cuanto al contenido de aminoácidos, la proteína en alpaca tiene 8 % de residuos de cisteína, 23.2 % de serina y 10.7% de glicina (Figura 23B). En humano se ha reportado entre 12 – 12.8% de cisteína, 20.3 – 26.4 % de serina y un contenido alto en glicina (Rogers *et al.*, 2006). Estas características colocan esta proteína en alpaca dentro del grupo HS. Tomando en cuenta el grado de identidad que guardan los ORF traducidos con las proteínas ortólogas en otras especies, además de los tamaños y la composición de aminoácidos presentadas se puede concluir que características las secuencias presentadas en este estudio corresponden a los genes deseados.

La identificación de los marcadores en los 5 genes se hizo mediante secuenciamiento directo de productos de PCR, por lo tanto el diseño de los cebadores constituyó una parte importante de este proceso. Todos los cebadores se

diseñaron utilizando el Software Primer3. En la medida de lo posible el contenido de GC se mantuvo entre 40 y 60%, los tamaños de los cebadores estuvieron entre 19 a 22 nt y se procuró que no exista una diferencia grande (máximo 2°C) entre las T_m . Además de estos parámetros se tuvo en cuenta para el diseño, el valor de ΔG para la formación de las distintas estructuras secundarias posibles (homodimeros, heterodimeros y hairpins). Para asegurar que la PCR sea eficiente se procuró que ninguno de los valores de ΔG fuera menor a -9 Kcal/mole. Todos los cebadores diseñados en este estudio fueron probados en tres muestras distintas de ADN genómico de alpaca. En todos los casos se observó productos de PCR que correspondían a las regiones deseadas. Sin embargo en algunos casos hubo la necesidad de diseñar más de un par de cebadores por gen. En un principio todos los cebadores se diseñaron con la intención de amplificar la región codante y parte de las regiones reguladoras 3'UTR y 5'UTR. Sin embargo debido a la presencia de inserciones/deleciones, en tres genes esto no fue posible. En los genes que presentaron las inserciones/deleciones, aquellos individuos heterocigotos para esta presentaban cromatogramas no informativos. Para solucionar el problema para cada gen se optó por diseñar un nuevo par de cebadores que excluyera las inserciones/deleciones. Al diseñar estos nuevos cebadores se obtuvieron los productos deseados, pudiéndose validar marcadores en los fragmentos de estos genes.

Los KRTAPs se caracterizan en otras especies por ser altamente variables, característica que también se observa en la alpaca. No es raro en esta familia de genes encontrar polimorfismos en cuanto al tamaño. Como se mencionó de los 5 genes con los que se trabajaron en este estudio, tres presentaron

inserciones/deleciones. En el gen KRTAP1-2 se identificó un indel de 30 nt en la región codante. En humano se han identificado varias variantes con respecto al tamaño que pertenecen a la subfamilia KRTAP1. Incluso durante un tiempo todas estas variantes fueron consideradas como distintos genes, hasta que se evaluó los patrones de segregación, donde se estableció que varios de estos “genes” correspondían a variantes de un único gen (Shimomura *et al.*, 2002). Por otro lado en el gen KRTAP6-1 se identificó un indel de 12 nt en la región codante del gen en alpaca. En oveja se ha reportado un indel de 57 pb en KRTAP6-1 en la región codante (Zhou *et al.*, 2014). Asimismo en el gen KRTAP9-2 de alpaca se identificó un indel de 15 nt en región codante. En cabra productora de fibra *cashemere* se ha reportado un indel de 30 nt en la región codante de este gen (Yu *et al.*, 2008). Además de Estos tres genes se han reportado polimorfismos en el tamaño en otros genes pertenecientes a la familia de los KRTAPs (Rogers y Schweizer, 2005). Sin embargo hasta el momento no se han reportado estos tipos de variaciones en los genes KRTAP11-1 y KRTAP13-1 de ninguna especie. En este estudio no se determinaron las frecuencias de estas inserciones/deleciones, únicamente se reporta su presencia. Sería interesante validarlas como marcadores genéticos.

Por otro lado en este estudio se identificaron 32 posiciones polimórficas en las secuencias de los genes analizados en alpaca. En el gen KRTAP1-2 en alpaca, se analizó un fragmento de 894 pb donde se identificaron 10 posiciones polimórficas. Tres de estas posiciones se ubican en la región codante y producen cambios no sinónimos en la proteína, mientras que las 7 restantes se ubican en la región 3'UTR (Tabla 25). Dos de los cambios en la secuencia de proteína (posiciones: 158 y 224)

se dan entre aminoácidos con propiedades fisicoquímicas similares, por lo que la estructura de la proteína no se va a ver muy afectada (Tabla 25). Por otro lado el tercer cambio en la proteína modifica el número de cisteínas, este cambio es importante no solo debido al cambio en las propiedades fisicoquímicas, sino debido a que se ha propuesto que la función de estas proteínas radica en su capacidad de formar puentes disulfuro (Rogers *et al.*, 2006). En el gen KRTAP1-2 en oveja merino se ha reportado un grado de variabilidad similar, se identificaron 10 posiciones polimórficas de las cuales 8 se ubican en la región codante. De estas 3 son no sinónimas, de las cuales una modifica también el número de cisteínas (Gong *et al.*, 2011b). La diferencia en el número de sustituciones en región codante probablemente se debe a que en oveja se analizó toda la región codante mientras que en este estudio solo se trabajó con una porción de esta. Se identificaron 9 haplotipos en este gen en oveja merino, de los cuales solo 6 presentaron frecuencias haplotípicas mayores al 5% (Gong *et al.*, 2014). Un resultado similar se observa en alpaca donde se identificaron 8 haplotipos y 6 de estos presentaron frecuencias mayores al 5% (Tabla 24). En alpaca todas las posiciones presentaron frecuencias alélicas mayores al 5% (Tabla 23). En cuanto a la variabilidad expresada mediante la heterocigosidad, todas las posiciones presentaron valores intermedios a altos (0.3 – 0.5) con excepción de la posición 158 (Tabla 26). Esto también se observó al analizar los valores de PIC (Tabla 26). Por otro lado todas las posiciones se encuentran en EHW (Tabla 26). Estos datos no han sido reportados para los marcadores identificados en otras especies. Las 10 posiciones identificadas en este gen presentan las características que tiene que tener un marcador genético, por lo tanto pueden ser considerados como tales.

En el gen KRTAP6-1 en alpaca se analizó un fragmento de 958 pb donde se identificaron 9 posiciones polimórficas. Todas las posiciones se ubicaron *upstream* de la región codante y de la secuencia de la caja TATA. Además se identificaron 12 haplotipos en el fragmento de este gen de los cuales 6 presentan frecuencias mayores al 5% (Tabla 30). En oveja se han reportado 3 posiciones polimórficas en el gen KRTAP6-1, ninguna en región codante. Además considerando estas posiciones se identificaron dos haplotipos, ambos con frecuencias mayores al 5% (Zhou *et al.*, 2015). La diferencia encontrada entre ambas especies probablemente se debe a la regiones del gen que fueron analizadas, en el estudio de Zhou y colaboradores se trabajó con toda la región codante y parte de la región 3'UTR mientras que en este estudio se trabajó con gran parte de la secuencia *upstream* de la caja TATA y una pequeña porción de la región codante. En ambos casos sin embargo no se identificaron sustituciones polimórficas en la región codante. Por otro lado, en alpaca, solo una de las posiciones (posición 77) presenta una frecuencia menor al 5% (Tabla 29). Los valores de heterocigosidad y PIC son bajos a intermedios y se encuentran alrededor de 0.1 y 0.3 para todas las posiciones, lo cual se explica por la baja presencia de individuos heterocigotos en varias de estas (Tabla 31). No se han encontrado reportes de estos índices en otras especies. También se determinó que 3 posiciones (posiciones 557, 502, 110) no se encuentran en EHW debido a un déficit en heterocigotos (Tabla 31), esto podría deberse a que existe algún tipo de fuerza que está afectando las frecuencias genotípicas. Estas fuerzas que producen la desviación del EHW probablemente sean selección artificial o *inbreeding*, debido a que son animales de crianza para producción de fibra. Sin embargo, otra opción es que se deba a la presencia de alelos nulos, un

estudio determino que en varios casos donde se han reportado marcadores que no están en EHW se debe a la presencia de alelos nulos, en especial aquellos casos donde hay un déficit en heterocigotos (Sen y Burnmeister, 2008). Es importante mencionar que este gen fue el que presento un mayor número de errores en el secuenciamiento. Solo 5 posiciones (212, 369, 506, 612 y 697) cumplen con los criterios para ser considerados como marcadores genéticos en este gen en alpaca.

En el gen KRTAP9-1 se trabajó con un fragmento de 836 pb, donde se identificaron 3 posiciones polimórficas. Dos de estas se ubican en la región codante y producen cambios no sinónimos en la proteína mientras que la restante se encuentra en la región 3'UTR (Tabla 35). Ambos cambios en la secuencia de la proteína se dan entre aminoácidos con propiedades fisicoquímicas diferentes, en este caso varían de acuerdo al tamaño del aminoácido como también en cuanto a la polaridad, probablemente estos cambios afectan la estructura de esta proteína. En dos variedades de cabra de fibra cashmere, se identificó un solo sitio polimórfico en la región codante de este gen (Wang *et al.*, 2012). De las tres posiciones polimórficas en alpaca, una de ellas (posición 598) presento frecuencias alélicas menores al 5% (Tabla 33). Al observar los índices de variabilidad, la posición 86 presenta los valores más altos de heterocigosidad (0.289) y PIC (0.24) (Tabla 36). Las otras dos posiciones presentan valores de heterocigosidad menores a 0.18, los valores de PIC también son bajos en estas dos posiciones (Tabla 36). Los valores de heterocigosidad y PIC reportados para la posición polimórfica en las dos variedades de cabra son más altos, ($H_o = 0.45$; $PIC = 0.35$) que los que se identificaron en alpaca (Wang *et al.*, 2012). En alpaca las tres posiciones se encuentran en EHW (Tabla 36), mientras que en cabra la posición identificada no lo está (Wang *et al.*,

2012). Tomando en consideración las características que debe tener un marcador genético, 2 de las posiciones polimórficas identificadas en alpaca pueden ser consideradas como tales.

En el gen KRTAP11-1 se trabajó con un fragmento de 1145 pb donde se identificaron 6 posiciones polimórficas. Cuatro de estas se ubican en región codante y producen cambios no sinónimos en la proteína, donde 2 sustituciones nucleotídicas modifican el número de cisteínas en la proteína (Tabla 41). Además en alpaca se identificaron 3 haplotipos distintos con frecuencias mayores al 5% (Tabla 40). Este mismo gen fue estudiado en oveja donde se identificaron 5 posiciones polimórficas en la región codante, de las cuales solo una produce un cambio en la secuencia de la proteína. Además se identificaron 6 haplotipos diferentes (Gong *et al.*, 2011a). Estos resultados estarían indicando una mayor variabilidad en KRTAP11-1 en oveja. Todas las posiciones polimórficas presentan frecuencias alélicas mayores al 5% en alpaca (Tabla 39). Los valores de heterocigosidad son altos, todos por encima de 0.39 menos la posición 956 que presenta un valor de 0.256 (Tabla 42). Los valores de PIC siguen un mismo patrón (Tabla 42). Por otro lado las 6 posiciones se encuentran en EHW (Tabla 42). No se reportan los índices de variabilidad para estas posiciones en oveja. Considerando estos valores, las 6 posiciones identificadas en el gen KRTAP11-1 en alpaca pueden ser consideradas como marcadores genéticos del tipo SNP.

En el gen KRTAP13-1 se analizó un fragmento de 826 pb donde se identificaron 4 posiciones polimórficas. Tres de estas se ubican en la región codante de gen y corresponden a sustituciones no sinónimas, dos de las cuales producen modificaciones en el número de cisteínas (Tabla 47). En tres variedades de cabra

cashmere se identificó una posición polimórfica en región codante, la cual produce cambios en la secuencia de la proteína (Fang *et al.*, 2010). En alpaca todas las posiciones presentan frecuencias alélicas iguales y mayores al 5% (Tabla 45). Los valores de heterocigosidad y PIC fueron altos en todas las posiciones, 0.45 y 0.37 respectivamente (Tabla 48). Estos valores son muy similares a los reportados en una de las variedades de cabra, donde se observaron valores de heterocigosidad de 0.46 y valores de PIC de 0.37 (Fang *et al.*, 2010). En alpaca todas las posiciones se encuentran en EHW (Tabla 48), mientras que en cabra las posición identificada no se encuentra en EHW en las tres variedades (Fang *et al.*, 2010). Estos resultados guardan relación con lo observado en el gen KRTAP9-2, donde en alpaca se pudo identificar una mayor variabilidad que en cabra. Tomando en cuenta estos resultados, las 4 posiciones identificadas en el gen KRTAP13-1 en alpaca pueden ser consideradas como marcadores genéticos del tipo SNP.

Además de evaluar si los marcadores dentro de un mismo gen se encuentran en desequilibrio de ligamiento, el mismo test se realizó entre las 32 posiciones identificadas. Para disminuir la presencia de falsos positivos debido al gran número de test realizados (496) se utilizó la corrección de Benjamini-Hochberg, la cual resulta menos conservadora que la corrección de Bonferroni (Benjamini y Hochberg, 1995). Al ajustar los datos se obtuvo que además de los marcadores dentro de cada gen, 28 pares están en desequilibrio de ligamiento (Tabla 58). De estos 28 pares, 18 corresponden a los genes KRTAP1-2 y KRTAP9-2. Los pares de posiciones de estos genes que no se encuentran en desequilibrio de ligamiento coincidentemente corresponden a aquellas posiciones que presentan las frecuencias alélicas más bajas o presentan muy pocos individuos homocigotos para un alelo (Tabla 51). Debido a

que se aplicó un *likelihood ratio test* (LTR), este es sensible al número de muestra cuando las frecuencias son bajas (Barnes y Gray, 2003). Esto indicaría que ambos genes están en desequilibrio de ligamiento. Se ha visto en otras especies que estos dos genes son parte de un mismo *cluster* en un mismo cromosoma por lo tanto es muy probable que estén ligados (Khan *et al.*, 2014). Por otro lado un estudio identificó la posición de diversos marcadores en los cromosomas de alpaca. En este estudio el gen IKZ3 presente en el *scaffold* 108 se ubica en el cromosoma 16 (Avila *et al.*, 2014). El gen KRTAP1-2 también se ubica en este mismo *scaffold* y por lo tanto debe encontrarse en el cromosoma 16 también. Teniendo en consideración esto y el hecho que el gen KRTAP1-2 se encuentra en desequilibrio de ligamiento con el gen KRTAP9-2, es probable que también este gen se ubique en este cromosoma. Los 10 pares restantes corresponden a los genes KRTAP6-1 y KRTAP11-1, esto era de esperarse debido a que ambos genes se ubican dentro del mismo *scaffold* en el genoma de referencia. Además en otras especies se ha observado que ambos genes, como también el gen KRTAP13-1, pertenecen al mismo *cluster* (Khan *et al.*, 2014). En el estudio de Avila y colaboradores se determinó la posición del gen UCN3 que se ubica en el *scaffold* 101, este gen está en el cromosoma 35 (Avila *et al.*, 2014). El gen KRTAP13-1 se ubica en el *scaffold* 101 y por lo tanto podría estar en el cromosoma 35. Debido a que en otras especies el gen KRTAP13-1 se encuentra en el mismo *cluster* con los genes KRTAP6-1 y KRTAP11-1, es probable que estos tres genes se ubiquen en el cromosoma 35 de alpaca. Sin embargo en este estudio no se presenta evidencia de las posiciones de estos genes en los cromosomas de alpaca, es necesario estudios posteriores para confirmar esta información.

En el primer componente de este estudio se identificaron 27 marcadores informativos de tipo SNP en 5 genes que codifican proteínas estructurales de fibra (KRTAPs) en alpaca. La metodología utilizada demostró ser útil, sin embargo hoy en día existen métodos más eficientes que permiten la caracterización de un número considerablemente mayor de marcadores SNP. Uno de estos métodos se denomina *genotyping by sequencing* (GBS), el cual permite la identificación y genotipificación de marcadores SNPs al mismo tiempo (Elshire *et al.*, 2011). Esta metodología ha resultado bastante útil en especies como el maíz donde se identificaron 200 000 marcadores (Elshire *et al.*, 2011). Esta tecnología se puede aplicar en especies que tienen un genoma secuenciado como especies que no lo tienen. Sin embargo es importante mencionar que utilizando esta metodología es más complejo identificar marcadores en genes específicos, por lo cual metodologías como la utilizada en este estudio son muy útiles también.

Los procesos de selección genética requieren de la identificación de regiones del genoma asociados a características de tipo productivas (QTLs), para lo cual es necesario contar con una gran cantidad de marcadores genéticos que permitan establecer la asociación fenotipo – genotipo (Mackay *et al.*, 2009). Esta asociación permite seleccionar individuos reproductores que tengan el efecto deseado sobre la característica que se desea mejorar.

Para el mapeo de QTLs existen hoy en día diversas aproximaciones, la que se utilizó en este estudio se denomina gen candidato. Consiste en seleccionar uno o varios genes de los cuales se tenga información a priori que están relacionados de alguna forma con la característica de interés (Tabor *et al.*, 2002). Otro tipo de aproximación muy utilizada es el *Genome Wide Association Studies* (GWAS), esta consiste en la

utilización de un gran número de marcadores que cubren gran parte del genoma (Cardon y Bell 2001). La ventaja de una aproximación de gen candidato es que por un lado en vez de utilizar marcadores al azar se seleccionan marcadores teniendo en cuenta la biología de la característica, además que esta aproximación no necesita de un número extenso de marcadores (Tabor *et al.*, 2002). Para este estudio se seleccionó esta aproximación debido a que existen muchos reportes de asociaciones entre la familia de genes seleccionada y diversas características de fibra. Si bien estos reportes son en otras especies, el folículo es un tejido muy conservado en el taxón de los mamíferos. Además el número de marcadores que han sido caracterizados en alpaca es muy bajo para utilizar otro tipo de aproximación.

Para la evaluación de la asociación de uno de estos 5 genes al diámetro de fibra, se eligió como gen candidato al gen KRTAP11-1. Se seleccionó este gen debido a varios factores. Este gen es el que presenta un mayor número de marcadores en región codante (4), además 2 de estos marcadores producen cambios significativos para la función de la proteína involucrando dos residuos de cisteína (Tabla 41). Por otro lado junto con el gen KRTAP13-1, tiene los marcadores más informativos y con las frecuencias alélicas más altas (Tablas 39; 45). Por otro lado este gen ha demostrado tener un rol muy importante en la formación de la fibra. Un estudio llevado a cabo por Fujimoto y colaboradores se enfocó en la relevancia que tiene KRTAP11.1 para la formación de pelo en ratones. Este estudio determinó que KRTAP11.1 tiene una regulación estricta a nivel post-traducciona l y que una sobreexpresión de este gen imposibilitaba un correcto ensamblaje de filamentos intermedios de queratina. En ratones *Foxn1*^{nu} que tienen una mutación en el gen *FOXN1*. *FOXN1* codifica un factor de transcripción que regula la expresión de un

cluster de genes KRTAPs y queratinas. Estos ratones presentan un pelo más delgado y rizado, que está asociado a un patrón de expresión anómalo de kRTAP11.1 (Fujimoto *et al.*, 2014). Además en un estudio de proteómica de fibra en oveja se determinó que este gen posee una alta y constante expresión durante toda la formación de la fibra (Plowman *et al.*, 2015). Además es una de las pocas subfamilias que solo consta de un único miembro y está presente en todas las especies de mamíferos donde se han estudiado esta familia de genes (Khan *et al.*, 2014).

