

UNIVERSIDAD DE COSTA RICA  
SISTEMA DE ESTUDIOS DE POSGRADO

**ASOCIACIÓN DE UN POLIMORFISMO EN LA POSICIÓN -1689 DE LA  
REGIÓN PROMOTORA DEL GEN DE  $\alpha$ -LACTALBÚMINA BOVINA CON LA  
COMPOSICIÓN PROTEÍNIC DE LA LECHE**

Tesis sometida a la consideración de la Comisión del Programa de Estudios de Posgrado en  
Biología para optar al grado y título de Maestría Académica en Biología con énfasis en  
Genética y Biología Molecular

JUAN MIGUEL CORDERO SOLÓRZANO

Ciudad Universitaria Rodrigo Facio, Costa Rica  
2013

## DEDICATORIA

A mi esposa

A mi madre

A mi padre y hermana

## AGRADECIMIENTOS

Al Dr. Bernal León por darme la oportunidad de desarrollar mi tema de tesis dentro del marco del proyecto de marcadores moleculares en bovinos.

Al Dr. Gustavo Gutiérrez por todas sus enseñanzas durante mi maestría, así como su paciencia y confianza, sin las cuales no habría podido realizar esta tesis.

A la Dr. Ir. Marleen Visker por toda su guía, apoyo y conocimientos transmitidos durante la realización de este trabajo.

Al Dr. Bernardo Vargas por sus invaluable aportes en el desarrollo del tema y análisis de resultados.

Al Prof. Dr. Ir. Johan van Arendonk por permitirme realizar el trabajo de laboratorio en el Departamento de Ciencias Animales de la Universidad de Wageningen.

A la Dr. Ir. Liesbeth van der Waaij e Ing. Bert Dibbits por toda su ayuda durante mi estancia en Países Bajos.

Al Ministerio de Ciencia y Tecnología por financiar mediante el proyecto FI-370-2011 la realización de esta tesis.

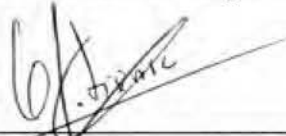
A todas las personas, compañeros y amigos que directa o indirectamente colaboraron en la elaboración de este documento.

“Esta tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Biología de la Universidad de Costa Rica, como requisito parcial para optar por el grado y título de Maestría en Biología con énfasis en Genética y Biología Molecular.”



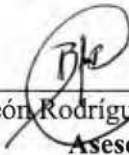
---

Marielos Mora López, M.Sc.  
**Representante de la Decana**  
**Sistema de Estudios de Posgrado**



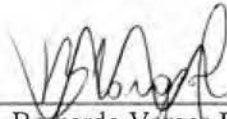
---

Gustavo Gutiérrez Espeleta, Ph.D.  
**Director de Tesis**



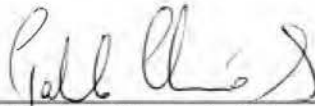
---

Bernal León Rodríguez, DMV., M.Sc.  
**Asesor**



---

Bernardo Vargas Leitón, Ph.D.  
**Asesor**



---

Gabriela Chavarría Soley, Ph.D.  
**Representante del Director**  
**Programa de Posgrado en Biología**



---

Juan Miguel Cordero Solórzano  
**Candidato**

## TABLA DE CONTENIDO

<b>DEDICATORIA</b>	<b>ii</b>
<b>AGRADECIMIENTOS</b>	<b>iii</b>
<b>HOJA DE APROBACIÓN</b>	<b>iv</b>
<b>TABLA DE CONTENIDO</b>	<b>v</b>
<b>RESUMEN</b>	<b>vii</b>
<b>LISTA DE CUADROS</b>	<b>viii</b>
<b>LISTA DE FIGURAS</b>	<b>ix</b>
<b>LISTA DE ABREVIATURAS</b>	<b>x</b>
<b>ESTADO ACTUAL DEL CONOCIMIENTO</b>	<b>1</b>
Uso de ganado bovino	1
Composición de la leche	2
Relación de $\alpha$ -lactalbúmina y síntesis de lactosa	4
Modificación del contenido de lactosa en la leche	6
Polimorfismos genéticos de las proteínas lácteas	7
Polimorfismos en el gen de $\alpha$ -lactalbúmina bovina ( <i>LALBA</i> )	8
Proyecto “Milk Genomics Initiative” de Países Bajos	9
<b>JUSTIFICACIÓN</b>	<b>11</b>
<b>HIPÓTESIS</b>	<b>12</b>
<b>PREDICCIONES</b>	<b>12</b>
<b>OBJETIVOS</b>	<b>13</b>
Objetivo General	13
Objetivos Específicos	13
<b>METODOLOGÍA</b>	<b>14</b>
Animales	14
Secuenciación	15
Genotipificación	17
Desequilibrio de ligamiento	18
Análisis estadístico	18
Heredabilidad	19

<b>RESULTADOS</b>	<b>21</b>
Amplificación del fragmento de la región promotora del gen <i>LALBA</i>	21
72 Sementales	21
1857 Vacas	21
Secuenciación	22
Genotipificación	22
72 Sementales	22
1857 Vacas	23
Desequilibrio de ligamiento	23
Polimorfismos en sementales	23
Ligamiento en muestras genotipificadas	25
Análisis estadístico	25
Heredabilidad	30
<b>DISCUSIÓN</b>	<b>31</b>
Estandarización de PCR y purificación	31
Secuenciación	32
Haplotipos y genotipificación	32
Análisis de asociación	33
$\alpha$ -Lactalbúmina	34
Caseínas totales	35
Heredabilidad	36
<b>CONCLUSIONES</b>	<b>37</b>
<b>RECOMENDACIONES</b>	<b>38</b>
<b>BIBLIOGRAFÍA</b>	<b>39</b>
<b>APÉNDICES</b>	<b>46</b>
Apéndice 1	46
Apéndice 2	47
Apéndice 3	48

## RESUMEN

Actualmente con herramientas de genética molecular y cuantitativa es posible seleccionar de forma precisa características deseables en animales domésticos, como es el caso del ganado bovino. En este proyecto se analizó un fragmento de 429 pb en la región promotora del gen *LALBA* en 2000 animales Holstein Friesian del proyecto “Milk Genomics Initiative” de Países Bajos por medio de PCR-RFLP y secuenciación. Se encontraron dos polimorfismos: -1691(g.G>C) y -1689(g.G>A). El SNP -1689 está en desequilibrio de ligamiento con otro polimorfismo en la posición +15(g.G>A), pero no está asociado con la producción de  $\alpha$ -lactalbúmina. Se confirmó la asociación del SNP +15 con  $\alpha$ -lactalbúmina y caseínas totales como se había reportado previamente. Esto permitirá seleccionar animales que brinden leche con mayor contenido proteico.