Cuando se realiza un estudio de asociación lo ideal es trabajar con una muestra grande, que tenga características homogéneas, que no esté relacionada y que la variable de interés siga una distribución normal (Balding, 2006; Foulkes, 2009). Si todos estos criterios se cumplen en la muestra, la asociación puede evaluarse mediante una regresión lineal simple o múltiple dependiendo de la presencia de co-variables (Foulkes, 2009). En caso alguno de estos criterios no se cumplan se pueden realizar correcciones o utilizar otras alternativas como por ejemplo modelos lineales mixtos que incluyen factores aleatorios (Yu *et al.*, 2006). Con el fin de garantizar una muestra homogénea todas las muestras provienen del mismo centro de crianza en Puno. Solo se utilizaron alpacas de la variedad huacaya y que produzcan fibra de color blanco. Estas alpacas además cuentan con el genotipo para 10 marcadores del tipo microsatelite y una ficha técnica con información sobre su peso, edad, sexo, presencia de parásitos entre otras características.

Una muestra para un estudio de asociación puede no ser independiente por dos motivos distintos. Un caso se da cuando existen relaciones de parentesco dentro de la muestra como por ejemplo una madre y una cría. El otro caso se da cuando hay

presencia de estructuración poblacional, en este caso la muestra puede ser subdividida en pequeños grupos que comparten ciertos alelos específicos. Es importante controlar ambos factores, de lo contrario esto puede llevar a la presencia de falsos positivos (Foulkes, 2009; Balding, 2006). Para determinar si existían individuos dentro de la muestra con relaciones de parentesco (madre – cría y padre – cría) se realizó un test de parentesco utilizando los genotipos de los 10 microsatelites. El panel de marcadores que se utilizó constituye una prueba de parentesco desarrollada y probada en la UBM. Se esperaba que no hubiera muchas relaciones de este tipo, dado que la muestra de 199 individuos fue colectada al azar de una población de 40000 alpacas. Solo se tomaron en cuenta aquellas muestras que presentaran un LOD score alto y presentaran un mismatch en los genotipos como máximo. Este mismatch se puede deber a un error de secuenciamiento. Se identificaron 73 pares (padre – cría o madre – cría), en algunos casos individuos se repetían por lo tanto solo fue necesario retirar de la muestra 47 individuos. Es importante mencionar que en algunos casos ambos individuos (padre/madre – cría) tenían la misma edad, lo cual hace que esta relación sea imposible. Aun así para evitar complicaciones en estos casos uno de los individuos era eliminado del estudio. Debido a que se usó un panel de marcadores que constituyen una prueba de paternidad validada se asume que la muestra restante no está relacionada a un nivel de confianza del 95%. Para determinar la presencia de estructuración poblacional existen diversos métodos confiables para utilizar, uno de estos es el algoritmo utilizado por el programa STRUCTRE. Los parámetros se seleccionaron de acuerdo a las recomendaciones presentadas en el manual de usuario. No se identificó la presencia de estructuración poblacional dentro de la muestra (Figura

41). Estos dos resultados permitieron la selección de una muestra independiente para el estudio de asociación.

Otro criterio importante para poder evaluar la asociación mediante una regresión lineal es que la variable dependiente presente una distribución normal, de lo contrario se tendría que utilizar una prueba no paramétrica. Para evaluar la distribución del diámetro de fibra en la muestra se aplicó la prueba de Shapiro-Wilk, de acuerdo a esta prueba el diámetro de fibra dentro de la muestra sigue una distribución normal ($p = 0.054$).

Otro criterio importante a considerar a la hora de realizar una prueba de asociación es si el tamaño de la muestra es lo suficientemente grande para garantizar un poder estadístico aceptable (Balding, 2006). El poder estadístico de una prueba se define como la probabilidad de rechazar la hipótesis nula, siendo la hipótesis alterna verdadera, generalmente el valor aceptable es de 0.8. Debido a la ausencia de métodos para determinar el poder estadístico se optó por realizar simulaciones de poblaciones que presentaran distintas frecuencias alélicas y distintos efectos del gen sobre la variable dependiente. El efecto del gen se determinó mediante el cálculo de heredabilidad en sentido estricto utilizando las formulas presentadas en el libro de Falconer y Mackay (Falconer y Mackay, 1996). Para cada población se evaluó mediante una regresión lineal múltiple (incluyendo sexo) si el gen se encontraba asociado a la característica de interés, luego se dividió el número de veces en las que el P-value era menor a 0.05 entre el total de evaluaciones. De acuerdo al resultado que se obtuvo luego de realizar las simulaciones, tomando en cuenta una muestra de 152 individuos, el modelo permite detectar efectos de hasta un 8% sobre

la varianza total con un poder estadístico de 0.8 cuando el marcador que se utilice presenta una frecuencia mínima de 0.4 (Figura 42).

Para evaluar si el gen KRTAP11-1 está asociado al diámetro de fibra se utilizó el genotipo de 2 de los 6 marcadores (SNP330 y SNP 956) validados en este gen. Solo se utilizaron estos dos marcadores, porque la combinación de estos marcadores representa toda la variabilidad identificada para este gen. Debido a que el SNP330 se encuentra en perfecto desequilibrio de ligamiento con los 4 marcadores no utilizados. Además se incluyó al modelo como covariables la edad y el sexo. Varios autores han reportado que la edad tiene un efecto importante sobre el diámetro de la fibra. En cuanto al sexo los reportes son un poco más ambiguos, hay autores que afirman que hay efecto del sexo sobre el diámetro mientras que hay otros que rechazan esta afirmación. No se incluyó data sobre nutrición debido a que estos animales provienen de un mismo criadero, se asume que todas las alpacas se alimentan del mismo pastizal. Además de acuerdo a las fichas técnicas ninguna alpaca sufría de desnutrición. Tampoco se incluyó data acerca de los parásitos en el modelo, debido a que las 152 alpacas presentaban sarna por lo tanto en este aspecto no habría diferencia entre los individuos. No se tenía información sobre el grado de parasitosis de cada animal.

Al ajustar la data al modelo propuesto se determinó que ninguno de los marcadores del gen KRTAP11-1 evaluados tiene un efecto significativo sobre el diámetro de fibra ($P_{330} = 0.8$ y $P_{956} = 0.723$) (Tabla 52). Es importante mencionar que en el caso del SNP956 la muestra utilizada no presenta un poder estadístico adecuado debido a que este marcador presenta un alelo con una frecuencia de 0.12 (Figura 42). Por lo tanto para poder evaluar el efecto de este marcador sobre el diámetro se debería

ampliar el número de muestra. En el caso del SNP330 si se alcanzó el poder estadístico deseado debido a que este marcador presenta frecuencias alélicas alrededor de 0.5. Los resultados indican que el gen KRTAP11-1 no tiene un efecto significativo (más del 8%) sobre el diámetro de fibra. Sin embargo debido a la importancia que se le atribuye a este gen en la formación de la fibra aun es un buen candidato a estar asociado a otras características como por ejemplo producción de fibra. En cuanto a la edad se determinó que si tiene un efecto significativo sobre el diámetro de fibra, como había sido reportado por otros autores ($P = 0.01$) (Tabla 52) (Quispe *et al.*, 2013). Mientras más vieja es la alpaca el diámetro tiende a engrosarse, esta relación se cree que está asociada con el número de esquilas al cual el animal es sometido. Con respecto al sexo tampoco se observó un efecto significativo sobre el diámetro de fibra ($P = 0.222$) (Tabla 52), sin embargo es importante mencionar que la ausencia de significancia pueda deberse a la diferencia entre el número de machos y hembras en el estudio.

En este estudio se validaron 27 marcadores informativos en 5 genes de una de las principales familias de proteínas estructurales en fibra de alpaca. Por otro lado este estudio ratifica lo encontrado en otros estudios de esta familia de genes; que los KRTAPs son genes con una alta variabilidad. Además, los tests de desequilibrio de ligamiento estarían indicando que la constitución de los *clusters* en alpaca es el mismo que en otras especies, sin embargo es necesario más estudios para confirmar esta información. También al realizar un estudio de asociación del gen KRTAP11-1 al diámetro de fibra, no se identificó una asociación significativa, esto no excluye la posibilidad que este gen no este asociado a otro tipo de características comerciales en alpaca. Este estudio es el primero en reportar marcadores

informativos en genes estructurales de la fibra en alpaca y constituye un paso importante en la investigación de los factores genéticos que afectan las propiedades de la fibra en este animal.

VII. Conclusiones

- Se caracterizaron *in silico* los genes KRTAP1-2, KRTAP6-1, KRTAP9-2, KRTAP11-1 y KRTAP13-1 en alpaca.
- Se identificaron 32 posiciones polimórficas en los genes KRTAP1-2 (10), KRTAP6-1 (9), KRTAP9-2 (3), KRTAP11-1 (6) y KRTAP13-1 (4) en alpaca.
- Se identificaron polimorfismos de tamaño en los genes KRTAP1-2, KRTAP6-1 y KRTAP9-2.
- 27 sitios polimórficos fueron validados como marcadores genéticos del tipo SNP en los genes KRTAP1-2 (10), KRTAP6-1 (5), KRTAP9-2 (2), KRTAP11-1 (6) y KRTAP13-1 (4).
- La familia de los genes KRTAP en alpaca ha demostrado ser muy variable, presentando polimorfismos puntuales y de tamaño, al igual que en otras especies.
- El gen KRTAP11-1 no tiene un efecto significativo sobre el diámetro de fibra en alpaca.

VIII. Recomendaciones

- Evaluar con un número de muestra más grande la asociación entre el SNP956 del gen KRTAP11-1 y el diámetro de fibra.
- Evaluar la asociación de los genes KRTAP1-2, KRTAP6-1, KRTAP9-2 y KRTA13-1 al diámetro de fibra como también a otras características de valor comercial.
- Identificar más marcadores en genes estructurales de fibra (KRTAPs y queratinas), así como en genes reguladores de los procesos biológicos del folículo en alpaca.
- Utilizar metodologías como GBS para la identificación de un gran número de marcadores esto permitirá realizar estudios de asociación del tipo GWAS, cubriendo gran parte del genoma de la alpaca.

IX. Referencias Bibliográficas

- Adams, N. R., Cronjé, P. B. (2003). A review of the biology linking fibre diameter with fleece weight, liveweight, and reproduction in Merino sheep. *Crop and Pasture Science*, 54(1), 1-10.
- Almeida, A. M., Plowman, J. E., Harland, D. P., Thomas, A., Kilminster, T., Scanlon, T., Clerens, S. (2014). Influence of feed restriction on the wool proteome: a combined iTRAQ and fiber structural study. *Journal of proteomics*, 103, 170-177.
- Antonini, M., Gonzales, M., Valbonesi, A. (2004). Relationship between age and postnatal skin follicular development in three types of South American domestic camelids. *Livestock Production Science*, 90(2), 241-246.
- Avila, F., Baily, M. P., Perelman, P., Das, P. J., Pontius, J., Chowdhary, R., Raudsepp, T. (2014). A comprehensive whole-genome integrated cytogenetic map for the alpaca (*Lama pacos*). *Cytogenetic and genome research*, 144(3), 196-207.
- Aylan-Parker, J., McGregor, B. A. (2002). Optimising sampling techniques and estimating sampling variance of fleece quality attributes in alpacas. *Small Ruminant Research*, 44(1), 53-64.
- Badajoz, L., García, V., Pezo, C. (2009). Descripción histológica del complejo folicular piloso en crías de alpacas. *Revista de Investigaciones Veterinarias del Perú*, 20(2), 154-164.
- Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*, 7(10), 781-791.
- Barnes, M. R., Gray, I. C. (Eds.). (2003). *Bioinformatics for geneticists*. John Wiley & Sons
- Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.
- Berk, A. (2000). *Molecular cell biology* (Vol. 4). New York: WH Freeman.
- Bustamante A.V., Mat M. L. , Zambelli A., Vidal-Rioja L. (2003) Isolation and characterization of 10 polymorphic dinucleotide microsatellite markers for llama and guanaco. *Molecular ecology Notes*, 3(1):68-69

- Cardon, L. R., Bell, J. I. (2001). Association study designs for complex diseases. *Nature Reviews Genetics*, 2(2), 91-99.
- Cervantes, I., Pérez-Cabal, M. A., Morante, R., Burgos, A., Salgado, C., Nieto, B., Gutiérrez, J. P. (2010). Genetic parameters and relationships between fibre and type traits in two breeds of Peruvian alpacas. *Small Ruminant Research*, 88(1), 6-11.
- Collins D.W., Jukes T.H., Rates of transition and transversion in coding sequences since the human-rodent divergence, *Genomics* 20 (1994) 386-396.
- Cordero, F., Contreras, P., Mayhua, M., Jurado, E., & Castrejón, V. (2011). Correlaciones fenotípicas entre características productivas en alpacas Huacaya. *Revista de Investigaciones Veterinarias del Perú*, 22(1), 15-21.
- Crossley, J. C., Borroni, C. G., Raggi, A. S. (2014). Correlation between mean fibre diameter and total follicle density in alpacas of differing age and colour in the Parinacota province of the Chilean high plain. *Journal of Applied Animal Research*, 42(1), 27-31.
- Cruz, A., Cervantes, I., Burgos, A., Morante, R., Gutiérrez, J. P. (2015). Estimation of genetic parameters for reproductive traits in alpacas. *Animal reproduction science*, 163, 48-55.
- Cruz, L. (2011). *Estimación de parámetros genéticos para caracteres productivos en alpacas (Vicugna pacos), Perú 2011* (Magister dissertation).
- Czaplicki, Z. (2012). Properties and Structure of Polish Alpaca Wool. *Institute of Natural fibres & Medicinal plant*. 1(90), 8–12.
- De Los Ríos, E. (2006). Producción textil de fibras de camélidos sudamericanos en el área altoandina de Bolivia, Ecuador y Perú. Organización de las Naciones Unidas para el Desarrollo Industrial (UNIDO).
- Elliott, K., Stephenson, T. J., Messenger, A. G. (1999). Differences in hair follicle dermal papilla volume are due to extracellular matrix volume and cell number: implications for the control of hair follicle size and androgen responses. *Journal of investigative dermatology*, 113(6), 873-877.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, 6(5), e19379.
- Evanno, G., Regnaut, S., Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, 14(8), 2611-2620.
- Excoffier, L., Slatkin, M. (1998). Incorporating genotypes of relatives into a test of linkage disequilibrium. *The American Journal of Human Genetics*, 62(1), 171-180.
- Falconer, D. S., Mackay, T. F., & Frankham, R. (1996). Introduction to quantitative genetics (4th edn). *Trends in Genetics*, 12(7), 280.
- Fang, Y., Liu, W. J., Zhang, F. Q., Shao, Y. G., Yu, S. G. (2010). The polymorphism of a novel mutation of KAP13. 1 gene and its associations with cashmere traits on

Xinjiang local goat breed in China. *Asian Journal of Animal and Veterinary Advances*, 5(1), 34-42.

FAO - Organización de las Naciones Unidas para la Agricultura y la Alimentación. (2005) Situación actual de los camélidos sudamericanos en Perú. Proyecto de cooperación técnica en apoyo a la crianza y aprovechamiento de los camélidos sudamericanos en la región andina (TCP/RLA/2914):1-62.

Feeley, N. L., Munyard, K. A. (2009) Correlation of single nucleotide polymorphisms in the melanocortin-1 receptor gene with colour in Vicugna pacos (Alpaca). *School of Biomedical Sciences, Curtin University of Technology, Perth, Western Australia*, 6845, 49(8), 675–681.

Feral J.P. (2002) How useful are the genetic markers in attempts to understand and manage marine biodiversity? *Journal of Experimental Marine Biology and Ecology*, 268:121-145.

Ferguson, M. B., McGregor, B. A., Behrendt, R. (2012). Relationships between skin follicle characteristics and fibre properties of Suri and Huacaya alpacas and Peppin Merino sheep. *Animal production science*, 52(7), 442-447.

Fernández Ugaz, D. A. (2015). Búsqueda de genes relacionados a la síntesis de la fibra y marcadores SSR en los ESTs de piel de alpaca Vicugna pacos. Tesis UNMSM.

Foulkes, A. S. (2009). Applied statistical genetics with R: for population-based association studies. *Springer Science & Business Media*.

Franco, F., San Martín, H., Ara, G., Olazábal, L., Carcelén, C. (2009). Efecto del nivel alimenticio sobre el rendimiento y calidad de fibra en alpacas. *Revista de Investigaciones Veterinarias del Perú*, 20(2), 187-195.

Fraser, R. D. B., MacRae, T. P., Rogers, G. E. (1972). *Keratins: their composition, structure, and biosynthesis*. Charles C. Thomas.

Fujimoto, S., Takase, T., Kadono, N., Maekubo, K., Hirai, Y. (2014). Krtap11-1, a hair keratin-associated protein, as a possible crucial element for the physical properties of hair shafts. *Journal of Dermatological Science*, 74(1), 39-47.

Galtier, N., Nabholz, B., Glémin, S., Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular ecology*, 18(22), 4541–50.

Gillespie, J. M. (1972). Proteins rich in glycine and tyrosine from keratins. *Comparative biochemistry and physiology. B, Comparative Biochemistry*, 41(4), 723–34.

Gillespie, J. M., Broad, A. (1972). Ultra-high-sulphur proteins in the hairs of the artiodactyla. *Australian journal of biological sciences*, 25(1), 138–45.

Gillet, E. M. (1999) Minimum sample sizes for sampling genetic marker distributions. In: Gillet, E. M. (ed.), Sampling strategies for marker analysis. *Final compendium of the research project development, optimisation and validation of*

molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research programme molecular tools for biodiversity.

Gong, H., Zhou, H., Dyer, J. M., Hickford, J. G. H. (2011a). Identification of the ovine KAP11-1 gene (KRTAP11-1) and genetic variation in its coding sequence. *Molecular biology reports*, 38(8), 5429–33. doi:10.1007/s11033-011-0697-2

Gong, H., Zhou, H., Hodge, S., Dyer, J. M., Hickford, J. G. (2015). Association of wool traits with variation in the ovine KAP1-2 gene in Merino cross lambs. *Small Ruminant Research*, 124, 24-29.

Gong, H., Zhou, H., Yu, Z., Dyer, J., Plowman, J. E., & Hickford, J. (2011b). Identification of the ovine keratin-associated protein KAP1-2 gene (KRTAP1-2). *Experimental dermatology*, 20(10), 815-819.

Goto, Y., Yue, L., Yokoi, A., Nishimura, R., Uehara, T., Koizumi, S., & Saikawa, Y. (2001). A novel single-nucleotide polymorphism in the 3'-untranslated region of the human dihydrofolate reductase gene with enhanced expression. *Clinical Cancer Research*, 7(7), 1952-1956.

Gregorius, H.-G. (1980) The probability of losing an allele when diploid genotypes are sampled. *Biometrics*, 36, 643-652.

Grigg, G. C., Beard, L. A., Augee, M. L. (2004). The evolution of endothermy and its diversity in mammals and birds. *Physiological and Biochemical Zoology*, 77(6), 982-997.

Guridi, M., Soret, B., Alfonso, L., Arana, A. (2011). Single nucleotide polymorphisms in the Melanocortin 1 Receptor gene are linked with lightness of fibre colour in Peruvian Alpaca (*Vicugna pacos*). *Animal genetics*, 42(6), 679–82.

Hamilton, M. (2009). *Population genetics*. First Edition. Great Britain. Wiley - Blackwell. Chapter 2, pp. 13; pp. 41; Chapter 9, pp. 283.

Handjiski, B. K., Eichmüller, S., Hofmann, U., Czarnetzki, B. M., Paus, R. (1994). Alkaline phosphatase activity and localization during the murine hair cycle. *British journal of dermatology*, 131(3), 303-310.

Hardy, M. H., Lyne, A. G. (1956). The pre-natal development of wool follicles in Merino sheep. *Australian Journal of Biological Sciences*, 9(3), 423-441.

Havill, L. M., Dyer, T. D., Richardson, D. K., Mahaney, M. C., Blangero, J. (2005). The quantitative trait linkage disequilibrium test: a more powerful alternative to the quantitative transmission disequilibrium test for use in the absence of population stratification. *BMC genetics*, 6(Suppl 1), S91.

Hesse, M., Zimek, A., Weber, K., Magin, T. M. (2003). Comprehensive analysis of keratin gene clusters in humans and rodents. *European Journal of Cell Biology*, 83(1), 19–26.

Höcker, H. (2002) Fibre morphology. In: *Wool: Science and technology*. (Edited by Simpson W. S. and Crawshaw G. H.), pp. 60-79. Woodhead Publishing Limited, Cambridge, England.

Hwang, K. A., Hwang, Y. L., Lee, M. H., Kim, N. R., Roh, S. S., Lee, Y Choi, K. C. (2012). Adenosine stimulates growth of dermal papilla and lengthens the anagen phase by increasing the cysteine level via fibroblast growth factors 2 and 7 in an organ culture of mouse vibrissae hair follicles. *International journal of molecular medicine*, 29(2), 195-201.

Ian William, P., Ian Robert, F. (2005). Major genes and QTL influencing wool production and quality: a review. *Genet. Sel. Evol*, 37(1), S97-S107.

Itenge-Mweza, T. O., Forrest, R. H. J., McKenzie, G. W., Hogan, A., Abbott, J., Amofo, O., Hickford, J. G. H. (2007). Polymorphism of the KAP1.1, KAP1.3 and K33 genes in Merino sheep. *Molecular and cellular probes*, 21(5-6), 338-42.

Jacobson, E. M., Concepcion, E., Oashi, T., Tomer, Y. (2005). A Graves' disease-associated Kozak sequence single-nucleotide polymorphism enhances the efficiency of CD40 gene translation: a case for translational pathophysiology. *Endocrinology*, 146(6), 2684-2691.

Jahoda, C. A., Reynolds, A. J., Oliver, R. F. (1993). Induction of hair growth in ear wounds by cultured dermal papilla cells. *Journal of investigative dermatology*, 101(4), 584-590.

Jones, L. N., Rivett, D. E. (1997). The role of 18-methyleicosanoic acid in the structure and formation of mammalian hair fibres. *Micron*, 28(6), 469-485.

Kadwell, M., Fernandez, M., Stanley, H. F., Baldi, R., Wheeler, J. C., Rosadio, R., Bruford, M. W. (2001). Genetic analysis reveals the wild ancestors of the llama and the alpaca. *Proceedings of the Royal Society of London B: Biological Sciences*, 268(1485), 2575-2584.

Kaufman, C. K., Zhou, P., Pasolli, H. A., Rendl, M., Bolotin, D., Lim, K. C., Fuchs, E. (2003). GATA-3: an unexpected regulator of cell lineage determination in skin. *Genes & development*, 17(17), 2108-2122.

Khan, I., Maldonado, E., Vasconcelos, V., O'Brien, S. J., Johnson, W. E., Antunes, A. (2014). Mammalian keratin associated proteins (KRTAPs) subgenomes: disentangling hair diversity and adaptation to terrestrial and aquatic environments. *BMC genomics*, 15(1), 1.

Langbein, L., Rogers, M. A, Winter, H., Praetzel, S., Schweizer, J. (2001). The catalog of human hair keratins. II. Expression of the six type II members in the hair follicle and the combined catalog of human type I and II keratins. *The Journal of biological chemistry*, 276(37), 35123-32.

Langbein, L., Rogers, M. A., Winter, H., Praetzel, S., Beckhaus, U., Rackwitz, H., Schweizer, J. (1999). The Catalog of Human Hair Keratins. I. Expression of the nine type I members in the hair follicle. *The journal of biological chemistry*. Vol. 274, No. 28, Issue of July 9, pp. 19874-19884.

- León-Velarde C.U, Guerrero J. 2001. Improving quantity and quality of Alpaca fiber; using simulation model for breeding strategies. <http://inrm.cip.cgiar.org/home/publicat/01cpb023.pdf>.
- Lupton C.J., McColl A. Stobart R.H. 2006. Fiber characteristics of the Huacaya Alpaca. *Small Rumin. Res.*, 64: 211-224.
- Mackay, T. F., Stone, E. A., Ayroles, J. F. (2009). The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics*, 10(8), 565-577.
- Mamani, G. (1991). Parámetros genéticos del peso vivo y vellón en alpacas Huacaya de La Raya, Puno. *VII Conv. Int. Especialistas en Cam. Sud., Jujuy, Argentina*.
- Marin, J. C., Zapata, B., Gonzalez, B. A., Bonacic, C., Wheeler, J. C., Casey, C., Spotorno, Á. E. (2007). Sistemática, taxonomía y domesticación de alpacas y llamas: nueva evidencia cromosómica y molecular. *Revista chilena de historia natural*, 80(2), 121-140.
- Martínez-Arias, R., Calafell, F., Mateu, E., Comas, D., Andrés, A, Bertranpetit, J. (2001). Sequence variability of a human pseudogene. *Genome research*, 11(6), 1071–85.
- Mc Elwee, K. J., Kissling, S., Wenzel, E., Huth, A., & Hoffmann, R. (2003). Cultured peribulbar dermal sheath cells can induce hair follicle development and contribute to the dermal sheath and dermal papilla. *Journal of investigative dermatology*, 121(6), 1267-1275.
- McGregor B.A. y Butler K.L. 2004. Sources of variation in fibre diameter attributes of Australian alpacas and implications for fleece evaluation and animal selection. *Aust. J. Agric. Res.*, 55: 433-442.
- McKenzie, G. W., Abbott, J., Zhou, H., Fang, Q., Merrick, N., Forrest, R., H. Hickford, J. G. (2010). Genetic diversity of selected genes that are potentially economically important in feral sheep of New Zealand. *Genet Sel Evol*, 42, 43.
- Mcpartlan, H. C., Matthews, M. E., Robinson, N. A., (1998) Alpaca microsatellites at the VIAS A1 and VIAS A2 loci. *Animal Genetics*, 29:158.
- Moore, K. E., Maloney, S. K., Blache, D. (2015). High follicle density does not decrease sweat gland density in Huacaya alpacas. *Journal of thermal biology*, 47, 1-6.
- Morante, R., Goyache, F., Burgos, A., Cervantes, I., Pérez-Cabal, M. A., Gutiérrez, J. P. (2009). Genetic improvement for alpaca fibre production in the Peruvian Altiplano: the Pacamarca experience. *Animal Genetic Resources Information*, 45, 37-43.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70(12), 3321–3.
- Nelson D. L., Cox M. M. (2009) *Lehninger principles of biochemistry*. W.H. Freeman and Company. New York. Fifth Edition: 288 – 292.

Obreque, V., Coogle, L., Henney, P., Baileye, E., Mancilla, R., Garcia-Huidobro, J., Hinrichsen, P., Cothran, E. (1998) Characterisation of 10 polymorphic alpaca dinucleotide microsatellites. *Animal Genetics*, 29:461-462.

Orwin, D. F. (1979). The cytology and cytochemistry of the wool follicle [Includes fiber development and properties]. *International review of cytology*.

Panigrahi, S., & Kushwaha, R. L. (2013). Alpaca Fibre Processing Methodology. *University of Saskatchewan*.

Paredes-Peralta, M. M., Alonso-Moraga, A., Analla, M., Machaca-Centty, J., Munoz-Serrano, A. (2011). Genetic parameters and fixed effects estimation for fibre traits in alpaca huacaya (*Lama pacos*). *Journal of Animal and Veterinary Advances*, 10(11), 1484-1487.

Parry, D. A D., Smith, T. A, Rogers, M. A, Schweizer, J. (2006). Human hair keratin-associated proteins: sequence regularities and structural implications. *Journal of Structural Biology*, 155(2), 361–9.

Parsons, Y. M., Cooper, D. W., Piper, L. R. (1994). Evidence of linkage between high-glycine-tyrosine keratin gene loci and wool fibre diameter in a Merino half-sib family. *Animal genetics*, 25(2), 105-108.

Paus, R., Cotsarelis, G. (1999). The biology of hair follicles. *N Engl J Med*, 341(7), 491-497.

Penedo M.C., Caetano A.R., Cordova K . (1999) Eight microsatellite markers for South American camelids. *Animal Genetics*, 30(2):166-7 61.

Penedo M.C., Caetano A.R., Cordova K.I. (1998) Microsatellite markers for South American camelids. *Animal Genetics*, 29(5):411-2 62.

Penedo M.C., Caetano A.R., Cordova K.I. (1999) Six microsatellite markers for South American camelids. *Animal Genetics*, 30(5):399

Picoult-newberg, L., Ideker, T. E., Pohl, M. G., Picoult-newberg, L., Ideker, T. E., Pohl, M. G., Boyce-jacino, M. (1999). Mining SNPs From EST Databases, *Genome Research* 167–174.

Plowman, J. E., Harland, D. P., Ganeshan, S., Woods, J. L., van Shaijik, B., Deb-Choudhury, S., Scobie, D. R. (2015). The proteomics of wool fibre morphogenesis. *Journal of structural biology*, 191(3), 341-351.

Plowman, J. E., Paton, L. N., & Bryson, W. G. (2007). The differential expression of proteins in the cortical cells of wool and hair fibres. *Experimental dermatology*, 16(9), 707-714.

Ponzoni R.W., R.J. Grimson, J.A. Hill, D.J. Hubbard, B.A. McGregor, A. Howse, I. Carmichael G.J. Judson. 1999. The inheritance of and association among some production traits in young Australian alpacas. En: <http://www.alpacas.com/AlpacaLibrary/InheritanceTraits.aspx>.

- Powell, B. C., Rogers, G. E. (1997). The role of keratin proteins and their genes in the growth, structure and properties of hair. *Exs*, 78, 59-148.
- Presciuttini, S., Valbonesi, A., Apaza, N., Antonini, M., Huanca, T., & Renieri, C. (2010). Fleece variation in alpaca (*Vicugna pacos*): a two-locus model for the Suri/Huacaya phenotype. *BMC genetics*, 11(1), 1- 8.
- Quispe, E. C. (2005). Mejoramiento genético y medioambiental de alpacas en la región de Huancavelica. *Proyecto de Inversión Pública a nivel de Perfil*.
- Quispe, E. C., Rodríguez, T. C., Iñiguez, L. R., & Mueller, J. P. (2009). Producción de fibra de alpaca, llama, vicuña y guanaco en Sudamérica. *Animal Genetic Resources Information*, 45, 1-14.
- Quispe, Q., Poma, A, Purroy, A. (2013). CARACTERISTICAS PRODUCTIVAS Y TEXTILES DE LA FIBRA DE ALPACAS DE RAZA HUACAYA/A REVIEW OF HUACAYA ALPACAS FIBER TRAITS. *Revista Complutense de Ciencias Veterinarias*, 7(1), 1.
- Rogers, G. (2004). Hair follicle differentiation and regulation. *International Journal of Developmental Biology*, 48(2-3), 163-170.
- Rogers, G. E. (1964). Structural and biochemical features of the hair follicle. In *The epidermis* (pp. 179-236). Academic Press New York.
- Rogers, G. E. (2006). Biology of the wool follicle: an excursion into a unique tissue interaction system waiting to be re-discovered. *Experimental dermatology*, 15(12), 931-949.
- Rogers, G. E., Powell, B. C. (1993). Organization and expression of hair follicle genes. *J. Invest. Dermatol.* 101, 50s-55s.
- Rogers, M. A, Langbein, L., Praetzel-Wunder, S., Winter, H., Schweizer, J. (2006). Human hair keratin-associated proteins (KAPs). *International review of cytology*, 251(06), 209-63.
- Rogers, M. A, Langbein, L., Winter, H., Beckmann, I., Praetzel, S., Schweizer, J. (2004). Hair keratin associated proteins: characterization of a second high sulfur KAP gene domain on human chromosome 21. *The Journal of investigative dermatology*, 122(1), 147-58.
- Rogers, M. A, Langbein, L., Winter, H., Ehmann, C., Praetzel, S., Korn, B., Schweizer, J. (2001). Characterization of a cluster of human high/ultrahigh sulfur keratin-associated protein genes embedded in the type I keratin gene domain on chromosome 17q12-21. *The Journal of biological chemistry*, 276(22), 19440-51.
- Rogers, M. A, Langbein, L., Winter, H., Ehmann, C., Praetzel, S., Schweizer, J. (2002). Characterization of a first domain of human high glycine-tyrosine and high sulfur keratin-associated protein (KAP) genes on chromosome 21q22.1. *The Journal of biological chemistry*, 277(50), 48993-9002.
- Rogers, M. A, Schweizer, J. (2005). Human KAP genes, only the half of it? Extensive size polymorphisms in hair keratin-associated protein genes. *The Journal of investigative dermatology*, 124(6), vii-ix.

- Rogers, M. A., Winter, H., Langbein, L., Wollschläger, A., Praetzel-Wunder, S., Jave-Suarez, L. F., Schweizer, J. (2007). Characterization of human KAP24.1, a cuticular hair keratin-associated protein with unusual amino-acid composition and repeat structure. *The Journal of investigative dermatology*, 127(5), 1197–204.
- Rufaut, N. W., Goldthorpe, N. T., Craven, A. J., Wallace, O. A., Wildermoth, J. E., Nixon, A. J. (2012). The germinative epithelium of sheep vibrissae and wool follicles has extensive proliferative potential but is dependent on the dermal papilla. *International journal of trichology*, 4(2), 75.
- Ruiz de Castilla, M., Alagon Huallpa, G., Quirita Bejar, C. R. (1992). Estudio de parámetros genéticos en alpacas Huacaya. *Informe de trabajos de investigación en alpacas y llamas de color*, 2, 1-28.
- Sasse, J., Mariasegaram, J., Jahabar, Ali M., Pullenayegum, S., Babu, R., Kinne, J., Wernery, U. (2000) Development of a microsatellite parentage and identity verification test for dromedary racing camels. *Abstracts of the 27th Conference of the international society of animal genetics (ISAG)*. Mineapolis. EE.UU.
- Schweizer, J., Bowden, P. E., Coulombe, P. A., Langbein, L., Lane, E. B., Magin, T. M. Wright, M. W. (2006). New consensus nomenclature for mammalian keratins. *The Journal of cell biology*, 174(2), 169–74.
- Sen, S., Burmeister, M. (2008). Hardy-Weinberg analysis of a large set of published association studies reveals genotyping error and a deficit of heterozygotes across multiple loci. *Human genomics*, 3(1), 1.
- Shete, S., Tiwari, H., Elston, R. C. (2000) On Estimating the Heterozygosity and Polymorphism Information Content Value. *Theoretical population biology*. Vol. 57, (3), 261-275.
- Shibuya, K., Kudoh, J., Obayashi, I., Shimizu, A., Sasaki, T., Minoshima, S., Shimizu, N., 2004a. Comparative genomics of the keratin-associated protein (KAP) gene clusters in human, chimpanzee and baboon. *Mamm. Genome*, 15, 179–192.
- Shimomura, Y., Aoki, N., Schweizer, J., Langbein, L., Rogers, M. A., Winter, H., Ito, M. (2002). Polymorphisms in the human high sulfur hair keratin-associated protein 1, KAP1, gene family. *Journal of Biological Chemistry*, 277(47), 45493-45501.
- Smith, E. J., Shi, L., Drummond, P., Rodriguez, L., Hamilton, R., Ramlal, S., Foster, J. (2001). Expressed sequence tags for the chicken genome from a normalized 10-day-old White Leghorn whole embryo cDNA library: 1. DNA sequence characterization and linkage analysis. *The Journal of heredity*, 92(1), 1–8.
- Sunnucks, P. (2000). Efficient genetic markers for population biology. *Trends in ecology & evolution*, 15(5), 199–203.
- Tabor, H. K., Risch, N. J., & Myers, R. M. (2002). Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Reviews Genetics*, 3(5), 391-397.

- Thomas, A., Harland, D. P., Clerens, S., Deb-Choudhury, S., Vernon, J. A., Krsinic, G. L., Dyer, J. M. (2012). Interspecies comparison of morphology, ultrastructure, and proteome of mammalian keratin fibers of similar diameter. *Journal of agricultural and food chemistry*, 60(10), 2434-2446.
- Valbonesi, A., Cristofanelli, S., Pierdominici, F., Gonzales, M., & Antonini, M. (2010). Comparison of fiber and cuticular attributes of alpaca and llama fleeces. *Textile Research Journal*, 80(4), 344-353.
- Van Tassell, C. P., Smith, T. P. L., Matukumalli, L. K., Taylor, J. F., Schnabel, R. D., Lawley, C. T., Sonstegard, T. S. (2008). SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries, *Nature methods* 5(3), 247–252.
- Vandebergh, W., & Bossuyt, F. (2012). Radiation and functional diversification of alpha keratins during early vertebrate evolution. *Molecular biology and evolution*, 29(3), 995-1004.
- Vásquez, R., Gómez-Quispe, O. E., & Quispe, E. (2015). Características Tecnológicas de la Fibra Blanca de Alpaca Huacaya en la Zona Altoandina de Apurímac: Technological Characteristics of the White Fibre of Huacaya Alpaca in The andean Region of Apurimac. *Revista de Investigaciones Veterinarias del Perú*, 26(2), 213-220.
- Vignal, A., Milan, D., SanCristobal, M., Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, 34(3), 275-306.
- Wang, B., Yang, W., McKittrick, J., Meyers, M. A. (2016). Keratin: Structure, mechanical properties, occurrence in biological organisms, and efforts at bioinspiration. *Progress in Materials Science*, 76, 229-318.
- Wang, X., Wang, L., Liu, X. (2012). *The quality and processing performance of alpaca fibres*. Rural Industries Research and Development Corporation (RIRDC).
- Wang, X., Zhao, Z. D., Xu, H. R., Qu, L., Zhao, H. B., Li, T., & Zhang, Z. Y. (2012). Variation and expression of KAP9. 2 gene affecting cashmere trait in goats. *Molecular biology reports*, 39(12), 10525-10529.
- Wheeler, J. C. (1995). Evolution and present situation of the South American Camelidae. *Biological Journal of the Linnean Society*, 54: 271-295.
- Wheeler, J. C. (2012). South American camelids - past, present and future. *Journal of Camelid Science*. 1–24.
- Wu, D. D., Irwin, D. M., Zhang, Y. P. (2008). Molecular evolution of the keratin associated protein gene family in mammals, role in the evolution of mammalian hair. *BMC Evolutionary Biology*, 8(1), 1.
- Wuliji T., Davis G.H., Dodds K.G., Turner P.R., Andrews R.N. y Bruce G.D. 2000. Production performance, repeatability and heritability estimates for live weight, fleece weight and fiber characteristics of alpacas in New Zealand. *Small Rumin. Res.*, 37: 189-201.

- Xavier, S. P., Gordon-Thomson, C., Wynn, P. C., McCullagh, P., Thomson, P. C., Tomkins, L., Moore, G. P. (2013). Evidence that Notch and Delta expressions have a role in dermal condensate aggregation during wool follicle initiation. *Experimental dermatology*, 22(10), 659-662.
- Yan, S., Yu, Z., Ning, L., Hai-dong, W., Jian-Shan, X., Shu-yuan, G., Xiao-yan, H. (2016). Let-7b promotes alpaca hair growth via transcriptional repression of TGF β R I. *Gene*, 577(1), 32-36.
- Yu, H., Wang, X., Chen, H., Wang, M., Zhao, M., Lan, X. Y., Wang, X. L. (2008). The polymorphism of a novel 30bp-deletion mutation at KAP9. 2 locus in the cashmere goat. *Small Ruminant Research*, 80(1), 111-115.
- Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., Kresovich, S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature genetics*, 38(2), 203-208.
- Yu, J., Yu, D., Checkla, D. M., Freedberg, I. M., Bertolino, A. P. (1993). Human hair keratins. *The society for investigative dermatology*. Vol. 101, No. 1.
- Yu, Z., Gordon, S. W., Nixon, A. J., Bawden, C. S., Rogers, M. A., Wildermoth, J. E., Pearson, A. J. (2009). Expression patterns of keratin intermediate filament and keratin associated protein genes in wool follicles. *Differentiation*, 77(3), 307-316.
- Zhang, Y., He, Y., Xue, P., & Chen, Y. (2011). Sequence variant in the KAP 1.1 gene associate with cashmere trait in two cashmere goat breeds. *African Journal of Agricultural Research*, 6(12), 2889-2894.
- Zhou, H., Gong, H., Li, S., Luo, Y., Hickford, J. G. H. (2015). A 57-bp deletion in the ovine KAP6-1 gene affects wool fibre diameter. *Journal of Animal Breeding and Genetics*, 132(4), 301-307.
- Zhu, A. J., Haase, I. and Watt, F. M. (1999). Signalling via b1 integrins and mitogen-activated protein kinase determines human epidermal stem cell fate in vitro. *Proc. Natl. Acad. Sci. USA* 96: 6728-6733
- Zhumbaeva, B. D., Gening, L. V., Gazaryan, K. G. (1992). Cloning and structural characterization of human hair sulfur-rich keratin genes. *Mol. Biol.* 26, 550–555.
- Zimek, A., Weber, K. (2005). Terrestrial vertebrates have two keratin gene clusters; striking differences in teleost fish. *European journal of cell biology*, 84(6), 623-635.