## LISTA DE CUADROS

Cuadro 1. Mezcla de reacción de PCR para amplificación de fragmento de 429pb	15
Cuadro 2. Mezcla de reacción de secuenciación	16
Cuadro 3. Mezcla de reacción para la digestión con enzima SduI	17
Cuadro 4. Variables dependientes analizadas en el modelo	20
Cuadro 5. Frecuencias de polimorfismos en región promotora de gen <i>LALBA</i> detectados por secuenciación	22
Cuadro 6. Frecuencias de polimorfismo -1689 en muestras de vacas, obtenidos por PCR-RFLP	23
Cuadro 7. Haplotipos abarcando 12 polimorfismos en diferentes regiones del gen <i>LALBA</i>	24
Cuadro 8. Haplotipos de los SNPs -1689 y +15 en las muestras de ganado	25
Cuadro 9. Valores de significancia para cada variable con respecto a $\alpha$ -LA utilizando modelo sin SNP +15	25
Cuadro 10. Valores de significancia para cada variable con respecto a $\alpha$ -LA utilizando modelo con SNP +15	26
Cuadro 11. Valores de significancia con respecto a índice de caseínas totales utilizando modelo con SNP +15	28
Cuadro 12. Valores de significancia con respecto a otras variables analizadas para los SNP +15 y -1689 y días de lactancia	29
Cuadro 13. Valores de heredabilidad para las variables asociadas significativamente con los SNP analizados	30



## LISTA DE FIGURAS

Figura 1. Amplificación del fragmento de 429 pares de bases en 7 muestras	21
Figura 2. Digestión del fragmento de 429 pares de bases en 8 muestras	22
Figura 3. Representación gráfica del análisis de ligamiento en los 12 SNP	24
Figura 4. Valores de concentración de $\alpha$ -LA según el genotipo	26
Figura 5. Comparación de los valores por genotipo de $\alpha$ -LA para el SNP -1689	27
Figura 6. Valores de concentración de lactosa según el genotipo para el SNP +15	27
Figura 7. Valores de índice de caseínas totales según el genotipo para el SNP +15	28

## LISTA DE ABREVIATURAS

**$\alpha$ -LA:**  $\alpha$ -Lactalbúmina

**ADN:** Ácido desoxirribonucleico

**ARNm:** Ácido ribonucleico mensajero

**$\beta$ -LG:**  $\beta$ -Lactoglobulina

**Ca<sup>2+</sup>:** Calcio divalente

**EDTA:** Ácido etilendiaminotetraacético

**EST:** Marcador de secuencia expresada, del inglés Expressed Sequence Tag

**EtBr:** Bromuro de etidio

**kb:** Kilobase

**kDa:** Kilodalton

***LALBA*:** Gen de  $\alpha$ -Lactalbúmina

**MM:** Marcador de peso molecular

**nm:** Nanómetro

**pb:** Par de bases

**PCR:** Reacción en cadena de la polimerasa, del inglés Polymerase Chain Reaction

**QTL:** Locus de rasgos cuantitativos, del inglés Quantitative Trait Locus

**RFLP:** Polimorfismo de longitud de fragmentos de restricción, del inglés Restriction Fragments Length Polymorphism

**SNP:** Polimorfismo de un solo nucleótido, del inglés Single Nucleotide Polymorphism

**TBE:** Solución tampón de tris base, ácido bórico y EDTA

**UTR:** Región no traducida, del inglés Untranslated Region

## ESTADO ACTUAL DEL CONOCIMIENTO

Se denomina mamíferos a los animales que poseen pelo y glándulas mamarias. Se clasifican dentro del dominio Eukarya, reino Animalia, filo Chordata, subfilo Vertebrata, clase Mammalia (Gillespie & Flanders 2009). En la actualidad hay cerca de 4 500 especies existentes (Benton 1999).

Las crías de la mayoría de especies nacen vivas (vivíparos) y son provistas con un alimento especial en el período crítico luego del nacimiento: la leche. Dicha sustancia es primordial para la supervivencia, desarrollo y crecimiento adecuado de los neonatos. Es la única fuente de agua, nutrientes orgánicos y minerales a los cuales tiene acceso la cría (Fox 2009).

La leche contiene factores de protección para el neonato, tales como inmunoglobulinas y ciertas enzimas (Baumrucker 2008, Goldman 2002). Sin embargo, la cantidad y composición varían de forma considerable entre especies e incluso individuos de la misma especie (Jensen 2005, Langer 2009). Esto responde al hecho de que debe suplir los requerimientos específicos de la cría.

El ser humano ha consumido la leche de otras especies por al menos 8 000 años (Solounias *et al.* 1995, Zenger *et al.* 2007), no como sustituto de la leche materna sino como un alimento más. Se han utilizado muchas especies para la producción de este líquido, pero actualmente el ganado bovino es la principal fuente con el 85% de la producción total (Lamy *et al.* 2012).

### **Uso de ganado bovino**

Dentro del orden Artiodactyla (ungulados de pezuñas pares o hendida) se incluye la familia Bovidae, la cual comprende el género *Bos* spp. Se han descrito ocho especies en este género, de las cuales destacan: *B. primigenus* (uro, ancestro del ganado bovino moderno; actualmente extinto), *B. taurus* (ganado doméstico) y *B. indicus* (ganado cebú) (Benton 1999).

Tanto el ganado cebú como el doméstico se utilizan en la industria bovina. Sin embargo, *B. indicus* es más resistente al estrés por calor y a diversas enfermedades (Berman 2011, Lamy *et al.* 2012). Por su parte *B. taurus* se utiliza principalmente para la obtención de leche, por lo que las razas lecheras pertenecen a esta especie. Hoy en día hay aproximadamente 1300 millones de especímenes de *B. taurus* y *B. indicus* alrededor del mundo (Fox 2009).

De los 200 millones de animales de producción lechera en el mundo, la raza Holstein Friesian representa el 35% siendo la más numerosa, seguido de Pardo Suizo 2% (4 millones), Jersey 1% (2 millones) y otros como Ayshire y Guernsey (Zenger *et al.* 2007, Sheehy 2009). La principal razón por la cual predomina Holstein Friesian es debido a que posee una mayor capacidad de producción de leche con respecto a otras razas (De Marchi *et al.* 2008, Prendiville *et al.* 2010).

### **Composición de la leche**

La leche es un líquido muy complejo que contiene cientos de especies moleculares. Sus principales constituyentes son: el agua, lípidos, azúcares y proteínas. Existen componentes traza como minerales, vitaminas, hormonas, enzimas y otros (Jensen 1995). La concentración de los componentes principales varía mucho según la especie; los lípidos pueden rondar de 2 a 55%, la proteínas de 1 a 20%, y la lactosa de 0 a 10% (Baumrucker 2008, Fox 2009).

El contenido de grasa en la leche era considerado el componente más preciado e incluso se utilizaba para determinar gran parte de su valor hasta hace algún tiempo. Esto quizá debido al desarrollo de metodologías sencillas para determinar su contenido total a finales del siglo XIX, mucho antes de que existieran procedimientos similares para las proteínas (Fox 2009).

En la leche, los lípidos se hayan en forma de glóbulos dispersos en el suero. Los ácidos grasos y monoglicéridos son convertidos a triglicéridos en el retículo endoplasmático de las células epiteliales ubicadas en el extremo basal. Los triglicéridos adquieren la forma de glóbulos en esta

organela y luego son liberados al citoplasma. Son estabilizados por una capa compleja de proteínas y fosfolípidos conocida como la membrana globular de grasa de la leche (MFGM por sus siglas en inglés) (Evers 2004).

El principal carbohidrato en la leche de la mayoría de especies es el disacárido lactosa, compuesto por una molécula de galactosa y otra de glucosa, unidos por un enlace  $\beta$ 1,4 glucosídico. Esta unión se produce por la acción del complejo enzimático lactosa sintetasa, formado por dos moléculas, una de ellas es UDP-galactosiltransferasa, que transfiere galactosa a partir de UDP-galactosa a cualquier molécula receptora en la biosíntesis de glicoproteínas y glicolípidos. La otra molécula es  $\alpha$ -lactalbúmina ( $\alpha$ -LA), la cual confiere especificidad al complejo ya que favorece la transferencia de galactosa hacia glucosa (Martin *et al.* 2002).