X. Anexos

Anexo 1. Secuencias a partir de las cuales se identificaron los genes en el genoma de referencia de alpaca.

>Secuencia gen KRTAP1-2 biblioteca de cDNA (543 pb)

```
ACCTCGCCCAGGCCTCCTGCTGCCGCCATCCTACTGTGGACAGTCCTACTGCCGCCAGCCTGC
TGCATCTGCTGTGAGCCCACCTGCTGTGAGCCCACCTGCTGAGAGCCCACCTGCTAAAAGTCAG
GTTTCTGATTTTCAACTTCAAATTTCAACTCCAGCTCAGTCCATGAACTAATAATCTCCTCAATC
ACCCCTGGGCCACTAACAAGTTCTTGAATTTTTATTGACTCTTTTTTGAAGGTTTACCAAATATT
TGGGCCTCACAGACCTGTCTGATTCCATTCAATGTAAAAAGATCTTGAGCGCTCCACTGAAGCT
GCCAAGATATAAGTTTGACCTAAGAAATGTGAAAACCAATTTGCCTCAGGATCAGCCTTCAGCA
AGCATCTTCTCCACGGCCCATTGCTTCCAAGGCTGTGGCAGAACCATCTGTCTTCTCAGAAACA
TTCATCTCTAACTCCACCATCTTGCAAAGTGCATTTCTATGATAGGGGAACAATATACCGAG
GTTTAATAAACTCTATCTTTGGT
```

>Secuencia gen KRTAP6-1 de oveja (Acc. Nr. NP_001180328.1)

```
ACACTCAAGTGACACTTCTACTCTCATTCTCTACCCGAGAACAACCTCAACAAGCAACACCATGT
GTGGCTACTACGAAACTACTATGGCGGCCCTCGGCTGTGGAAGCTACGGCTATGGAGGCCTGG
GCTGTGGCTGGCTCCTGCTACGGCTCTGGCTTCCGCAGGCTGGGCTGTGGCTATGGCTGTGGCT
ATGGCTATGGCTCCGCTCTCTGTGGAAGTGGCTATGGCTATGGCTCCCGCTCTCTGTGGAA
GTGGCTATGGATGCGGCTTGGCTATGGCTCTGGCTTTGGCTACTACTATTGAGGATGCCACGGA
AGACTCTCATCCTCTATACCTGGACACCAGGATCACCAGTTCTGAATGAACCCATACATTCTC
GTCCTGCTAAAGACTTGCCTTCCGGTGCCCTCTACATCTGATACACCAACCCTTCCCTTTTTGCAT
TTGATGTCAAAGAGTGAAGAATGTGAAATTCATTAAGAAGTCTAGTCATGATCTCTCATTAG
AATGATGCCATCACAGAACTCTTTTTGACTTCCATGCAGAAGACTTTTCTATTTTCTAATA
AACTTGTTCAATATGGAATCATAAA
```

> Secuencia gen KRTAP9-2 biblioteca de cDNA (142 pb)

```
ACCACCCACGTGCTGCTGCCTGCCTGGGTGCCTGGCCCAGGGCTGTGGATCCAGCTGCTGCCA
GCCTCGCTGCCGCCCTGTCTGCTGTCAGACCACGTGCTGCCGCCCTAGCTGTGTGTCCAGCTGCT
GCCAGCCCTCCTG
```

>Secuencia gen KRTAP11-1 biblioteca de cDNA (467 pb)

```
ACCAGCAGTCTTGGGTGTCCAGCTGCCGAAGAACTTGCTAAGTGTGCAGGAGCCAGTGAGCGA
ATCAAGCCTCTATGACCTGTCAGCTGTGTTTCCAGGATCTTCCAACATGCTGTCTGTCCCTGAAG
AATTCTTCATTGCTGACTGCTATTCTAACTGCCTGATTGCTGGCTGCCAGCCCTGAATAAGCTGC
CTTTGGCCATCTAAGCGTTTCTGGCCAGACCAATCTATTTAAGGGTTTGATCACCGGTGGC
ACGTATACCTCTGGAGTTTCCAGAAATTTACCACCCACGCGCAAGTCTAATGGTTTTGGCAT
GTCTTGACCTTTCTGCTTTGTCTGGCTTCTGGCTTCTGTTTTGTGCCTGTGAAAAGGGAACTTGT
```

CTCGCTCTGTGTTTCTCAATAAACCTTCATTACTTGGCATTGTAAAAAAAAAAAAAAAAAAAA
AAAAAAGCTTGT

>Secuencia gen KRTAP13-1 biblioteca de cDNA (720)

GTCACCATGTCCTACAACCTGCTGCTCTGGAACTTCTCCTCTCGCTCCTTTGGGGGCCACCTGCG
CTGCCGAGGCTCCTCCTGTGGCTCTTCTACCCAGCAACCTGGTCTACACCACGGACCTCTGCT
CTCCAGCACCTGCCAGCAGGACTCCTCTCTCCACAGCGGCTGTCAGGAGATCTGCTCTGAGCCT
ATCAGGTGCCAGACGTTCCAGGTGGAGTCCAGTCCCTGCCAGTCATCCTGCTACCGCTGGAAGA
CCTCCACGCTCTGCCGTCCCTGGCAGACGACTTACTCTGGGTCTCTGGGCTTTGGCTCCAGAGGT
TTTCAATCTTTGGTTGTGGCTTCCCGTCTCTGGGCTTTGGATCCAGTGGTTTCCAATCAGTGGGT
TGTGGTCCCAGGGCTTTCTCATCTGTAAGATGGAGATCCAGCTTTGACAGTCCAACTACTTCTC
TTCTAGGAGCTGGCAGCCTGCTTCTTTCCAACCAATCTTTAGATCTGGCTTCTTCTAATCTTATGC
TTGAAAACAACACCGACTGTATCTAGAAATTTGATGCAATATCTCCTATATCAAGATCTATACTAT
CTATTGGTCTCCTGTCCTACTGTTAGCTTACAAATTTGACTTCTAGTGAATCTATTATAGATTGG
GAATAATTTGAATACCAGTAACTGGAAATGCAATCAGTAAAGAAAGATGTCAAACTTCTTCCTT
GT

Anexo 2. Relación de las 96 muestras utilizadas para la validación de los marcadores

ID de la muestra	ID de la alpaca	Región de procedencia	ID de la muestra	ID de la alpaca	Región de procedencia
A1	253	Puno	A7	2956	Cusco
B1	B.LEE	Puno	B7	76	Cusco
C1	925	Puno	C7	2315	Cusco
D1	180	Puno	D7	704	Cusco
E1	084-1	Puno	E7	2367	Cusco
F1	458	Puno	F7	51	Cusco
G1	082-1	Puno	G7	2794	Cusco
H1	086-1	Puno	H7	60	Cusco
A2	087-1	Puno	A8	3682	Cusco
B2	088-1	Puno	B8	2510	Cusco
C2	089-1	Puno	C8	2408	Cusco
D2	144R	Puno	D8	621	Junin
E2	20R	Puno	E8	E1611	Junin
F2	19R	Puno	F8	19536	Junin
G2	149R	Puno	G8	AB--A-11	Junin
H2	150R	Puno	H8	025C	Junin
A3	151R	Puno	A9	19930	Junin
B3	155R	Puno	B9	AB--A-7	Junin
C3	161R	Puno	C9	7654	Huacavelica
D3	162R	Puno	D9	13met2	Huancavelica
E3	163R	Puno	E9	7643	Huancavelica
F3	190R	Puno	F9	7644-1	Huancavelica
G3	189R	Puno	G9	8693	Huancavelica
H3	196R	Puno	H9	8295	Huancavelica
A4	158R	Puno	A10	8643	Huancavelica
B4	1192	Puno	B10	2076	Huancavelica
C4	309	Puno	C10	pal-12-2	Huancavelica
D4	3075	Puno	D10	pal14	Huancavelica
E4	429	Puno	E10	pal15	Huancavelica
F4	245	Puno	F10	pal-20	Huancavelica
G4	781	Puno	G10	pal-28	Huancavelica
H4	1187	Puno	H10	pal11	Huancavelica
A5	122	Puno	A11	pal6	Huancavelica
B5	127	Puno	B11	pal122	Huancavelica
C5	6121	Puno	C11	pal 108	Huancavelica
D5	693	Puno	D11	pam-13	Apurimac
E5	6146	Puno	E11	pam-34	Apurimac
F5	672	Puno	F11	pam-24	Apurimac
G5	444	Puno	G11	pam-6	Apurimac
H5	49V	PUNO	H11	pam-36	Apurimac
A6	13V	Puno	A12	pam-38	Apurimac
B6	2V	Puno	B12	pam-68	Apurimac
C6	3629	Cusco	C12	pam-56	Apurimac
D6	2690	Cusco	D12	ar4	Arequipa
E6	3602	Cusco	E12	ar5	Arequipa
F6	2234	Cusco	F12	ar10	Arequipa
G6	58	Cusco	G12	ar3	Arequipa
H6	3645	Cusco	H12	ar18	Arequipa

Anexo 3. Alineamientos entre las proteínas predichas para KRTAP6-1 y KRTAP11-1 de alpaca y las proteínas ortólogas en humano.

Alineamiento entre la proteína predicha a partir del gen KRTAP6-1 de alpaca y la proteína KAP6-1 en humano (Acc. Nr. NP_853633.1) (66.3% de identidad)

```

V.pacosKAP6.1      1 MC-DYYGNYYGGRGYGCC-----GYGGLGCGYGSCYGCGRRLGCGYG      42
  || .|||||||..|||.|      |||||||||||||.|||||||
H.sapiensKAP6     1 MCGSYYGNYGTPGYGFCYGGGLGYGYGGLGCGYGSCCGCFRRL-----      45

V.pacosKAP6.1     43 CGYGCGYGFGRSLCGCGYGC GSGYGS GF GYYY      75
  |||||:|||||||.||||      |||.||||
H.sapiensKAP6    46 ---GCGYGYGRSLCGYGYGC----GSGSGYYY      71
  
```

Alineamiento entre la proteína predicha a partir del gen KRTAP11-1 de alpaca y la proteína KAP11-1 en humano (Acc. Nr. AAI30558.1) (82.8% de identidad)

```

V.pacosKAP11.     1 MSYNCSTRNCSSRLIGGQYSVPVTPWTASTQDADCLSGIHLPSFQTGS      50
  ||:|||||||||.|||:..|||.|.|.|||.||.|||||||
H.sapiensKAP1     1 MSFNCSTRNCSSRPIGGRCIVPAQVTTTSTTDADCLGGICLPSFQTGS      50

V.pacosKAP11.     51 WLLDHCQETCCEPTVCQPTCYQQTSCVSSPVQVTCRQTNCVSNPCSTAY    100
  |||||||||||||.|||||:|||||:|.|||||||.||:|||||.|
H.sapiensKAP1     51 WLLDHCQETCCEPTACQPTCYRRTSCVSNPCQVTCRQTTCISNPCSTTY    100

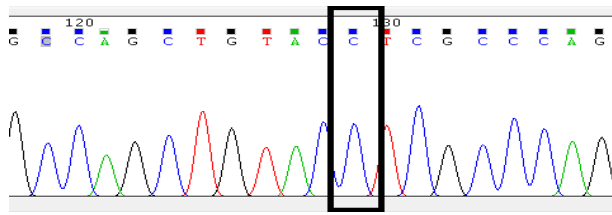
V.pacosKAP11.     101 SRPLSFVSSGCRPLGGISTVCQPVRVSVTVCRPVGGVSTICQPACGVSRT    150
  ||||:|||||:|||||:|||||.:.|||:|||||:|||||||
H.sapiensKAP1     101 SRPLTFVSSGQPLGGISSVCQPVGGI STVCQPVG VSTVCQPACGVSRT    150

V.pacosKAP11.     151 YQQSCVSSCRRTC      163
  |||||||||||
H.sapiensKAP1     151 YQQSCVSSCRRTC      163
  
```

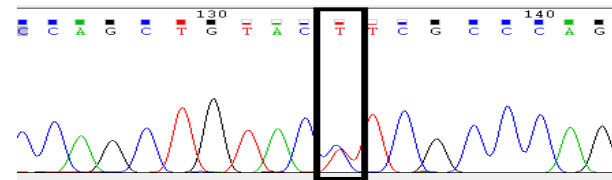
Anexo 4. Cromatogramas de los genotipos para cada una de las posiciones polimórficas identificadas

KRTAP1-2

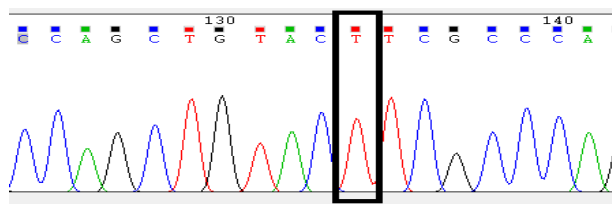
Posición 158



Individuo homocigoto CC

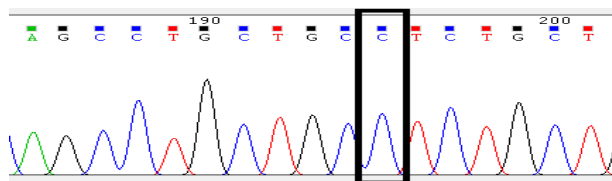


Individuo heterocigoto CT

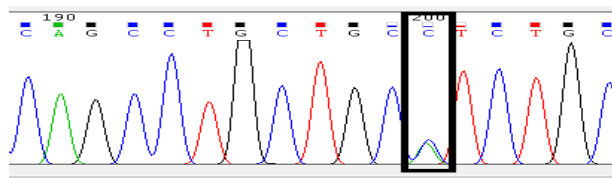


Individuo homocigoto TT

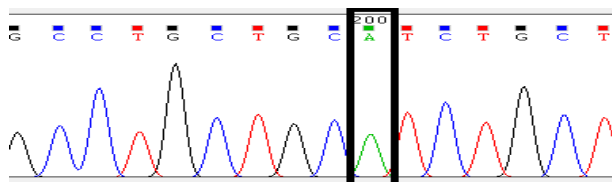
Posición 224



Individuo homocigoto CC

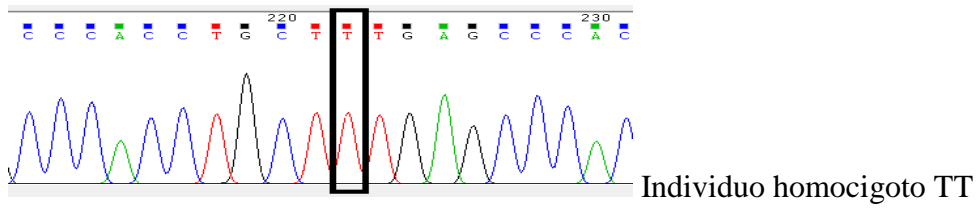
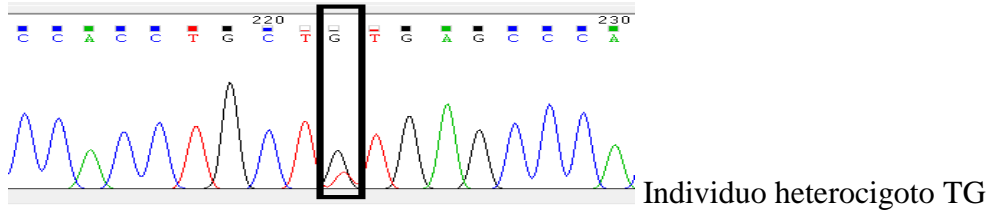
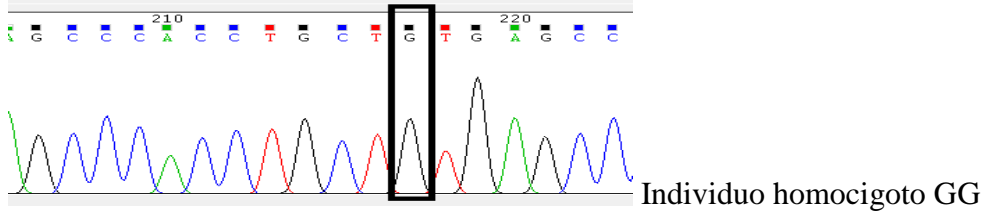


Individuo heterocigoto AC

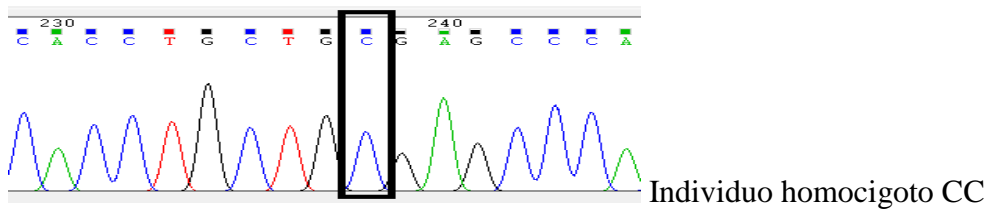
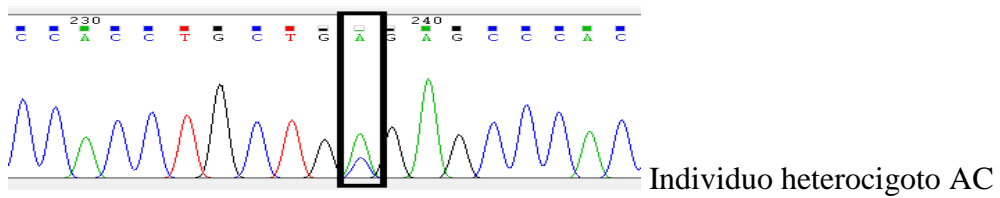
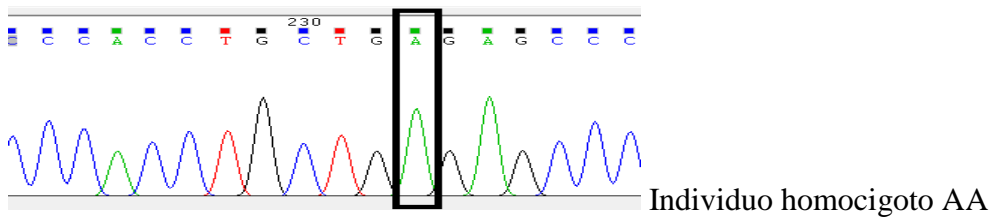