Las proteínas afectan las propiedades de la leche y la mayoría de productos lácteos de manera más significativa que cualquier otro constituyente. La fracción proteínica de la leche está compuesta por un alto número de proteínas específicas. El grupo principal de proteínas son las caseínas, de las cuales existen cuatro tipos:  $\alpha_{s1}$ ,  $\beta$ ,  $\alpha_{s2}$  y  $\kappa$ . A las demás proteínas encontradas en la leche se les denomina proteínas del suero (Jensen 1995, Sheehy *et al.* 2009).

Las proteínas del suero se dividen en dos grupos; las lactalbúminas que son solubles en solución saturada con sulfato de magnesio ( $MgSO_4$ ) o saturada al 50% con sulfato de amonio ( $(NH_4)_2SO_4$ ) y las lactoglobulinas que precipitan en estas condiciones. La fracción de lactoglobulinas contiene principalmente inmunoglobulinas. La fracción de lactalbúminas se compone principalmente de  $\alpha$ -lactalbúmina ( $\alpha$ -LA) y  $\beta$ -lactoglobulina ( $\beta$ -LG), además de otras proteínas menores, incluida albúmina de suero sanguíneo y lactoferrina (Ha & Zemel 2003).

La  $\beta$ -LG representa el 50% de las proteínas del suero y 12% de la proteína total en la leche bovina. En su estado natural es muy resistente a la proteólisis, lo que sugiere que su rol no es principalmente nutricional. Dentro de sus funciones conocidas está adherirse al retinol (vitamina A) para protegerlo de la oxidación y transportarlo del estómago al intestino donde es transferido a un

receptor, además de unirse a ácidos grasos estimulando así su degradación (Evers 2004). A pesar de esto, es el mayor alérgeno presente en la leche bovina para bebés humanos, por lo cual debe estar ausente en fórmulas (Selo *et al.* 1999, Natale *et al.* 2004).

Cerca del 20% de la proteína de suero bovino (3.5% de la proteína total de la leche) es  $\alpha$ -LA. Se compone de 123 aminoácidos con una masa aproximada de 14.2 kDa. Contiene cuatro residuos de triptófano por mol, lo que le confiere una absorbancia específica de 20 a 260 nm. Contiene cuatro puentes de disulfuro intramoleculares por mol pero no tiene cisteína, fosfato o carbohidratos (McKenzie & White 1991, Brew 2003).

### **Relación de $\alpha$ -lactalbúmina y síntesis de lactosa**

Hasta el momento se han descrito cuatro variantes de  $\alpha$ -LA en bovinos (Visker *et al.* 2012) pero sólo las variantes B y D están presentes en *Bos taurus*. Su estructura primaria es homóloga a la lisozima, de sus 123 residuos, 54 son idénticos a los aminoácidos correspondientes en la lisozima de las gallinas, y otros 23 aminoácidos son estructuralmente similares (Ha & Zemel 2003). Se cree que la  $\alpha$ -LA evolucionó a partir de la duplicación de genes de lisozima.

Dicha proteína es un componente del complejo enzimático lactosa sintetasa, el cual cataliza la síntesis de lactosa.  $\alpha$ -LA es una metaloproteína que contiene un ion de  $\text{Ca}^{2+}$  por mol en un área con cuatro residuos de ácido aspártico (Asp). Posee la característica de que si es desnaturizada por calor recupera su estructura original una vez disminuida la temperatura. Esto no ocurre si se desnaturiza disminuyendo el pH a valores menores o iguales a 5, ya que los residuos de Asp se protonan y pierden su habilidad de unirse al calcio. La proteína sin el metal se desnaturiza y no recupera su estructura (Sheehy *et al.* 2009). Dicha propiedad se utiliza para aislar  $\alpha$ -LA del suero.

Existe una correlación positiva entre la concentración de lactosa y  $\alpha$ -LA en la leche; por ejemplo, el león marino californiano (*Zalophus californianus*) no sintetiza lactosa y carece de  $\alpha$ -LA

(Pilson & Kelly 1962, Johnson *et al.* 1972). Además, experimentos con ratas han demostrado que a los animales que se les brinda una dieta con insuficiente cantidad de proteínas producen menos leche con menor cantidad de lactosa y  $\alpha$ -LA, sin embargo la cantidad de galactosiltransferasa se mantiene constante (Grimble *et al.* 1987).

La síntesis de lactosa lleva agua hacia las vesículas del aparato de Golgi (ya que la UDP-galactosiltransferasa es una enzima adherida a la membrana de esta organela) y por lo tanto afecta el volumen de la leche y la concentración de caseína, la cual es empacada en estas vesículas (Bleck *et al.* 2009, Kuhn & White 2009). Existe una correlación inversa entre la concentración de lactosa y la de cantidad de caseína en la leche.

Este carbohidrato posee un bajo nivel de dulzura; comprende tan sólo un 16% con respecto a la sacarosa en una solución al 1%, por lo que no tiene mucho valor como edulcorante, la principal aplicación de los azúcares en las comidas. Sin embargo, juega un papel fundamental en la producción de derivados lácteos ya que sirve como fuente de carbono para bacterias que sintetizan ácido láctico (Sheehy *et al.* 2009).

La lactosa sirve como fuente rápida de energía para el neonato, proveyendo el 30% de las calorías en la leche bovina y sirviendo como alternativa para los lípidos, los cuales son más densos en cuanto a cantidad de energía (1 g de grasa equivale a 9 calorías mientras que la misma cantidad de carbohidratos corresponde a 4 calorías). De forma general, la leche con concentraciones más altas de lactosa, tienden a tener menor cantidad de lípidos (Van Kneysel *et al.* 2007).

La leche mantiene la misma presión osmótica que la sangre (isotónica) y al ser la lactosa responsable del 50% de la presión osmótica de la leche, su concentración se mantiene esencialmente constante. La leche con menos lactosa, suele tener mayor concentración de sales para mantener la presión osmótica al mismo nivel que la sangre; se observa una correlación negativa entre estos dos factores (Fox 2009). Esto indica que la  $\alpha$ -lactalbúmina posee un rol importante en la regulación del volumen de la leche.

## Modificación del contenido de lactosa en la leche

Los mamíferos no pueden absorber disacáridos en el intestino delgado, estos deben ser hidrolizados a monosacáridos. La enzima responsable de digerir la lactosa es la  $\beta$ -galactosidasa (conocida también como lactasa o  $\beta$ -gal) secretada por las células del intestino delgado (Sheehy *et al.* 2009). Las crías de la mayoría de mamíferos secretan cantidades abundantes de esta enzima, pero conforme el animal se desarrolla, la producción disminuye y eventualmente se vuelve insuficiente (en humanos suele ocurrir entre los 8 y 10 años) (Zenger *et al.* 2007). Si la lactosa no es hidrolizada y llega al intestino grueso, absorbe agua e induce diarrea, además de ser metabolizada por enterobacterias, provocando gases, calambres y flatulencia (Ingram y Swallow 2007).

Para evitar esto se fabrica leche deslactosada prehidrolizando la lactosa mediante el uso de  $\beta$ -gal exógena proveniente de bacterias como *Lactobacillus casei* y *L. acidophilus*. Sin embargo, esto altera el sabor al liberar los monosacáridos galactosa y glucosa que son dulces al paladar (Sheehy *et al.* 2009).