Individuo homocigoto AA

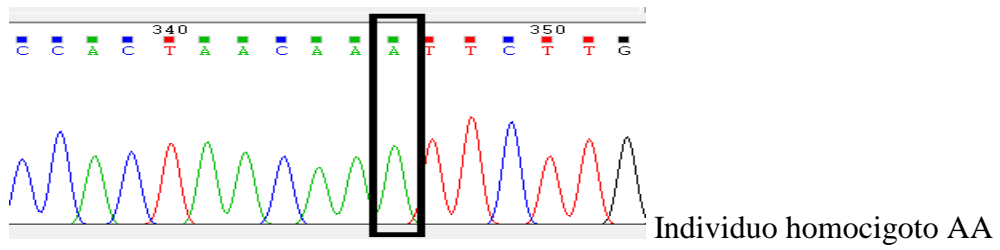
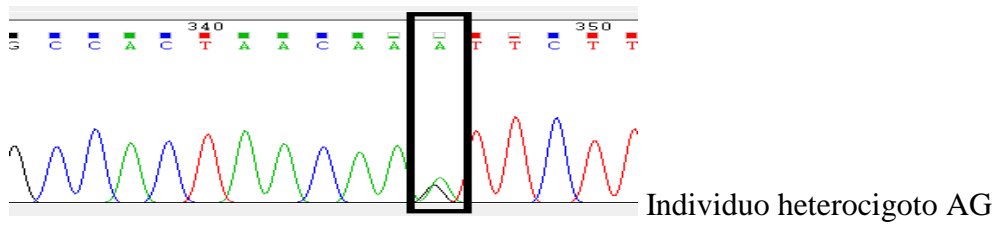
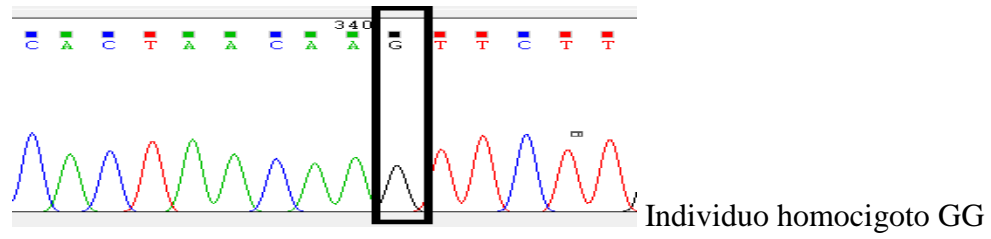
Posición 246



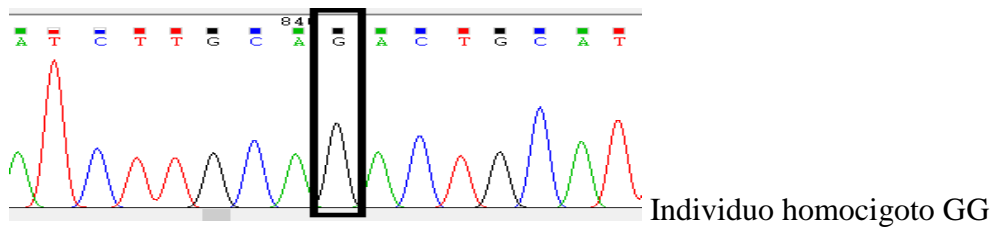
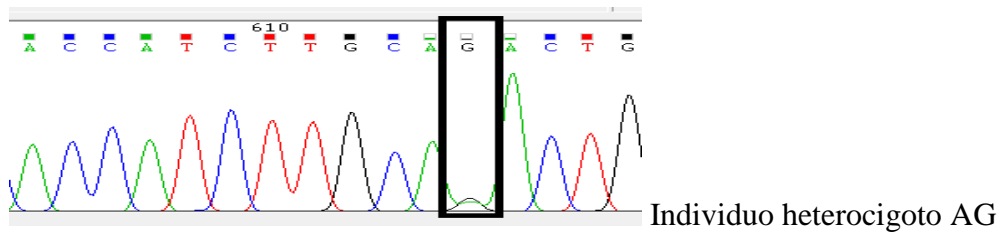
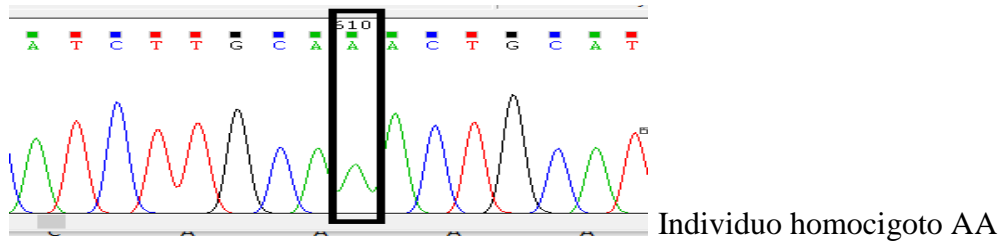
Posición 262



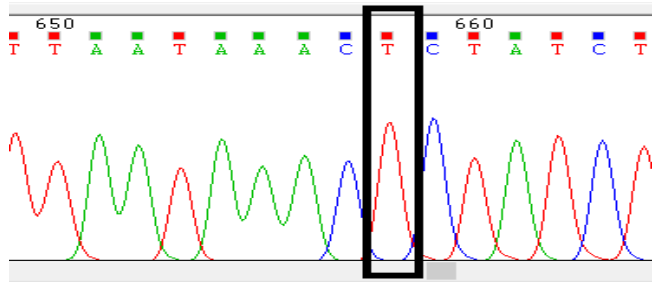
Posición 370



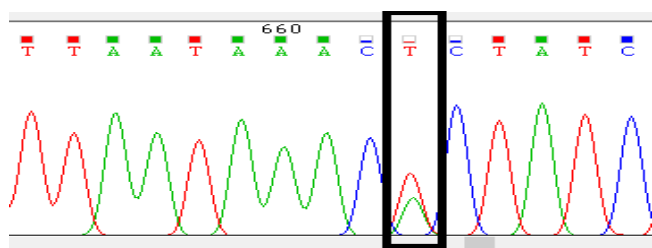
Posición 639



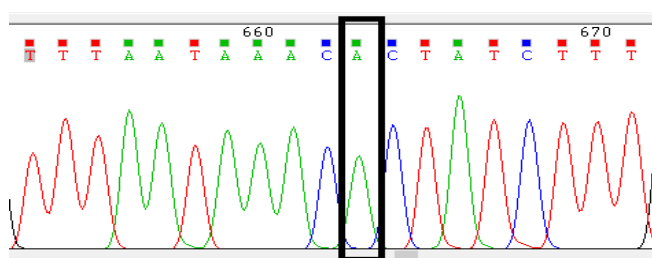
Posición 687



Individuo homocigoto TT

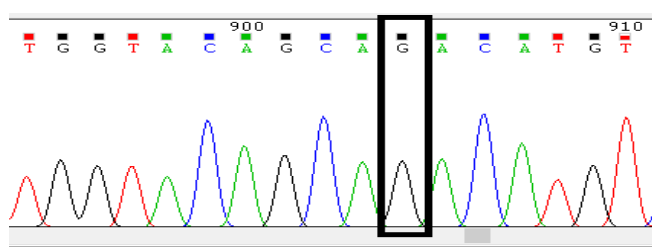


Individuo heterocigoto AT

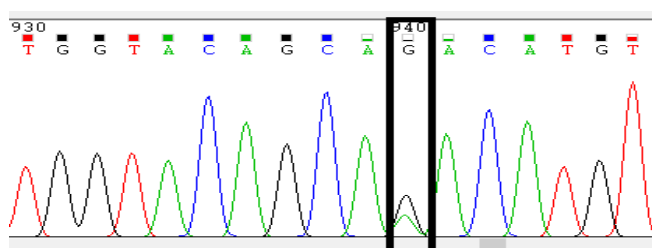


Individuo homocigoto AA

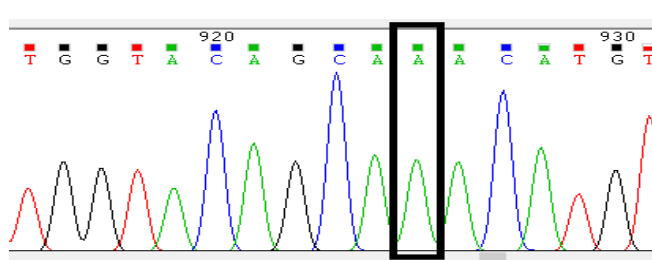
Posición 705



Individuo homocigoto GG

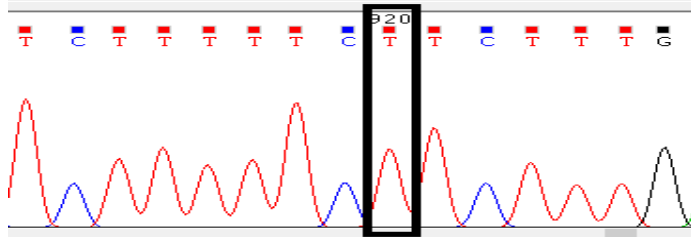


Individuo heterocigoto AG

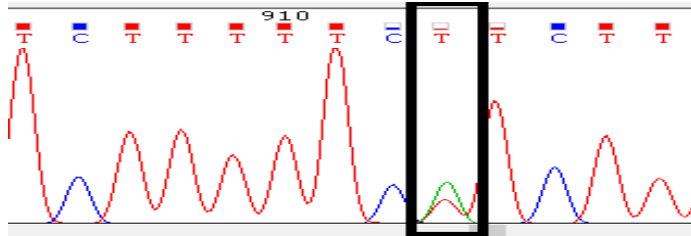


Individuo homocigoto AA

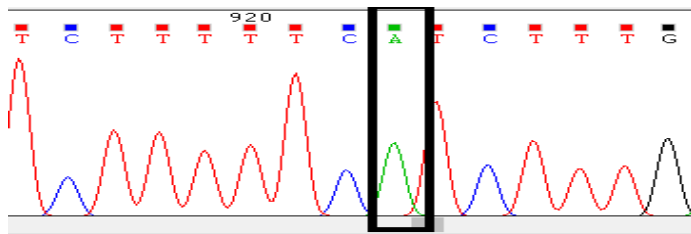
Posición 721



Individuo homocigoto TT

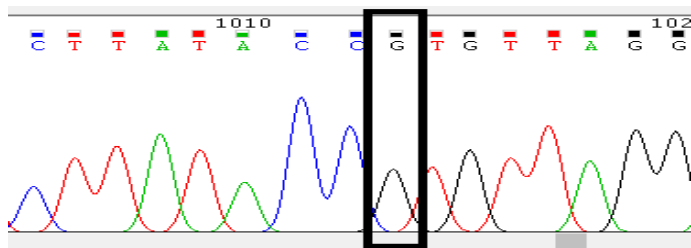


Individuo heterocigoto AT

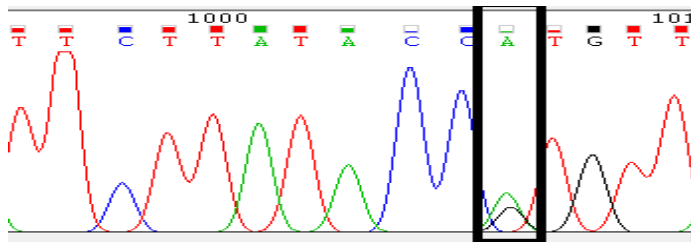


Individuo homocigoto AA

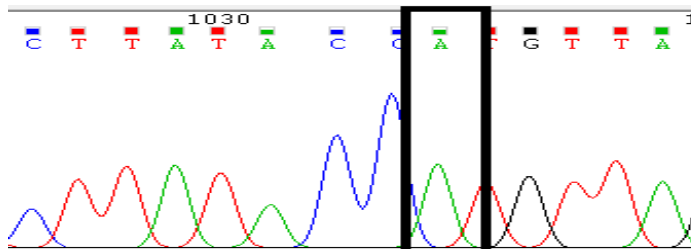
Posición 814



Individuo homocigoto GG



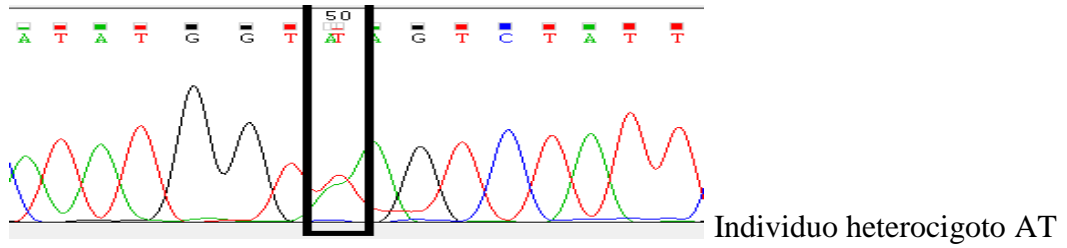
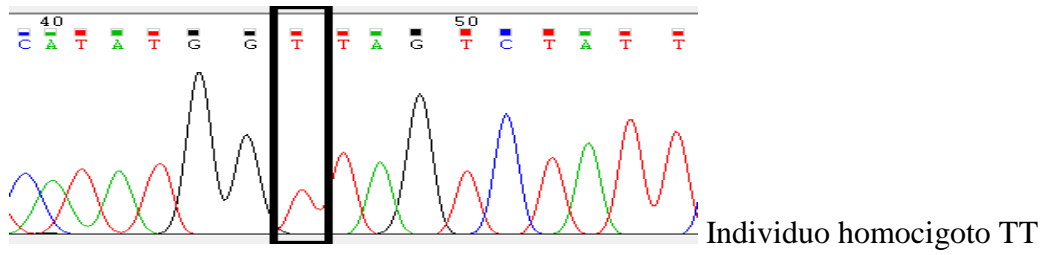
Individuo heterocigoto AG



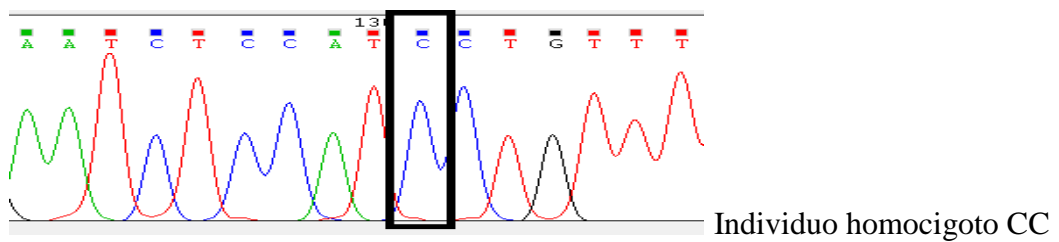
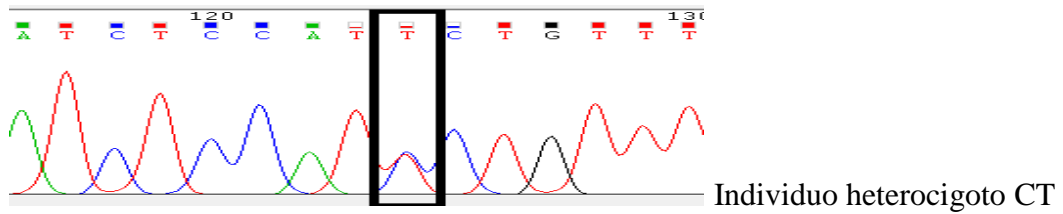
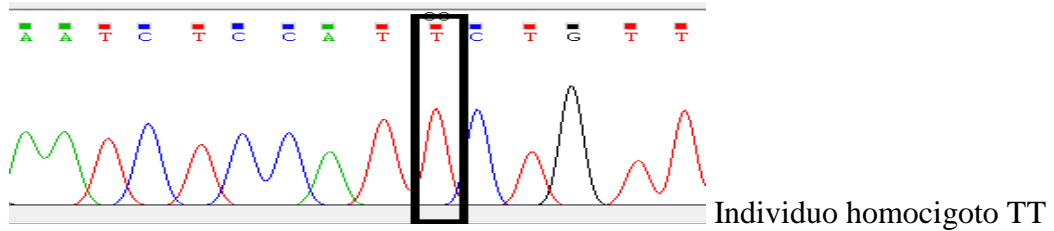
Individuo homocigoto AA

KRTAP6-1

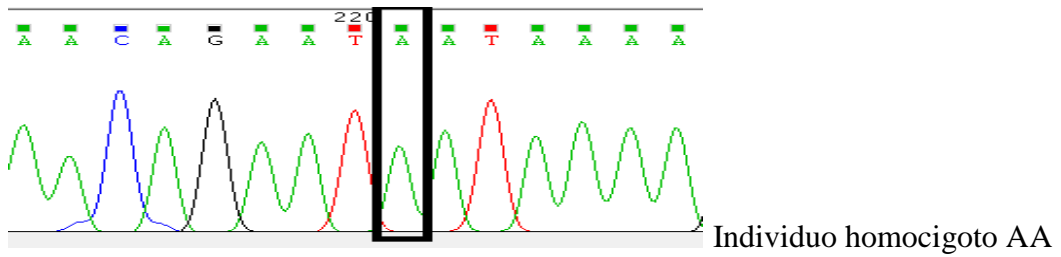
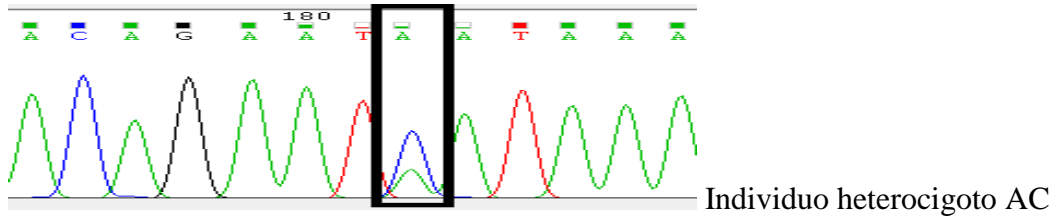
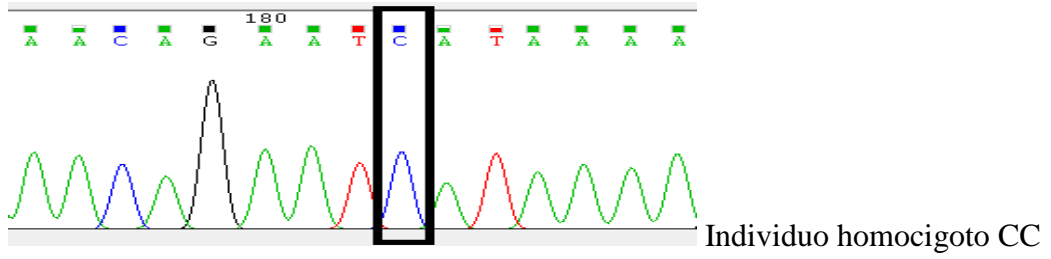
Posición 77



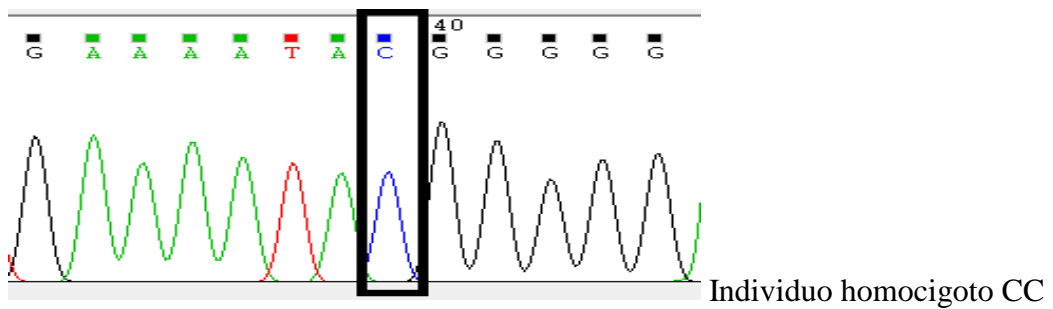
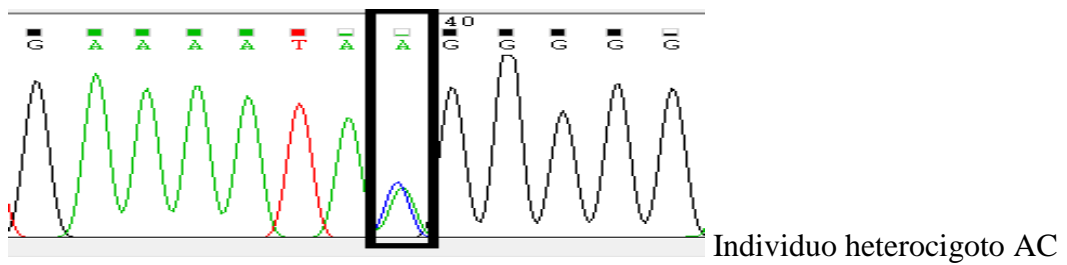
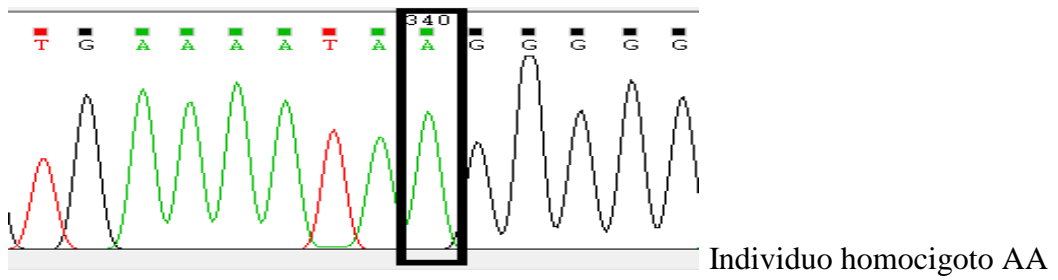
Posición 110