Además de lo expuesto anteriormente, la lactosa es el componente con el valor más bajo en términos económicos y tiene un costo energético por parte del animal para sintetizarlo. Es importante mantener un balance energético adecuado (calorías consumidas versus gastadas) y si se utilizan calorías adicionales, se debe proveer alimento adicional, por lo que sería ventajoso también desde el punto de vista productivo, disminuir la concentración de lactosa (Van Kneysel *et al.* 2007).

En muchos procesos de la industria de los lácteos se remueve totalmente el agua. Es por eso que resultaría ventajoso disminuir el contenido de agua reduciendo la concentración de lactosa. Sin embargo, si se reduce demasiado, la viscosidad de la leche sería muy alta dificultando su salida de las glándulas mamarias (Boston *et al.* 2001).

Considerando entonces que  $\alpha$ -LA regula la síntesis de lactosa y ésta a su vez afecta el volumen de leche y la cantidad de caseína, resultaría conveniente reducir la producción de  $\alpha$ -LA para obtener



mayor cantidad de proteína y menos agua (reduciendo el estrés para el animal) lo que incrementaría el valor de dicha leche en el mercado (Schopen *et al.* 2011, Visker *et al.* 2011).

### **Polimorfismos genéticos de las proteínas lácteas**

Desde el descubrimiento de dos variantes de  $\beta$ -LG en 1955, se han realizado innumerables estudios en busca de polimorfismos en las proteínas de la leche. Las variantes genéticas son de especial interés ya que van asociadas a la producción, composición y otras propiedades de la leche. Al ser heredados de forma mendeliana, es posible seleccionar variantes específicas (Martin *et al.* 2002).

La genética de bovinos, enfocada en la búsqueda de loci de rasgos cuantitativos (QTL, por sus siglas en inglés) ha sido un área fundamental. Por medio de la genética clásica en conjunto con la genética molecular, más de 1000 genes ya habían sido mapeados a mediados de los años noventa. Posteriormente se realizaron estudios más comprensivos para generar marcadores de secuencias expresadas (EST por sus siglas en inglés), proveyendo un índice de más de 100000 secuencias. Más recientemente el proyecto de secuenciación del genoma bovino permitió conocer las 2700 millones de bases (pb) que componen el material genético de los bovinos (Khatkar *et al.* 2007), representando un avance importante en el descubrimiento de QTLs a través de la genómica comparativa, funcional y asociaciones de genoma completo.

Identificar polimorfismos en el genoma bovino es importante para entender los procesos que definen las características de la leche. El proyecto del genoma permitió conocer un gran número de polimorfismos de un solo nucleótido (SNP por sus siglas en inglés). Y con la posterior secuenciación de otras razas (Holstein, Angus, Jersey, Brahman, entre otras) este número se incrementó aún más (Khatkar *et al.* 2007).

Los primeros estudios en razas lecheras utilizando SNPs determinaron que la raza predominante –Holstein Friesian– posee una variabilidad genética muy baja (Zenger *et al.* 2007). El promedio de las frecuencias de alelos fue de 0.29 basado en un análisis de 9195 SNP en 1000 animales. También en otro estudio se determinó que el desequilibrio de ligamiento ocurre en mayor medida que en humanos. Al construir los haplotipos identificaron 727 bloques (cubriendo el 2.18% del genoma) con una longitud promedio de 2.9 kilobases (kb) (Khatkar *et al.* 2007).

El genoma bovino se compone de 29 pares cromosomas autosómicos y 1 par de cromosomas sexuales. Aproximadamente el 95% del volumen de proteínas en la leche de los rumiantes, se encuentra codificada en seis genes que abarcan  $\alpha$ -LA,  $\beta$ -LG y cuatro variantes de caseína ( $\alpha_{s1}$ ,  $\beta$ ,  $\alpha_{s2}$  y  $\kappa$ ). El gen de  $\alpha$ -LA (*LALBA*) en bovinos se encuentra en el locus 5q21, contiene cuatro exones, tres intrones y posee una longitud de 2 kb aproximadamente (Vilotte *et al.* 1987, Martin *et al.* 2002).

### **Polimorfismos en el gen de $\alpha$ -lactalbúmina bovina (*LALBA*)**

Un SNP ubicado en la región no traducida (UTR) al extremo 5' del gen *LALBA*, conocido como +15 (g.G>A) relativo al punto de inicio de la transcripción (tsp), inicialmente reportado por Bleck y Bremel (1993a), se ha asociado con diferencias en ciertos componentes de la leche (Bleck y Bremer 1993b, Lundén y Lindersson 1998, Schopen *et al.* 2011, Visker *et al.* 2011). En un estudio reciente realizado por Visker *et al.* (2011) se ha asociado este SNP con la concentración de  $\alpha$ -LA y cantidad relativa de caseína.

Se ha reportado también que otro SNP corriente arriba del gen *LALBA* en la posición -1689, correspondiente a una sustitución de guanina por adenina, podría estar asociado con la composición de la leche (Lundén y Lindersson 1998). Sin embargo, este polimorfismo parece encontrarse en

desequilibrio de ligamiento con el SNP+15 (Voelker *et al.* 1997) y no está del todo claro si hay una asociación o no entre el SNP-1689 y características de la leche.

Dos estudios independientes obtuvieron resultados distintos, uno encontrando asociación (Lundén y Lindersson 1998) y el otro no (Voelker *et al.* 1998). En el caso de Lundén y Lindersson (1998) se analizaron únicamente 78 animales mientras que Voelker *et al.* (2008) analizaron 162 individuos Holstein, pero no explican el modelo estadístico utilizado.

### **Proyecto “Milk Genomics Initiative” de Países Bajos**

En el 2005 comenzó un proyecto en la Universidad de Wageningen, Países Bajos cuyo objetivo es identificar los genes que contribuyen a la variabilidad genética en rasgos relacionados con la calidad de la leche, especialmente la composición de grasas y proteínas (<http://www.milkgenomics.nl>) Combina estudios en el campo de la leche, genética cuantitativa, genómica y bioinformática.

En dicho proyecto se tomaron muestras de leche y sangre de una población de 2000 vacas Holstein Friesian en su primera lactancia. Los animales pertenecían a 400 hatos comerciales de Países Bajos y se muestrearon 5 vacas por hato. Estas muestras se han utilizado para estimar parámetros genéticos así como para detectar QTLs y realizar análisis de asociación.

Se ha determinado la composición detallada de la leche en las muestras para variables como grasa, proteínas, lactosa, minerales, anticuerpos, urea, conteo de células somáticas y otros, utilizando técnicas como espectroscopía infrarroja, cromatografía de gases y electroforesis capilar de zona (CZE por sus siglas en inglés).

Asimismo, se han estudiado polimorfismos en genes candidatos buscando asociación con la composición de grasa y proteínas principalmente. Dichos estudios contemplan polimorfismos identificados en investigaciones previas así como polimorfismos detectados en este proyecto.

Esta tesis se enmarca dentro del proyecto “Milk Genomics Initiative” y pretende determinar si el SNP -1689 está presente en las 2000 muestras de vacas Holstein Friesian y si hay o no asociación entre este polimorfismo y diferentes características de la leche.

Fue financiada mediante el proyecto “Uso de marcadores moleculares para identificación y verificación de parentescos en ganado bovino”, que se desarrolla en el Servicio Nacional de Salud Animal (SENASA) de Costa Rica, en conjunto con la Universidad Nacional (UNA) y la Universidad de Costa Rica (UCR), cuyo objetivo es estudiar la variabilidad genética del hato costarricense para implementar ensayos de paternidad e identidad.

## JUSTIFICACIÓN

Con una producción mundial aproximada de 600 millones de toneladas al año (FAO Statistics Division 2010), la leche de ganado bovino representa un recurso nutricional importante para el ser humano. Dicho recurso debe adaptarse a las necesidades y gustos de la población humana; baja en grasa y alta en proteínas, con menos lactosa, sin  $\beta$ -lactoglobulina (especialmente para fórmulas), con cierta textura y otras propiedades bioquímicas y organolépticas (Gillespie y Flanders 2009).

Adaptar esas propiedades requiere de un estudio detallado de los elementos implicados en su variación. La genética clásica en conjunto con la genética molecular ha permitido determinar que seis genes ( $\alpha$ -LA,  $\beta$ -LG y cuatro variantes de caseína) codifican para el 95% del volumen de proteínas en la leche de los rumiantes (Vilotte *et al.* 1987, Martin *et al.* 2002). Identificar polimorfismos en estos genes es de vital importancia para entender los procesos que definen la composición proteínica de la leche (Khatkar *et al.* 2007).

Estudiar el polimorfismo propuesto en este proyecto compondría un eslabón más en el proceso de esclarecer los factores genéticos que afectan la concentración de  $\alpha$ -lactalbúmina en la leche bovina. Esta proteína es de particular interés ya que regula la síntesis de lactosa. Si se logra reducir este carbohidrato, podría obtenerse leche con mayor cantidad de proteína, menos agua (lo que disminuye el estrés para el animal) y más digerible para personas poco tolerantes a la lactosa. Además de las implicaciones nutricionales, una leche con mayor contenido proteínico se vende a un precio más alto, favoreciendo el desarrollo de la actividad lechera (Schopen *et al.* 2011).

Así mismo, la elaboración de leche sin lactosa, la cual suele hacerse digiriendo este disacárido con la enzima  $\beta$ -galactosidasa, altera el sabor al liberar los monosacáridos galactosa y glucosa que son dulces al paladar. Sin embargo, si la concentración de lactosa es menor, al realizar este proceso, podría generarse una leche deslactosada con un sabor más similar al original.

A nivel nacional, este proyecto significa un primer esfuerzo por incluir la genética molecular dentro de los programas de mejoramiento de ganado lechero. A pesar de que los análisis de laboratorio son realizados con animales Holstein Friesian de Países Bajos, en Costa Rica la industria lechera utiliza en su mayoría esta misma raza, por lo que la información generada podría aplicarse en nuestro país.

## HIPÓTESIS

El polimorfismo -1689 en la región promotora del gen *LALBA*, afecta su expresión, disminuyendo la cantidad de  $\alpha$ -LA sintetizada.

## PREDICCIONES

- El polimorfismo -1689 de la región 5'UTR del gen *LALBA* está presente en las muestras de ganado Holstein Friesian del proyecto "Milk Genomics Initiative".
- Existe desequilibrio de ligamiento (LD) entre el SNP -1689 y +15 del gen *LALBA*.
- Hay asociación entre el polimorfismo -1689 y la composición proteínica de la leche.

## OBJETIVOS

### Objetivo General

- Determinar si el polimorfismo -1689 en el promotor del gen de  $\alpha$ -Lactabúmina bovina está asociado con la composición proteínica de la leche en ganado Holstein Friesian de Países Bajos.

### Objetivos Específicos

1. Secuenciar un fragmento de 429 pb en la región promotora del gen de  $\alpha$ -LA en 72 sementales Holstein Friesian holandeses para determinar si el polimorfismo -1689 está presente y en desequilibrio de ligamiento con el SNP +15.
2. Analizar 1857 muestras de ganado Holstein Friesian del proyecto “Milk Genomics Initiative” utilizando un ensayo de PCR-RFLP para genotipificar el SNP -1689.
3. Establecer si existe asociación entre el polimorfismo -1689 y la composición de la leche en las muestras del proyecto “Milk Genomics Initiative” de Países Bajos.

## METODOLOGÍA

Las muestras se analizaron en el Departamento de Ciencias Animales de la Universidad de Wageningen, Países Bajos donde se desarrolla el proyecto “Milk Genomics Initiative”.

### **Animales**

Se utilizaron extracciones de ADN genómico, realizadas previamente a partir de semen de 72 toros Holstein Friesian holandeses (Visker *et al.* 2012), de los cuales 22 son sementales probados y el resto sementales jóvenes, para determinar la presencia del polimorfismo -1689 en el gen de  $\alpha$ -lactalbúmina.

Posteriormente se analizaron extracciones de ADN a partir de sangre (procesadas con el kit DNeasy 96 Blood & Tissue de QIAGEN siguiendo las instrucciones del fabricante) obtenidas de un proyecto anterior (Schopen *et al.* 2009) de 1857 vacas Holstein Friesian de primera lactancia. Dichos animales provienen de 398 hatos en Países Bajos muestreados para el proyecto “Milk Genomics Initiative” el cual cuenta con la información del fenotipo de la leche (ver apartado Análisis estadístico para detalles) y genealogía para cada una de las muestras (Stoop *et al.* 2007).

Previo a esto, las 1857 extracciones fueron purificadas utilizando el kit OneStep™ PCR Inhibitor Removal (Zymo Research) siguiendo las instrucciones del fabricante, ya que se extrajeron desde el 2005 y se removieron inhibidores potenciales para garantizar su funcionalidad en PCR. Después se cuantificó la concentración de ADN de 15 muestras seleccionadas al azar utilizando un NanoDrop 2000 (Thermo Scientific).



## Secuenciación

Se analizaron 72 sementales para determinar si presentaban el polimorfismo -1689 y confirmar si la técnica de genotipificación por PCR-RFLP es precisa (debe haber un 100% de concordancia entre los resultados) y poder utilizarla con las 1857 vacas.

Se amplificó un fragmento de 429 pares de bases comprendiendo las posiciones -1951 a -1523 de la región promotora del gen de  $\alpha$ -lactalbúmina, utilizando los imprimadores descritos por Voelker *et al.* (1997) con la mezcla OneTaq MasterMix 2X (BioLabs Inc.) (Cuadro 1). Antes de amplificar las 72 muestras se realizaron varias pruebas para estandarizar el PCR.

CUADRO 1

Mezcla de reacción de PCR para amplificación de fragmento de 429pb

Reactivo	Concentración de trabajo	Concentración final	Volumen ( $\mu$ l)
OneTaq Master Mix	2X	1X	6
Iniciador 1951-F	50 $\mu$ M	1 $\mu$ M	0.25
Iniciador 1523-R	50 $\mu$ M	1 $\mu$ M	0.25
ADN Muestra	—	—	5.5

El programa de corrida para la amplificación se nombró “Knor” y las condiciones fueron las siguientes: 5 min a 95 °C, 35 ciclos de 30 s a 95 °C, 30 s a 56 °C y 30 s a 72 °C con una extensión final de 5 min a 72 °C en un termociclador TPersonal (Biometra GmbH).

Para verificar la amplificación, las muestras se corrieron en un gel de agarosa al 1 % teñido con EtBr en TBE 1X por 1 hora y 20 minutos a 80 voltios, agregando 2  $\mu$ l de buffer de carga 6X azul (BioLabs Inc.) a cada reacción y posteriormente cargando 2  $\mu$ l de muestra en el gel, como marcador de peso molecular se usó 2  $\mu$ l de marcador de 100 pb (BioLabs Inc.).