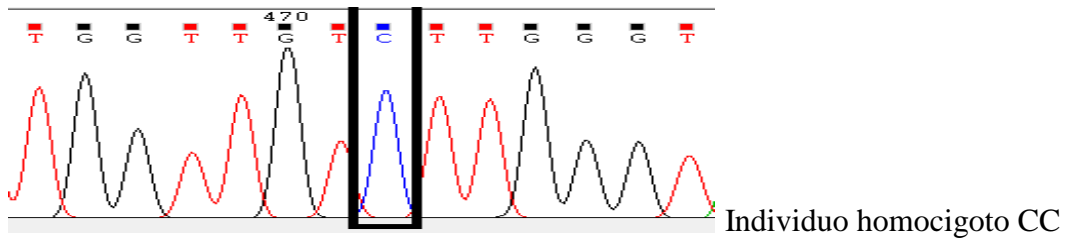
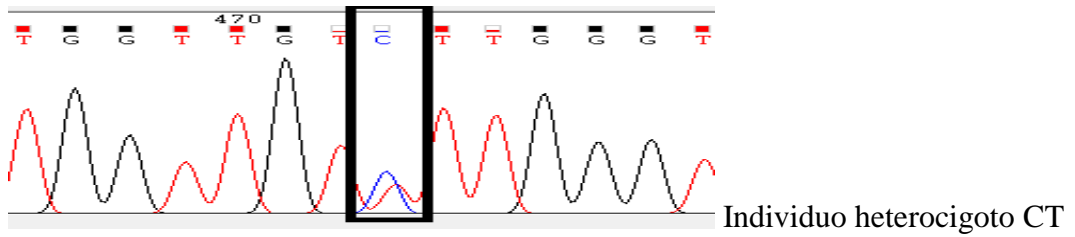
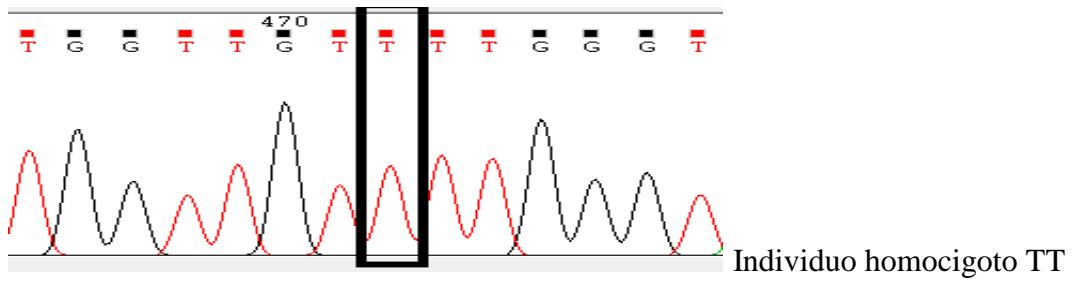
Posición 212



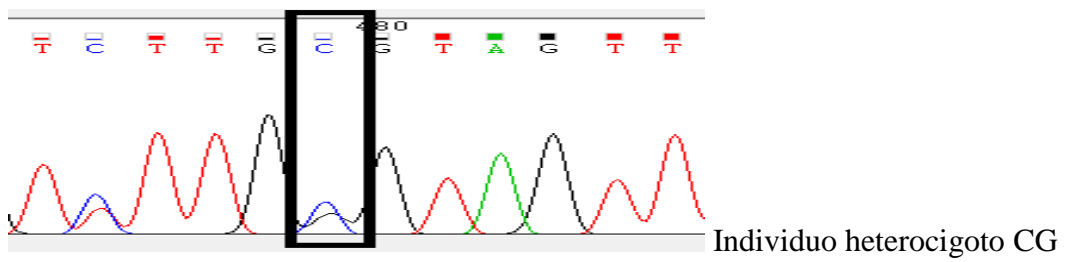
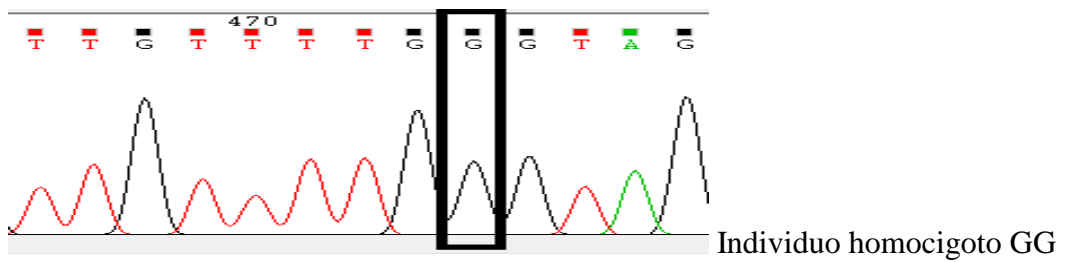
Posición 369



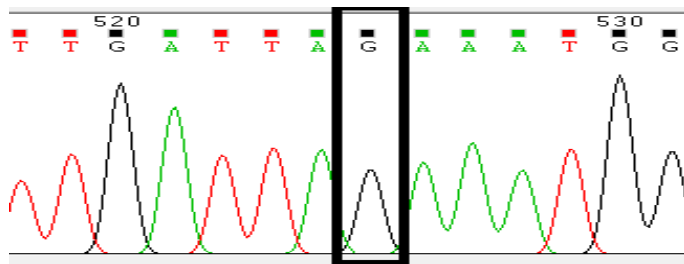
Posición 502



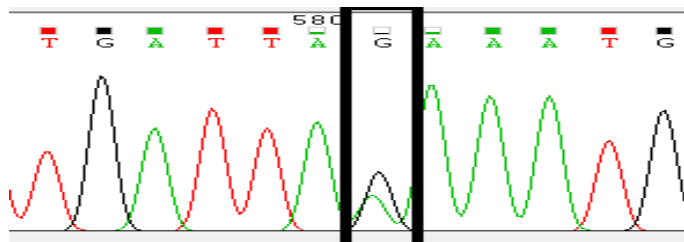
Posición 506



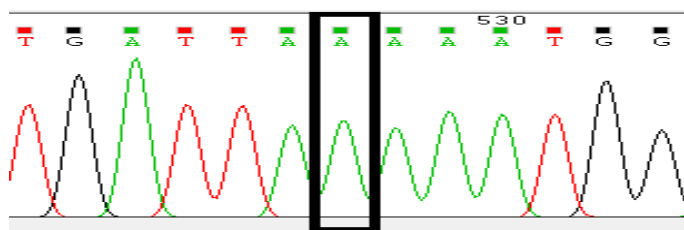
Posición 557



Individuo homocigoto GG

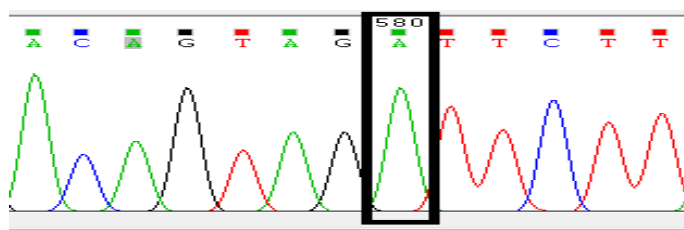


Individuo heterocigoto AG

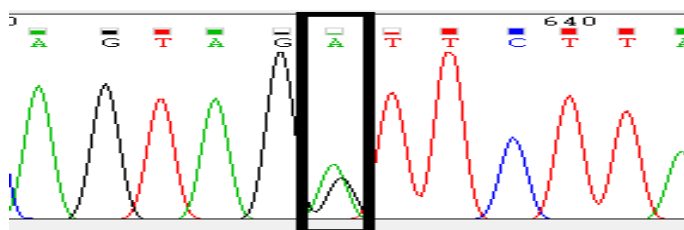


Individuo homocigoto AA

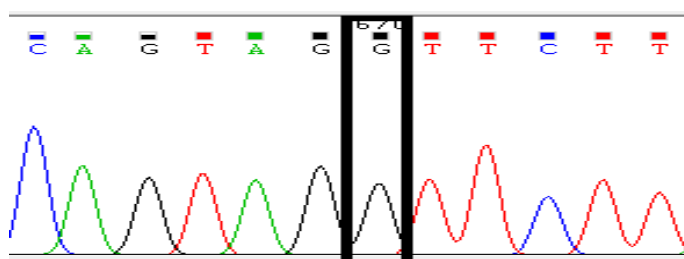
Posición 612



Individuo homocigoto AA

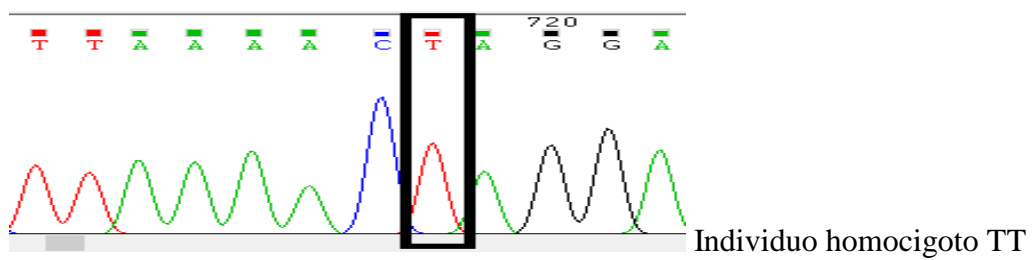
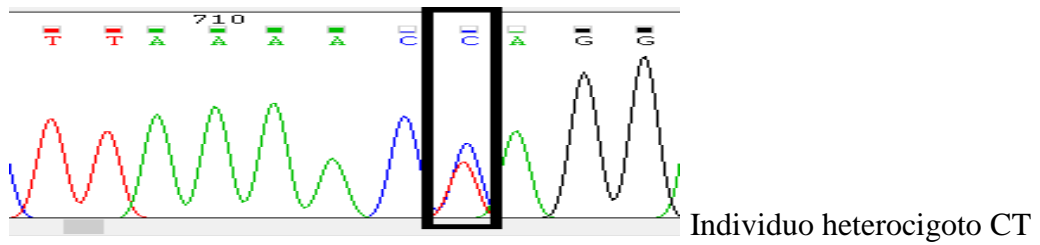
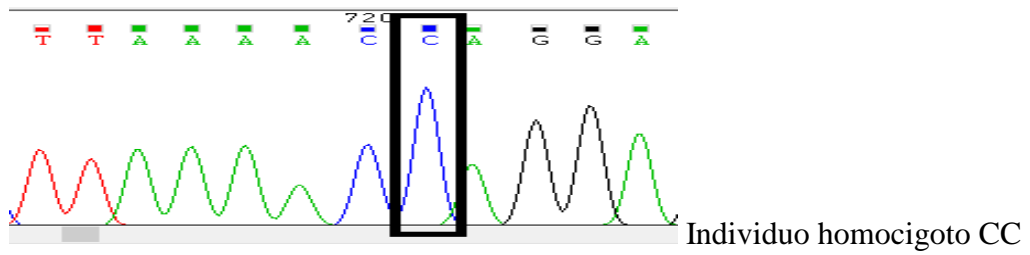


Individuo heterocigoto AG



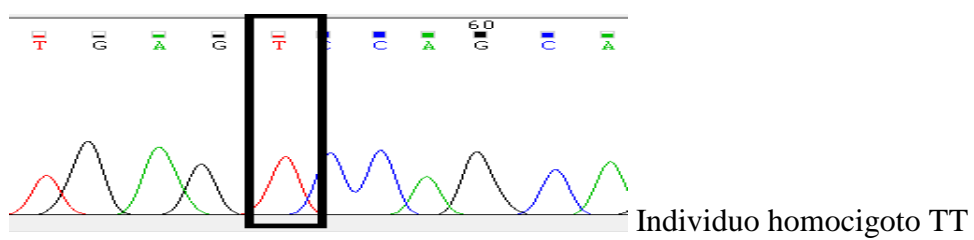
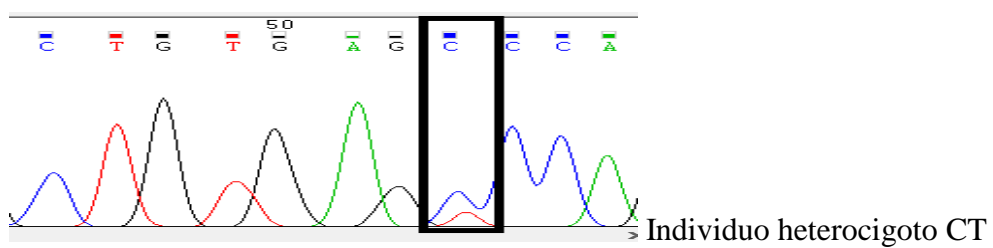
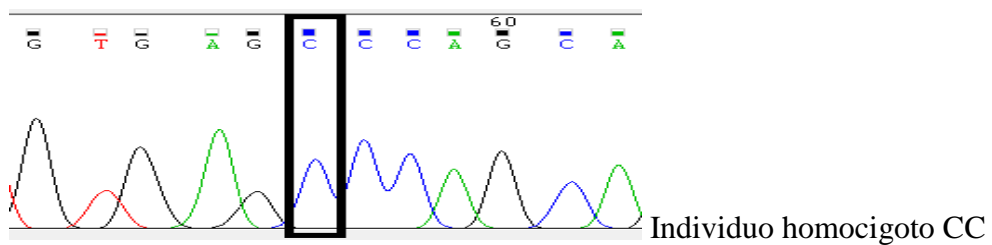
Individuo homocigoto GG

Posición 697

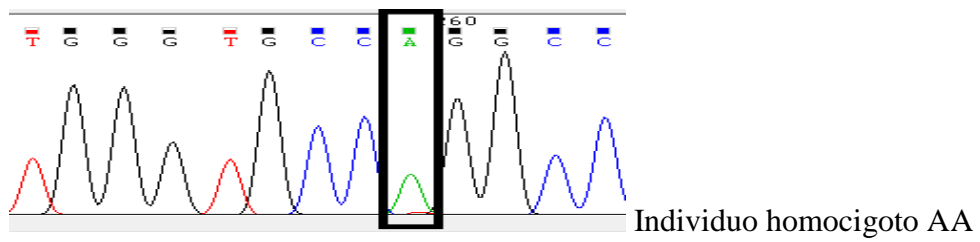
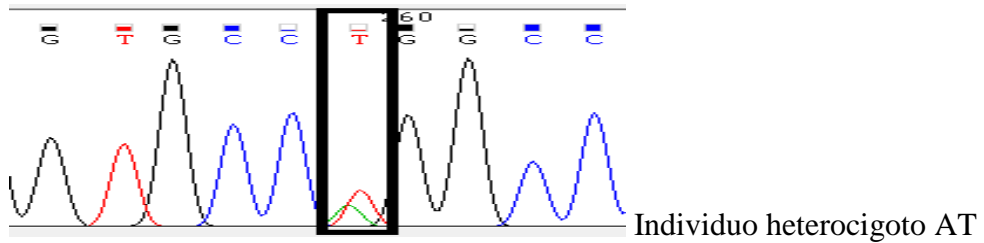
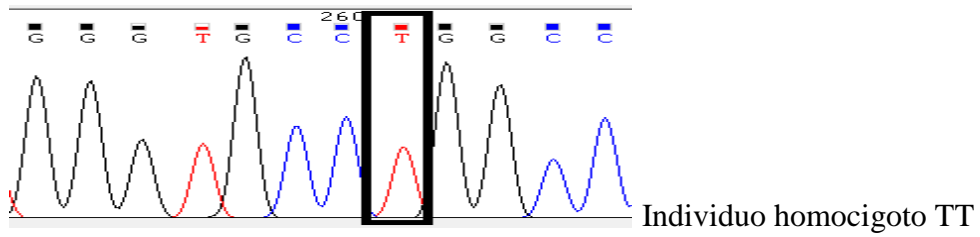


KRTAP9-2

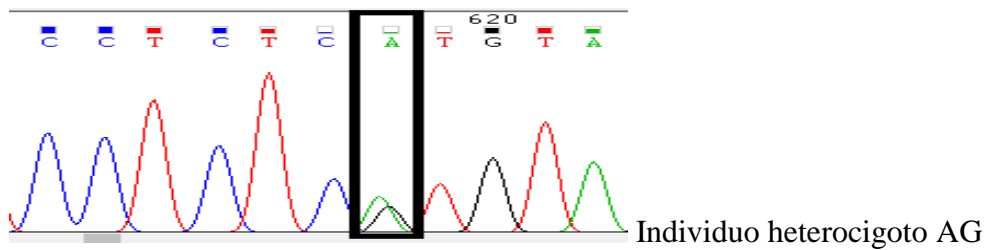
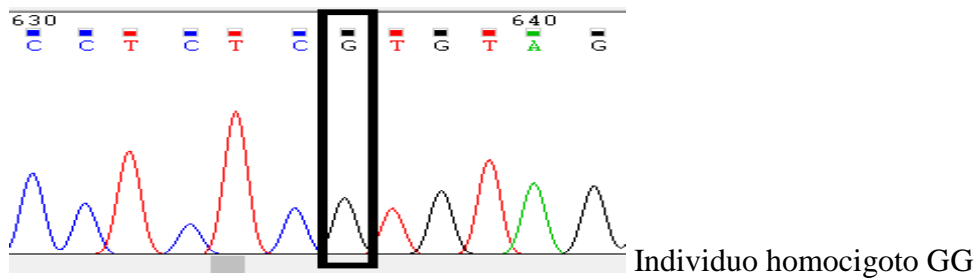
Posición 86



Posición 291

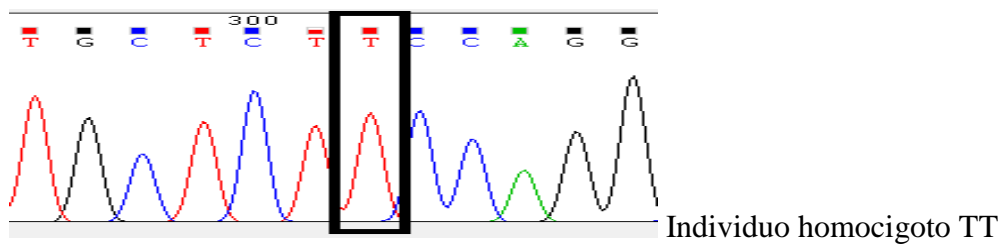
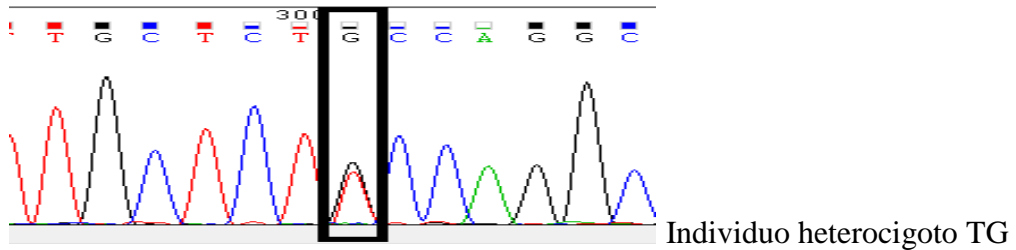
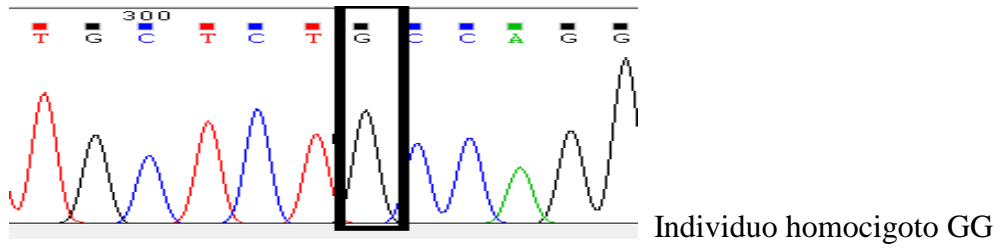