Luego de confirmar la amplificación, las reacciones de PCR se purificaron por medio de placa de 96 pozos MultiScreen HTS PCR (EMD Millipore) siguiendo las instrucciones del fabricante con modificaciones (Apéndice 1). Luego las amplificaciones purificadas se montaron en la reacción de secuenciación utilizando el kit Big Dye® Terminator v3.1 (Cuadro 2).

CUADRO 2

## Mezcla de reacción de secuenciación

Reactivo	Concentración de trabajo	Concentración final	Volumen (μl)
Agua libre de nucleasas	—	—	5.34
Solución tampón de secuenciación	5X	0.75X	1.5
Mezcla Ready Reaction	2.5X	0.25X	1
Iniciador sentido o antisentido	50 μM	0.8 μM	0.16
ADN Muestra	—	—	2

El programa de secuenciación en el termociclador se corrió con las siguientes condiciones: 1 min a 96 °C, 30 ciclos de 10 s a 96 °C, 5 s a 62 °C y 4 min a 60 °C, y se nombró “BDMG”.

Posterior a la reacción de secuenciación se realizó la segunda purificación con acetato de sodio (NaAc), EDTA y etanol (Apéndice 2). Con esta purificación las muestras quedaron resuspendidas en formamida y se desnaturalizaron calentando la placa a 95 °C por 5 minutos e inmediatamente colocándola en hielo por 2 minutos. Luego de eso se centrifuga y se coloca en el secuenciador para el procesamiento.

Las muestras fueron analizadas en un secuenciador ABI 3730 (Applied Biosystems) utilizando un capilar de 36 cm, polímero POP7, módulo de corrida para secuenciación estándar con Big Dye v3.1 sin modificaciones y protocolo de análisis para secuenciación *de Novo* con algoritmo de identificación de bases KB.

Las secuencias consenso fueron construidas utilizando el software SeqScape 2.7 (Applied Biosystems) tomando como secuencia de referencia las bases 1 a 429 de la secuencia M90645 de Genbank.

### Genotipificación

Las 72 muestras de sementales del paso anterior y las 1857 muestras de vacas se tipificaron utilizando un ensayo de PCR-RFLP descrito por Voelker *et al.* (1997) donde se amplifica el mismo fragmento de 429 pares de bases y luego es digerido con la enzima de restricción Bsp1286I (SduI) (Thermo Scientific), utilizando la mezcla de reacción descrita en el Cuadro 3.

CUADRO 3

Mezcla de reacción para la digestión con enzima SduI

Reactivo	Concentración de trabajo	Concentración final	Volumen ( $\mu$ l)
Agua libre de nucleasas	—	—	4.99
Solución tampón para SduI	10X	1X	1
Enzima SduI	10 U/ $\mu$ l	0.1 U/ $\mu$ l	0.1
Producto de PCR	—	—	4

La mezcla se incubó a 37 °C por 3 horas seguido de 20 minutos a 65 °C en el termociclador TPersonal con el programa denominado “Digest37”. Para verificar la digestión, las muestras se corrieron en un gel de agarosa al 1,5 % teñido con EtBr en TBE 1X por 1 hora y 20 minutos a 80 voltios, agregando 2  $\mu$ l de buffer de carga 6X azul (BioLabs Inc.) a cada reacción y posteriormente cargando 9  $\mu$ l de digestión en el gel, como marcador de peso molecular se usó 2  $\mu$ l de marcador de 100 pb (BioLabs Inc.).

Cuando la muestra cuenta con el genotipo AA para el polimorfismo -1689 no hay digestión y debe observarse la banda de 429 pb, si tiene el genotipo GG, la enzima digiere el amplicón y se observan únicamente fragmentos de 260 y 169 pares de bases, en caso de ser heterocigota los tres fragmentos deben ser visibles.

### Desequilibrio de ligamiento

Para establecer si existe desequilibrio de ligamiento entre los polimorfismos +15 y -1689 de la región promotora del gen de  $\alpha$ -lactalbúmina se utilizó el software PHASE v2.1.1 (Stephens & Scheet 2005) para construir los haplotipos y HaploView v4.2 (Barrett *et al.* 2005) para calcular D', LOD y generar el gráfico de desequilibrio de ligamiento. Se tomó la información de genotipificación de una investigación previa (Visker *et al.* 2012) donde se analizaron otros 10 polimorfismos (incluido el SNP +15) en los mismos 72 sementales. Se probaron 10 000, 100 000 y 1 000 000 de iteraciones para el análisis con PHASE.

Posteriormente se realizó el mismo análisis para las 1857 muestras de vacas, pero tomando en cuenta únicamente los polimorfismos +15 y -1689.

### Análisis estadístico

Para determinar si existe asociación entre el polimorfismo -1689 y la producción de  $\alpha$ -lactalbúmina en leche, además de otros componentes, se usó el paquete ASREML-R (Butler *et al.* 2006) dentro del software R (R Development Core Team 2011).

Para dicho análisis se utilizó el siguiente modelo lineal mixto tomado de Schopen *et al.* (2011):

$$y_{ijklmno} = \mu + b_1 * lactst_i + b_2 * e^{-0.05slactst_i} + b_3 * ca_j + b_4 * ca_j^2 + season_k + scode_l + SNP_o + animal_m + herd_n + e_{ijklmn}$$

Donde  $y_{ijklmno}$  es la variable dependiente (Cuadro 4),  $\mu$  es la media global,  $lactst_i$  una covariable describiendo el efecto de días en lactancia mediante modelo exponencial,  $ca_j$  la covariable del efecto de la edad al primer parto (años) mediante modelo cuadrático,  $season_k$  corresponde al efecto fijo con tres clases para la temporada de parto (Junio-Agosto 2004, Setiembre-Noviembre 2004 y Diciembre 2004-Febrero 2005),  $scodse_l$  es el efecto fijo tomando en cuenta posibles diferencias a nivel genético entre sementales probados y jóvenes (3 clases),  $SNP_p$  corresponde al efecto fijo del polimorfismo -1689 (AA, AG y GG),  $animal_m$  es el efecto genético aditivo aleatorio del animal m,  $herd_n$  es el efecto aleatorio del hato y  $e_{ijklmn}$  el efecto residual aleatorio. También se incluyó otra variable dentro del modelo con el efecto fijo del polimorfismo +15 para descartar cualquier posible falsa asociación del polimorfismo -1689 con la concentración de  $\alpha$ -LA.

Para el umbral de significancia, el cálculo se realizó tomando los valores de  $p$  obtenidos del modelo para el SNP -1689 en cada una de las características fenotípicas de la leche. Se analizó la concentración de  $\alpha$ -LA principalmente, pero al contar con un fenotipo detallado de la leche por el proyecto “Dutch Milk Genomics Initiative” con los valores de grasa, proteínas totales, caseínas totales, caseínas individuales  $\alpha$ S1,  $\alpha$ S2,  $\beta$  y  $\kappa$ , lactosa, urea,  $\beta$ -LG y producción de leche, se analizó cada una de ellas. Si el valor de  $p$  es menor a 0.01 se considera significativamente asociado a la característica.

## Heredabilidad

Para las variables que resulten con asociación significativa se calculará la heredabilidad ( $h^2$ ) y la heredabilidad intrahato ( $h_{herd}^2$ ) utilizando las siguientes fórmulas:

$$h^2 = \frac{\sigma_a^2}{\sigma_{herd}^2 + \sigma_a^2 + \sigma_e^2} \quad h_{herd}^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

Donde  $\sigma_{h, c}^2$  es la varianza del hato (herd),  $\sigma_a^2$  es la varianza aditiva genética (animal) y  $\sigma_e^2$  es la varianza del residual.

CUADRO 4

Variables dependientes analizadas en el modelo

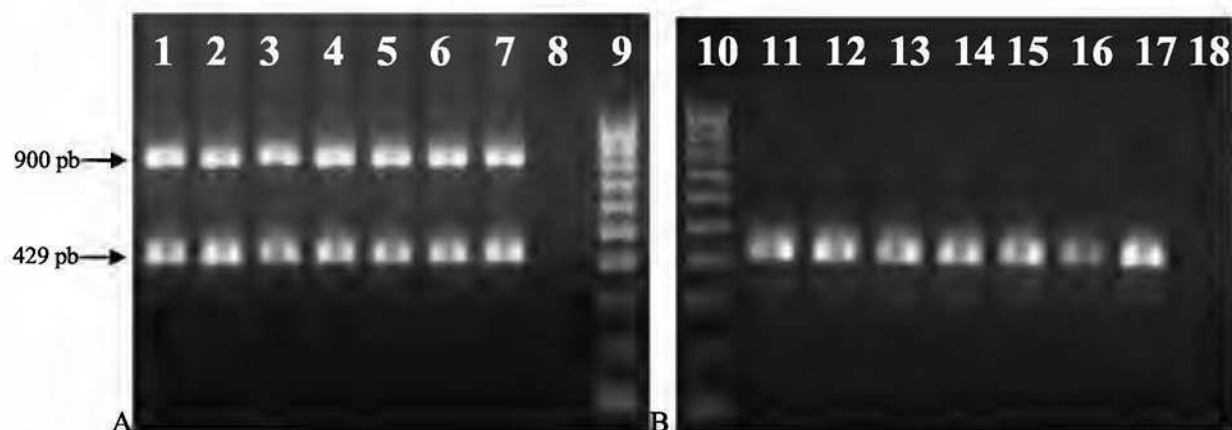
Variable	Unidad	Detalle
Porcentaje de grasa	%	masa grasa / masa total de leche
Porcentaje de proteínas totales	%	masa proteínas totales / masa total de leche
Índice grasa / proteínas totales	-	porcentaje grasa / porcentaje proteínas totales
Porcentaje de lactosa	%	masa lactosa / masa total de leche
Producción de grasa	kg	-
Producción de proteínas totales	kg	-
Producción de lactosa	kg	-
Concentración de urea	mg/dl	masa urea / volumen total de leche
Producción de leche	kg	-
Porcentaje de $\alpha$ -lactalbúmina	%	masa $\alpha$ -LA / masa proteínas totales
Porcentaje de $\beta$ -lactoglobulina	%	masa $\beta$ -LG / masa proteínas totales
Porcentaje de caseína $\alpha$ 1	%	masa $\alpha$ 1 / masa proteínas totales
Porcentaje de caseína $\alpha$ 2	%	masa $\alpha$ 2 / masa proteínas totales
Porcentaje de caseína $\kappa$	%	masa $\kappa$ / masa proteínas totales
Porcentaje de caseína $\beta$	%	masa $\beta$ / masa proteínas totales
Porcentaje de caseínas totales	%	masa caseínas totales / masa proteínas totales
Producción de caseínas totales	kg	--
Índice de caseína	—	porcentaje de caseínas totales / (porcentaje de proteínas del suero + porcentaje de caseínas totales)

## RESULTADOS

### Amplificación del fragmento de la región promotora del gen *LALBA*

#### *72 Sementales*

Para las pruebas iniciales de estandarización se tomaron al azar 7 muestras de los 72 sementales y se amplificaron con la mezcla descrita en el cuadro 1 corriendo el programa “Knor” encontrándose una fuerte banda no esperada a la altura de los 900 pb junto con la banda de 429 pb. En pruebas posteriores se aumentó gradualmente la temperatura de alineamiento hasta que dicha banda desapareció por completo a 62 °C (Fig. 1.).



**Fig. 1.** Amplificación del fragmento de 429 pares de bases en 7 muestras (A) Corrida con temperatura de alineamiento de 56 °C, Carriles 1-7 muestras, 8 Control negativo, 9 Marcador de peso molecular (MM) de 100 pb. (B) Corrida con alineamiento a 62 °C, Carril 10 MM 100 pb, 11-17 muestras, 18 Control negativo.

#### *1857 Vacas*

De las extracciones purificadas con el kit OneStep™ PCR Inhibitor Removal, se tomaron 15 muestras al azar y se cuantificaron, dando valores de 40 a 215 ng/μl con relaciones 268/280 entre 1.88 y 1.94 y de 260/230 entre 1.08 y 1.85 (Apéndice 3). Al amplificar el fragmento de 429 pares de bases en estas 15 muestras, se utilizó el programa con 62 °C de temperatura de alineamiento y en todas se observó únicamente la banda del tamaño esperado.

## Secuenciación

De las 72 muestras de sementales, sólo en 1 no se pudo obtener información de la secuencia debido a que no amplificó en la etapa de PCR.

Se encontraron dos polimorfismos; uno en la posición -1691 correspondiente a una sustitución de guanina por citosina y otro en la posición -1689 con sustitución de guanina por adenina.

CUADRO 5

Frecuencias de polimorfismos en región promotora de gen *LALBA* detectados por secuenciación

Polimorfismo	Frecuencia genotípica			Frecuencia alélica	
	Homocigotos AA	Heterocigotos AB	Homocigotos BB	A	B
-1691	0.76 (GG)	0.18 (GC)	0.06 (CC)	0.85 (G)	0.15 (C)
-1689	0.35 (AA)	0.48 (AG)	0.17 (GG)	0.59 (A)	0.41 (G)

## Genotipificación

### 72 Sementales

La tipificación por RFLP reveló los mismos resultados que la secuenciación para el polimorfismo -1689 (Cuadro 5), por lo que hubo un 100% de concordancia entre ambas técnicas para estas muestras.



**Fig. 2.** Digestión del fragmento de 429 pares de bases en 8 muestras. Carril 1 MM 100 pb, carriles 2, 5 y 6 muestras homocigotas (AA), carriles 3, 4 y 7 heterocigotas (AG), carriles 8 y 9 homocigotas (GG).



*1857 Vacas*

En el caso de las 1857 muestras, 159 no pudieron ser analizadas ya sea porque no amplificaron del todo o porque la banda era muy débil y no podía utilizarse en el RFLP.

## CUADRO 6

Frecuencias de polimorfismo -1689 en muestras de vacas, obtenidos por PCR-RFLP

Frecuencia genotípica			Frecuencia alélica	
Homocigotos AA	Heterocigotos AB	Homocigotos BB	A	B
0.31 (AA)	0.51 (AG)	0.18 (GG)	0.57 (A)	0.43 (G)

Adicional a la genotipificación con enzimas, se secuenciaron 95 muestras seleccionadas al azar entre las 302 que según la tipificación tenían el genotipo AA, para una verificación adicional y también hubo un 100% de concordancia.

**Desequilibrio de ligamiento***Polimorfismos en sementales*

Tomando 10 polimorfismos analizados por Visker *et al.* 2012 para estas mismas muestras en el gen *LALBA* (incluyendo +15), más los 2 polimorfismos detectados en este estudio se construyeron los haplotipos (Cuadro 7).

Condensando los haplotipos a sólo los polimorfismos -1691, -1689 y +15, se observa que únicamente hay 4 variantes: GAA, GAG, GGG y CAG.