Posición 598

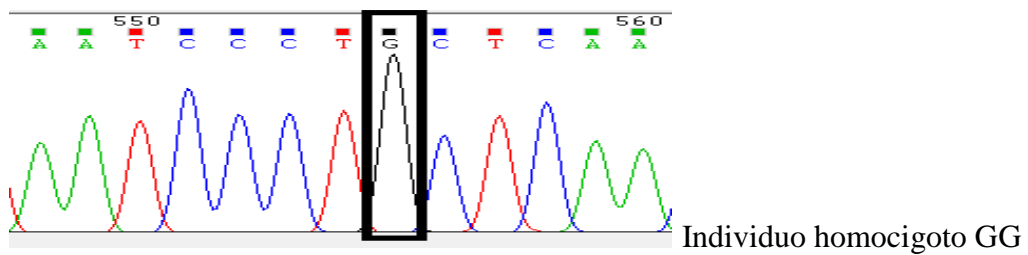
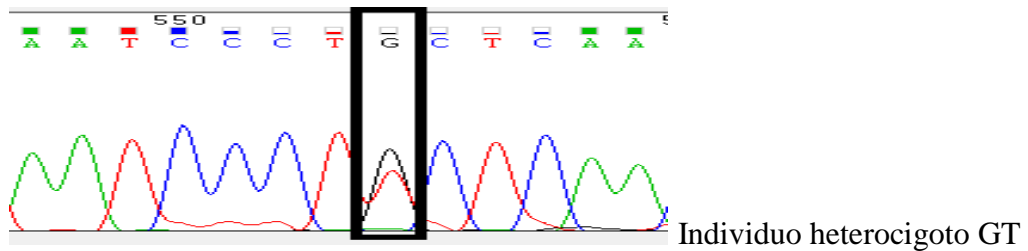
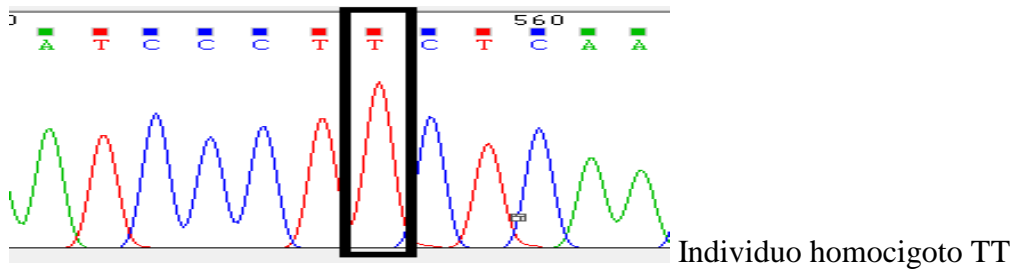


KRTAP11-1

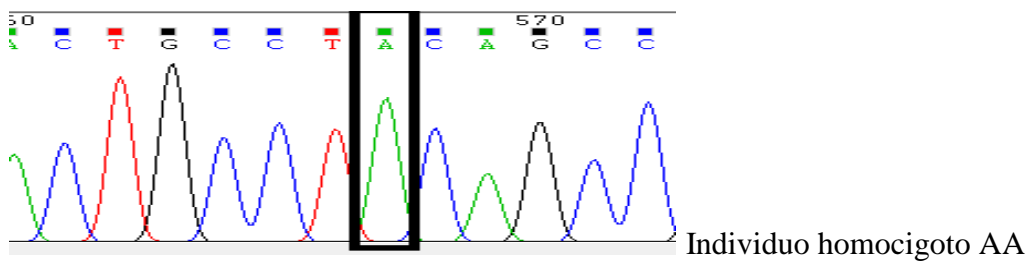
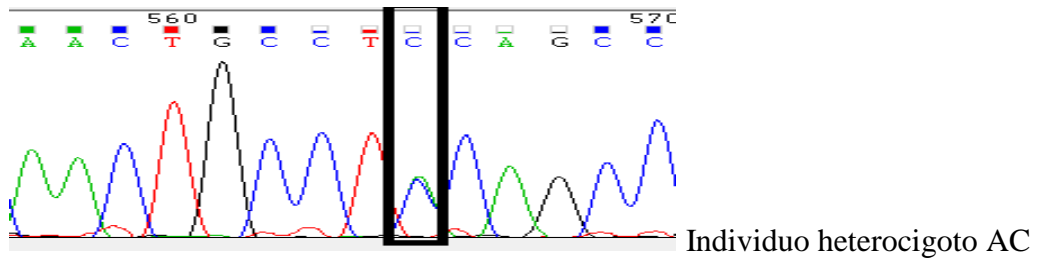
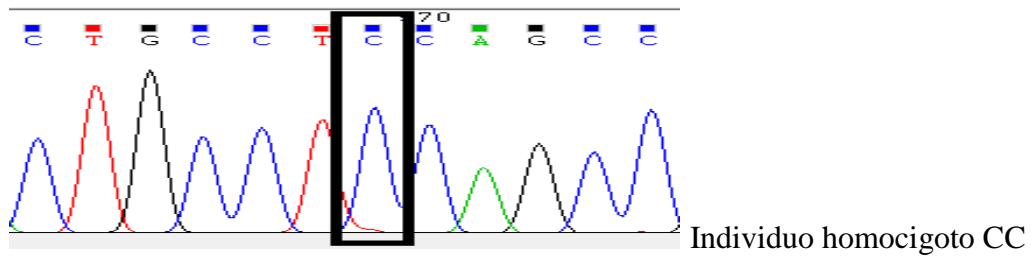
Posición 330



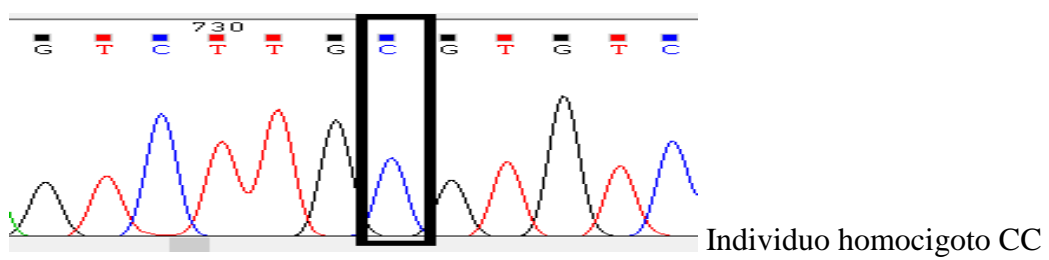
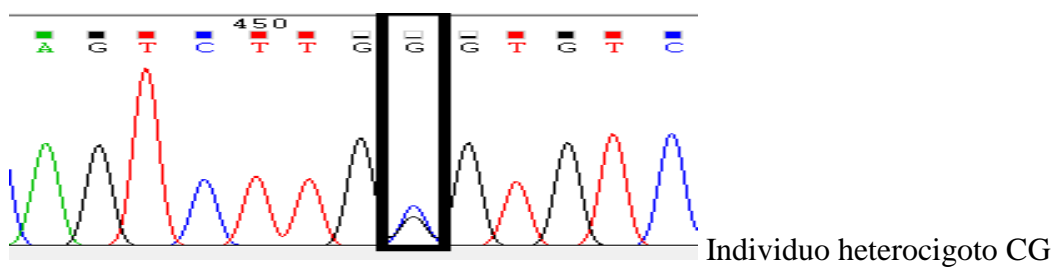
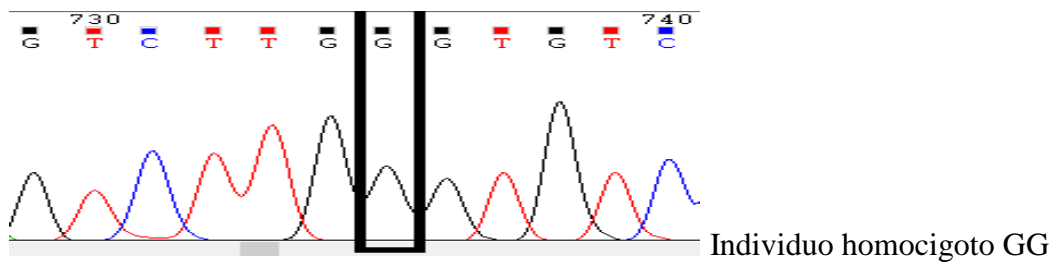
Posición 583



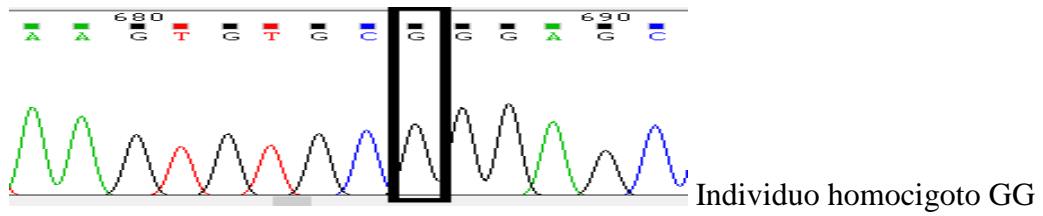
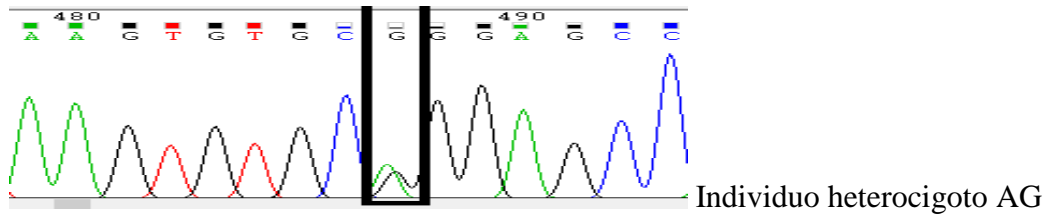
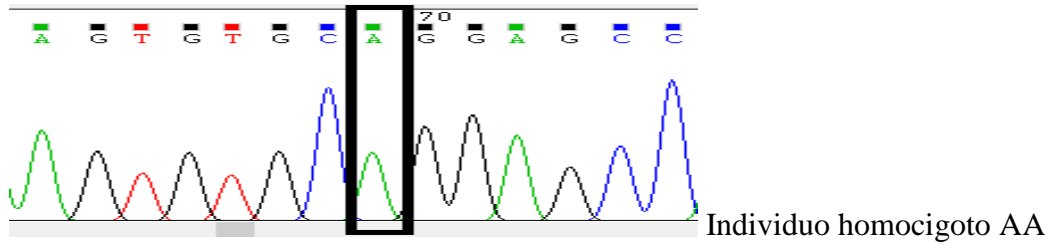
Posición 595



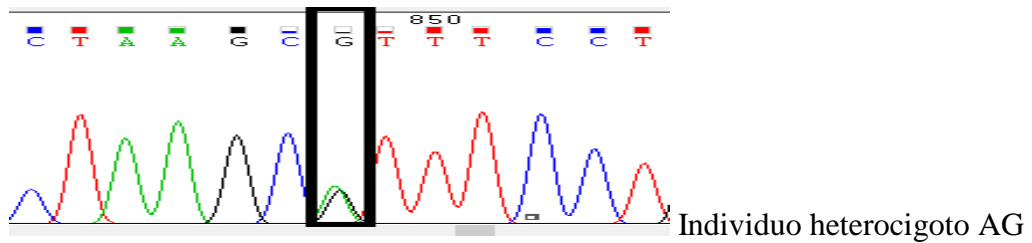
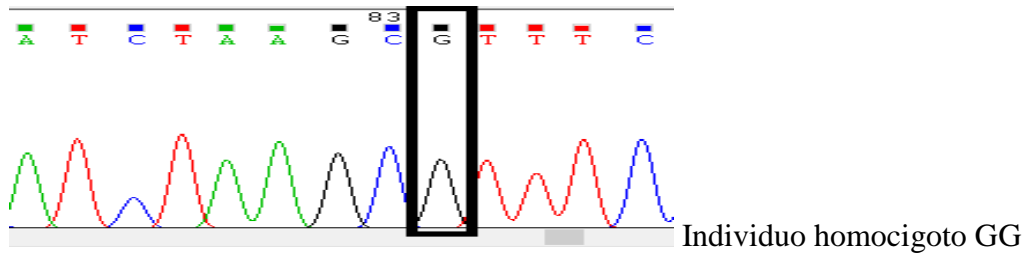
Posición 761



Posición 795

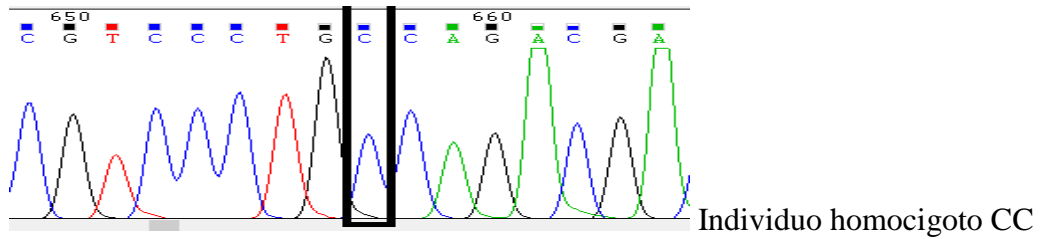
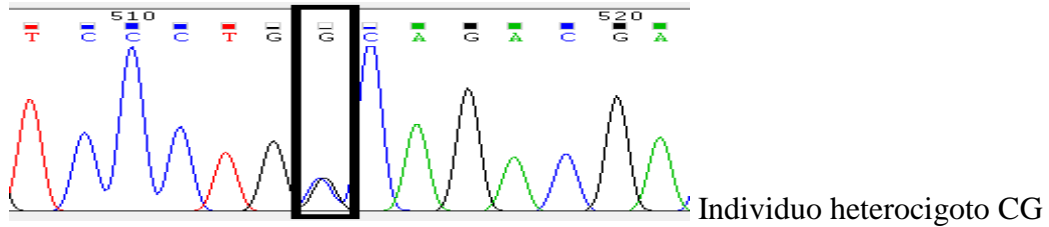
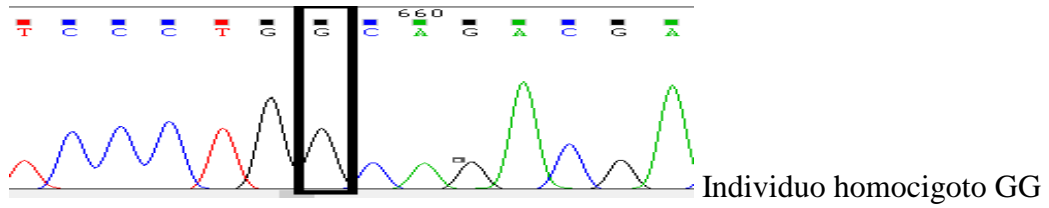


Posición 956

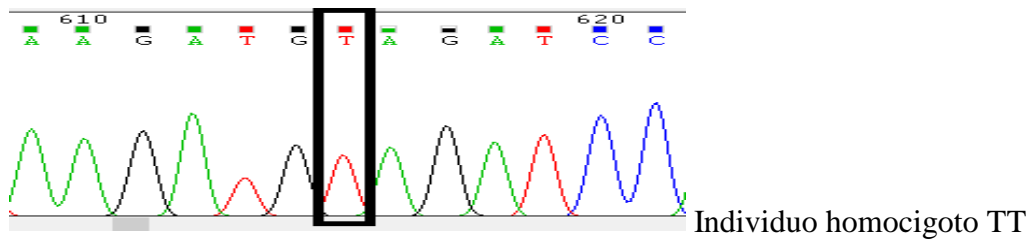
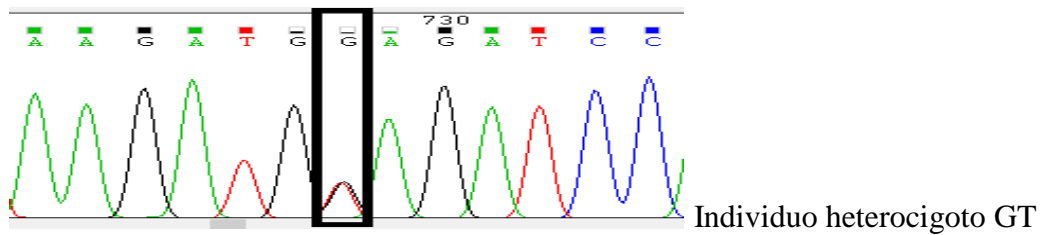
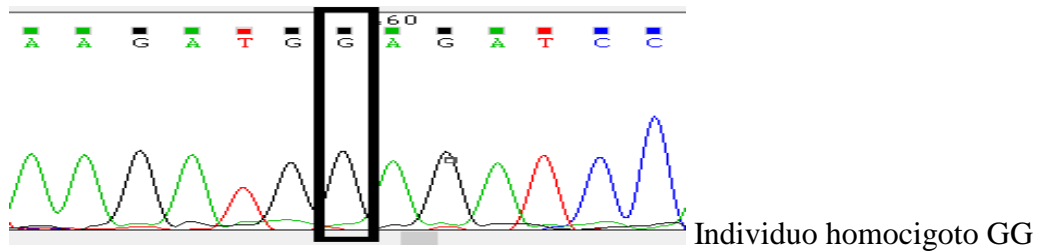


KRTAP13-1

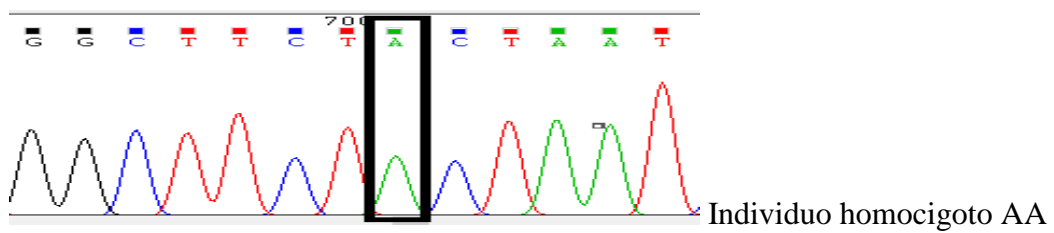
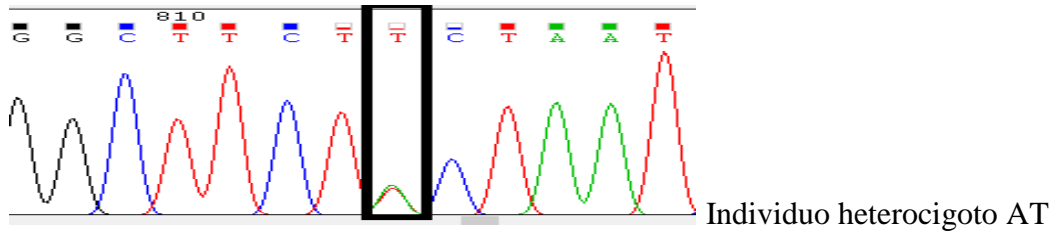
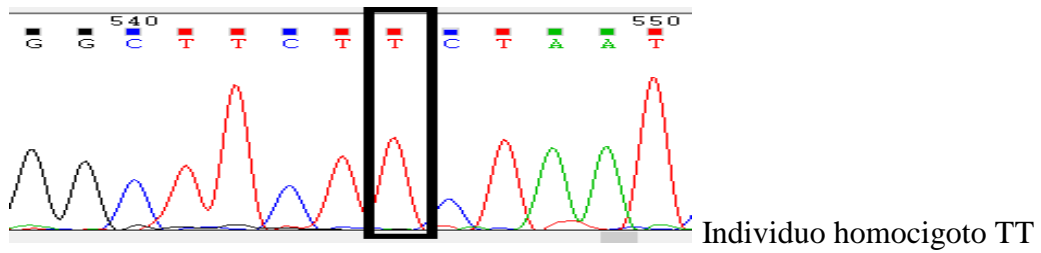
Posición 392



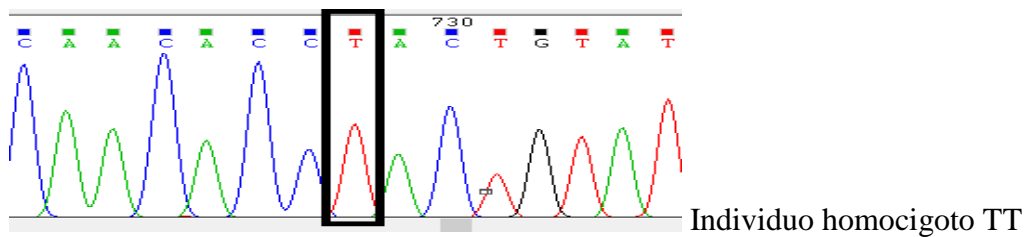
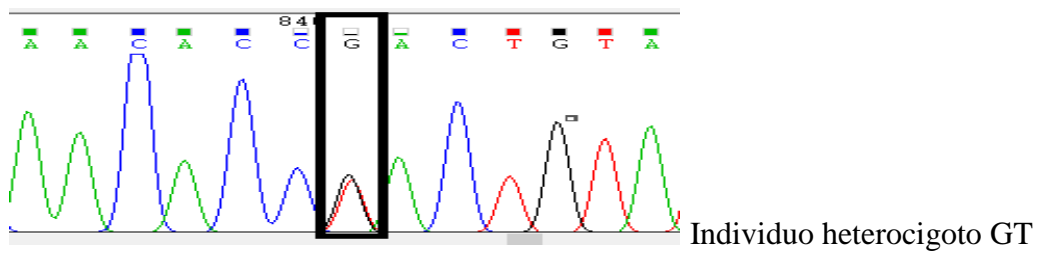
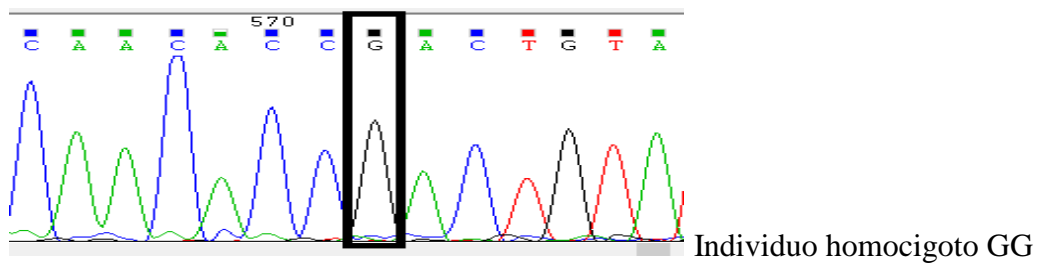
Posición 534



Posición 620



Posición 647



Anexo 5. Resultados del análisis de estructuración poblacional

Determinación de K

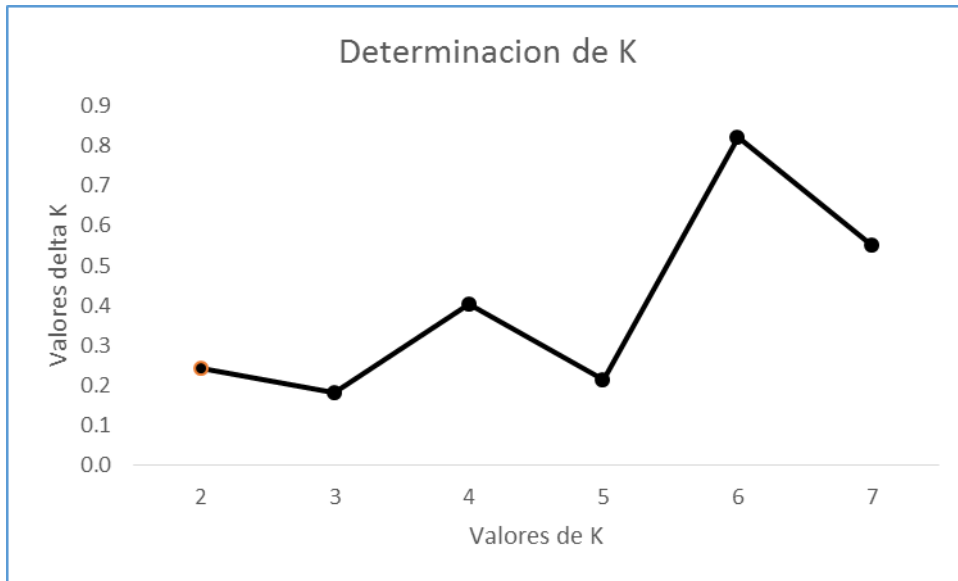
Datos obtenidos para cada una de las simulaciones en STRUCTURE

Nr. de simulacion	Ln P (D) K = 1	Ln P (D) K = 2	Ln P (D) K = 3	Ln P (D) K = 4	Ln P (D) K = 5	Ln P (D) K = 6	Ln P (D) K = 7
1	-6226.0	-6309.7	-6357.3	-6782.5	-6821.1	-6835.6	-6336.6
2	-6225.8	-6352.6	-6243.7	-6719.2	-7071.1	-6510.7	-7578.8
3	-6225.8	-6391.1	-6271.6	-6582.1	-6451.5	-6977.8	-7212.8
4	-6226.1	-6388.5	-6442.8	-6346.5	-6615.4	-6445.7	-6582.5
5	-6225.3	-6321.6	-6572.7	-6544.8	-6632.8	-6767.6	-7106.5
6	-6226.1	-6400.4	-6543.6	-6638.8	-6870.5	-6558.6	-6714.0
7	-6226.0	-6324.8	-6409.8	-6469.0	-7302.9	-7074.9	-7238.9
8	-6225.9	-6339.1	-6572.2	-6924.2	-6878.6	-7150.8	-7819.3
9	-6225.8	-6251.9	-6259.0	-6824.0	-6578.0	-6966.0	-7229.0
10	-6226.4	-6314.7	-6548.7	-6347.4	-6755.0	-6361.8	-7085.8
11	-6226.4	-6463.6	-6641.2	-6514.1	-6721.7	-6896.6	-6802.4
12	-6226.1	-6386.5	-6684.9	-6501.9	-6933.8	-7230.0	-7465.3
13	-6225.6	-6401.5	-6382.9	-6746.2	-6610.3	-6838.6	-6794.7
14	-6225.5	-6400.2	-6694.0	-6904.5	-6571.5	-6560.2	-6746.1
15	-6225.8	-6239.5	-6575.0	-6676.6	-7065.6	-6884.6	-6309.3
16	-6225.5	-6328.2	-6573.2	-6590.6	-6328.4	-6469.8	-6373.1
17	-6225.9	-6378.1	-6448.6	-6650.2	-6286.8	-6505.8	-6652.3
18	-6227.6	-6268.7	-6425.1	-6789.5	-6252.9	-6688.0	-7084.7
19	-6226.3	-6356.0	-6538.5	-6497.6	-7211.5	-6518.1	-7484.7
20	-6226.3	-6290.0	-6385.5	-6680.2	-6599.1	-6891.4	-7282.1
Promedio	-6226.0	-6345.3	-6478.5	-6636.5	-6727.9	-6756.6	-6994.9

Calculo del delta de K (Evanno et al., 2005)

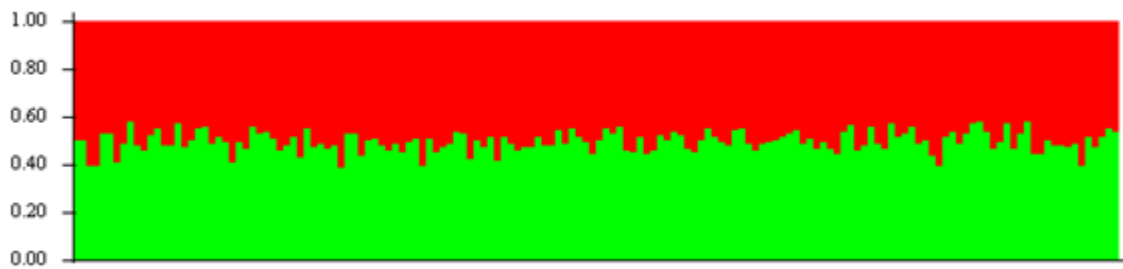
K	L(K)	StDev	L'(K)	L''(K)	[L''(K)]	Delta K
1	-6226.0	0.5	-	-	-	-
2	-6345.3	57.3	-119.3	-13.9	13.855	0.2
3	-6478.5	136.2	-133.2	-24.8	24.8	0.2
4	-6636.5	165.2	-158.0	66.6	66.6	0.4
5	-6727.9	294.4	-91.4	62.7	62.7	0.2
6	-6756.6	255.3	-28.7	-209.6	209.61	0.8
7	-6994.9	432.3	-238.3	238.3	238.3	0.6

Grafico Delta de K vs K

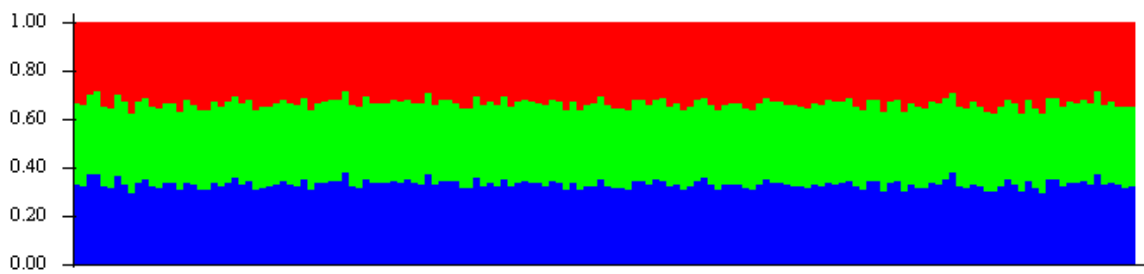


Diagramas de barras

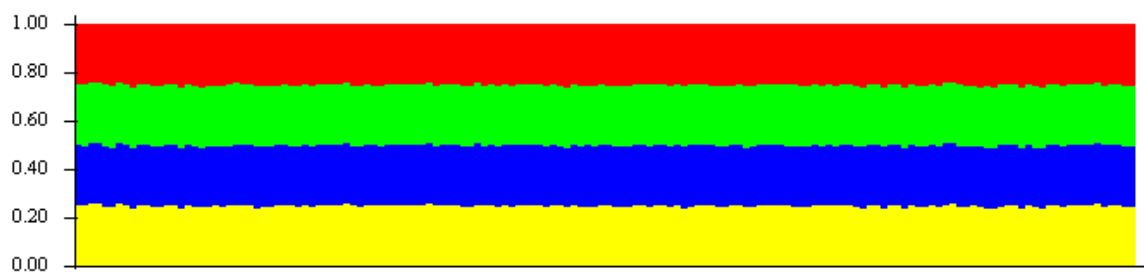
Para $K = 2$



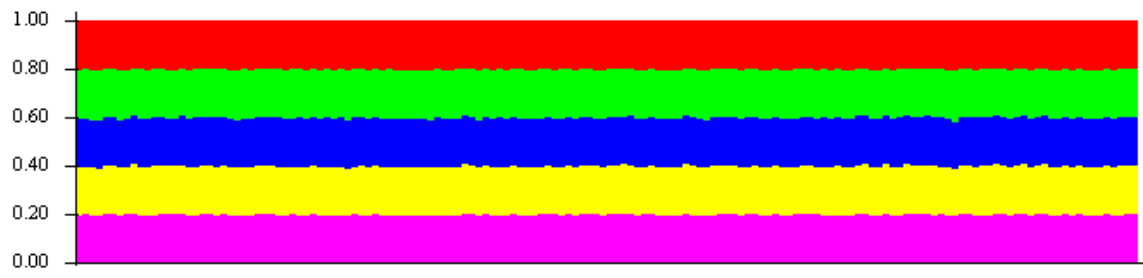
Para $K = 3$



Para $K = 4$



Para $K = 5$



Para $K = 6$



Para $K = 7$



Anexo 6. Datos de entrada para el estudio de asociación

ID Alpaca	ID Muestra	Edad	Sexo	Snpl	Snpl2	Diametro de Fibra	ID Alpaca	ID Muestra	Edad	Sexo	Snpl	Snpl2	Diametro de Fibra
1R	A1	1	1	1	1	16.5	81R	E10	3	1	2	1	20.1
2R	B1	1	1	0	0	15.6	82R	F10	2	1	1	1	19.0
24R	C1	2	1	1	0	17.9	83R	G10	1	1	1	0	19.7
27R	D1	4	1	1	0	25.4	84R	H10	1	1	2	0	19.9
28R	E1	2	1	0	0	17.4	87R	A11	2	1	2	0	20.6
29R	F1	4	1	1	0	25.2	88R	B11	1	1	2	0	20.0
30R	G1	3	1	1	0	19.0	90R	C11	1	1	0	0	23.3
31R	H1	2	1	1	1	26.4	91R	D11	1	1	0	0	23.1
33R	A2	2	1	1	0	19.7	92R	E11	1	1	1	0	18.3
35R	B2	2	1	2	0	15.8	94R	F11	1	1	1	0	21.2
36R	C2	4	1	2	0	23.4	95R	G11	2	1	1	0	24.1
37R	D2	4	1	0	0	24.5	97R	H11	1	1	0	0	22.4
38R	E2	1	1	1	1	21.5	98R	A12	2	1	1	0	28.1
39R	F2	3	1	2	0	24.0	99R	B12	2	1	2	1	21.7
40R	G2	3	1	0	0	13.6	100R	C12	1	1	0	0	23.0
41R	H2	4	1	1	1	20.6	123R	D12	1	1	0	0	17.6
43R	A3	4	1	1	0	23.7	125R	E12	2	1	1	0	23.7
46R	B3	2	1	0	0	20.8	206R	F12	4	1	0	0	24.9
48R	C3	3	1	1	0	22.3	209R	G12	2	1	1	0	19.5
50R	D3	1	1	1	1	21.3	210R	H12	4	1	1	0	22.0
51R	E3	3	1	1	0	19.1	126R	A1	2	1	1	1	25.4
52R	F3	2	1	1	0	21.9	127R	B1	2	1	2	1	24.4
54R	G3	2	1	1	0	16.3	128R	C1	1	1	0	0	19.7
55R	H3	2	1	2	0	19.3	131R	E1	2	1	1	0	23.4
58R	A4	4	1	2	0	23.0	132R	F1	1	1	1	1	19.4
25V	B4	1	1	1	0	20.6	133R	G1	1	1	1	0	23.4
12V	C4	3	1	1	0	17.3	134R	H1	2	1	1	0	23.1
13V	D4	2	1	2	1	18.6	135R	A2	1	1	1	0	20.8
14V	E4	4	1	2	0	21.6	137R	B2	1	1	1	0	21.6
15V	F4	4	1	1	0	24.2	138R	C2	4	1	1	0	29.9
16V	G4	4	1	0	0	18.1	139R	D2	4	1	2	0	19.7
18V	H4	4	1	2	1	22.2	140R	E2	2	1	1	0	18.4
19V	A5	3	1	1	0	24.9	141R	F2	2	1	1	0	23.4
20V	B5	4	1	1	1	23.6	142R	G2	1	1	1	1	25.5
5V	C5	1	2	1	0	20.4	143R	H2	2	1	0	0	18.7
22V	D5	2	2	1	0	21.2	21R	A3	1	1	1	0	14.0
11V	E5	2	2	1	0	21.1	147R	C3	3	1	1	0	18.9
24V	F5	3	2	1	0	20.7	152R	D3	1	1	0	0	17.3
47V	G5	2	2	0	0	23.4	153R	E3	2	1	1	0	17.8
49V	H5	4	2	0	0	23.2	154R	F3	2	1	1	0	18.3
52C	A6	1	1	1	1	15.9	156R	G3	2	1	2	0	18.9
53C	B6	1	1	1	0	15.4	157R	H3	2	1	0	0	18.8
54C	C6	1	1	2	0	15.4	160R	A4	4	1	1	0	21.1
55C	D6	1	1	1	0	19.1	164R	B4	2	1	0	0	28.1
18C	E6	2	1	1	0	19.1	167R	C4	2	1	1	1	20.4
19C	F6	4	1	0	0	22.2	168R	D4	2	1	1	0	22.4
59C	G6	1	1	2	0	22.6	170R	E4	3	1	1	0	18.3
20C	H6	2	1	2	0	21.7	171R	F4	3	1	1	0	22.7
21C	A7	4	1	1	0	25.9	172R	G4	4	1	2	0	20.1
22C	B7	2	1	2	0	21.4	173R	H4	4	1	1	0	20.0
70C	C7	1	2	1	0	18.6	174R	A5	4	1	2	1	19.1
75C	D7	1	2	2	0	27.8	175R	B5	3	1	1	0	22.0
80C	E7	1	2	0	0	20.8	176R	C5	4	1	2	0	19.9
81C	F7	2	2	1	0	18.8	177R	D5	4	1	2	1	20.5
68C	G7	2	2	1	1	18.8	178R	E5	2	1	1	0	19.5
30C	H7	3	2	1	0	26.3	179R	F5	3	1	0	0	19.3
31C	A8	2	2	0	0	16.4	181R	G5	4	1	1	0	23.3
32C	B8	3	2	0	0	24.6	183R	H5	3	1	1	0	19.6
33C	C8	2	2	1	0	23.2	184R	A6	2	1	1	0	20.5
34C	D8	2	1	1	0	19.7	185R	B6	4	1	1	1	20.0
060R	E8	1	1	2	0	20.3	188R	C6	4	1	2	0	19.9
061R	F8	1	1	2	1	19.6	191R	D6	4	1	1	1	17.9
062R	G8	1	1	1	0	22.9	193R	E6	4	1	1	1	19.7
064R	H8	1	1	2	1	20.6	195R	F6	2	1	1	1	19.5
65R	A9	3	1	2	2	28.8	197R	G6	4	1	2	0	20.7
70R	B9	1	1	0	0	18.4	198R	H6	2	1	1	0	22.6
71R	C9	1	1	1	1	17.6	200R	A7	4	1	2	0	19.9
72R	D9	1	1	1	1	20.7	207R	B7	1	1	1	0	18.3
73R	E9	2	1	1	0	28.1	101R	C7	4	1	0	0	23.0
74R	F9	1	1	1	0	21.3	104R	D7	2	1	0	0	20.6
75R	G9	1	1	1	0	27.8	110R	F7	4	1	0	0	24.2
76R	H9	1	1	1	0	20.2	201R	G7	2	1	1	0	18.7
77R	A10	1	1	1	0	20.4	202R	H7	3	1	0	0	18.2
78R	B10	1	1	0	0	17.6	203R	A8	4	1	1	0	20.8
79R	C10	1	1	1	1	19.7	204R	B8	4	1	1	0	15.4
80R	D10	1	1	1	0	18.8	205R	C8	2	1	1	0	24.3

Explicacion de la tabla anexo 6: En la tabla del Anexo 5 la edad esta expresada como una variable categórica con escala ordinal que refleja la dentadura del animal, donde: el valor 1 corresponde a dientes de leche, el valor 2 corresponde a 2 dientes, el valor 3 corresponde a 4 dientes y el valor 4 corresponde a boca llena. En el sexo, el valor 1 corresponde a hembras y el valor 2 corresponde a machos. La variable SNP1 corresponde al haplotipo de la combinación de los SNPs 330, 583, 595, 761 y 795. Donde se identifican 2 alelos (TGACG y GTCGA). El valor 0 corresponde a la ausencia del alelo TGACG, el valor 1 corresponde a la presencia de un solo alelo TGACG y el valor 2 corresponde a la presencia de dos alelos TGACG. El SNP2 corresponde al genotipo del SNP 956, donde el valor 0 corresponde a la ausencia del alelo A, el valor 1 corresponde a la presencia de un alelos A y el valor 2 corresponde a la presencia de 2 alelos A. Por último se muestra el diámetro de fibra medido en micrómetros.