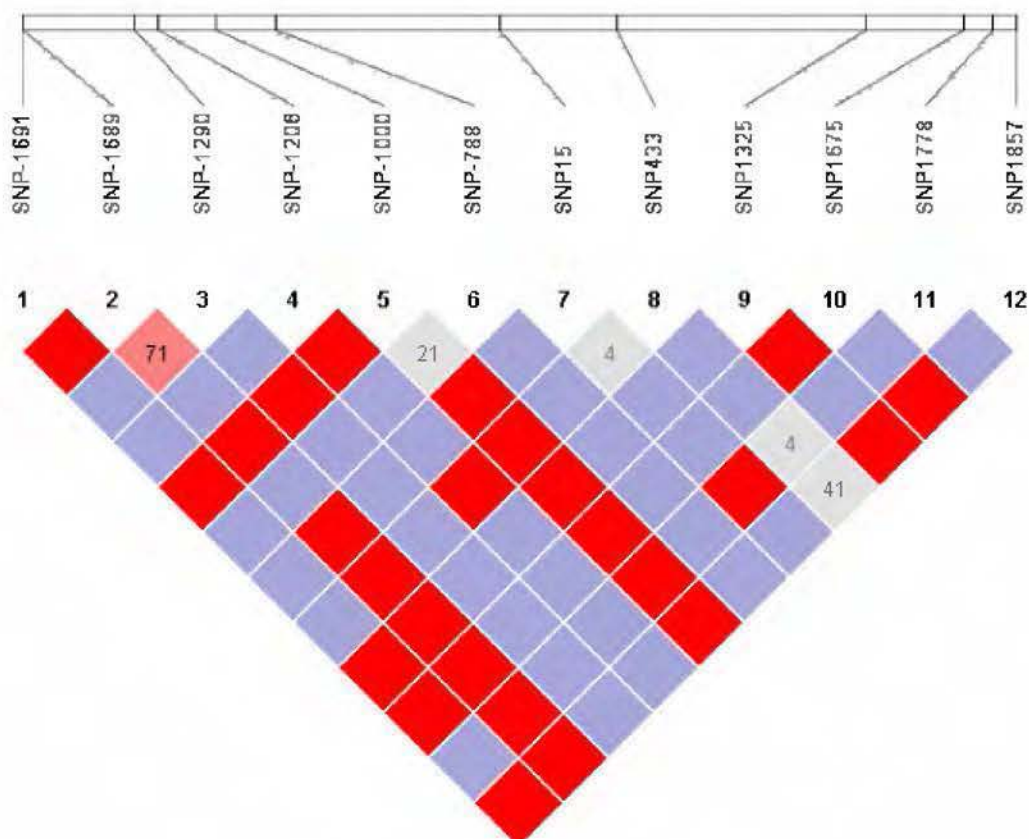
Los valores de D' y LOD para -1689 con +15 fueron 1 y 4.44 respectivamente, para -1691 con -1689, fue de 1 y 2.84. El gráfico de desequilibrio de ligamiento (Fig 3.) muestra el ligamiento entre los polimorfismos.

CUADRO 7

Haplotipos abarcando 12 polimorfismos en diferentes regiones del gen *LALBA*

Nº	Secuencia <sup>1</sup>	Frecuencia
1	GAGCTGGTGTGTTG	0.02
2	GAGCTGGTGTGTTA	<0.01
3	GAGCTGATGTGTCG	0.26
4	GAGCTTGTGTGTTG	0.01
5	GAGCCTGTGTGTTG	<0.01
6	GAGATGGTGTGTTG	0.05
7	GAGATGGCGTTG	0.08
8	GGGCCGGTGTGTTG	0.24
9	GGTCCGGTGTGTTG	0.17
10	CAGCTGGTACTA	0.15

1. Las posiciones resaltadas en azul, verde y amarillo corresponden a los polimorfismos -1691, -1689 y +15 respectivamente.



**Fig. 3.** Representación gráfica del análisis de ligamiento en los 12 SNP. Cada cuadro simboliza los valores de ligamiento entre dos SNP; los cuadros celestes indican un D' de 1 con un LOD menor a 2, los rojos representan un D' de 1 con LOD mayor o igual a 2 y los blancos un D' menor a 1 y LOD menor a 2. El número en los cuadros indica el porcentaje de D'.

### Ligamiento en muestras genotipificadas

Para las muestras tipificadas de las vacas, los valores de D' y LOD para -1689 y +15 fue de 0.992 y 131.72 respectivamente. Se obtuvo 4 haplotipos, pero uno de ellos con muy baja frecuencia (Cuadro 8).

CUADRO 8

Haplotipos de los SNPs -1689 y +15 en las muestras de ganado

Nº	Secuencia	Frecuencia
1	AA	0.23
2	AG	0.33
3	GG	0.44
4	GA	<<0.01

### Análisis estadístico

Las características de la leche se analizaron de dos maneras; incluyendo el efecto del SNP +15 en el modelo y otra excluyéndolo.

Cuando se excluye el SNP +15, la variable porcentaje de  $\alpha$ -lactalbúmina presenta una asociación significativa (Cuadro 9) con el SNP -1689, sin embargo, una vez que éste se incluye la significancia desaparece.

CUADRO 9

Valores de significancia para cada variable con respecto a  $\alpha$ -LA utilizando modelo sin SNP +15

Fuente de variación	g.l.	Denominador g.l.	F incremental	F condicional	p <sup>1</sup>
Intercepto	1	1598.5	2277.00	5.96	0.015 *
Días de lactancia (lactst)	1	1623.4	61.26	19.15	1.3x10 <sup>-5</sup> ***
Días de lactancia – exponencial (e <sup>lac</sup> )	1	1523.8	0.99	0.63	0.427
Edad de primer parto (ca)	1	1574.1	4.88	1.41	0.236
Edad de primer parto – cuadrático (ca <sup>2</sup> )	1	1562.7	1.74	1.79	0.182
Temporada de parto (season)	2	1572.6	2.3x10 <sup>-3</sup>	2.9x10 <sup>-3</sup>	0.997
Semental (scode)	1	286.3	0.30	0.36	0.549
SNP -1689	2	1525.8	6.02	6.02	0.002 **

1. Significancia: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

CUADRO 10

Valores de significancia para cada variable con respecto a  $\alpha$ -LA utilizando modelo con SNP +15

Fuente de variación	g.l.	Denominador g.l.	F incremental	F condicional	p <sup>1</sup>
Intercepto	1	1592.0	2392.00	6.29	0.012 *
Días de lactancia (lactst)	1	1621.3	64.73	19.21	1.2x10 <sup>-5</sup> ***
Días de lactancia – exponencial (e <sup>lac</sup> )	1	1515.9	1.06	0.61	0.434
Edad de primer parto (ca)	1	1567.1	5.15	1.38	0.241
Edad de primer parto – cuadrático (ca <sup>2</sup> )	1	1555.5	1.84	1.78	0.183
Temporada de parto (season)	2	1567.0	2.4x10 <sup>-3</sup>	0.01	0.991
Semental (scode)	1	289.1	0.31	0.57	0.450
SNP +15	2	1508.1	49.42	44.63	1.5x10 <sup>-19</sup> ***
SNP -1689	2	1516.1	1.57	1.57	0.209

1. Significancia: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘.’ 1

En los valores predichos se observa como el SNP +15 para el genotipo AA tiene una mayor concentración de  $\alpha$ -LA ( $p < 0.01$ ), mientras que para el SNP -1689 todos los genotipos poseen un promedio similar ( $p > 0.05$ ) (Fig. 4.).

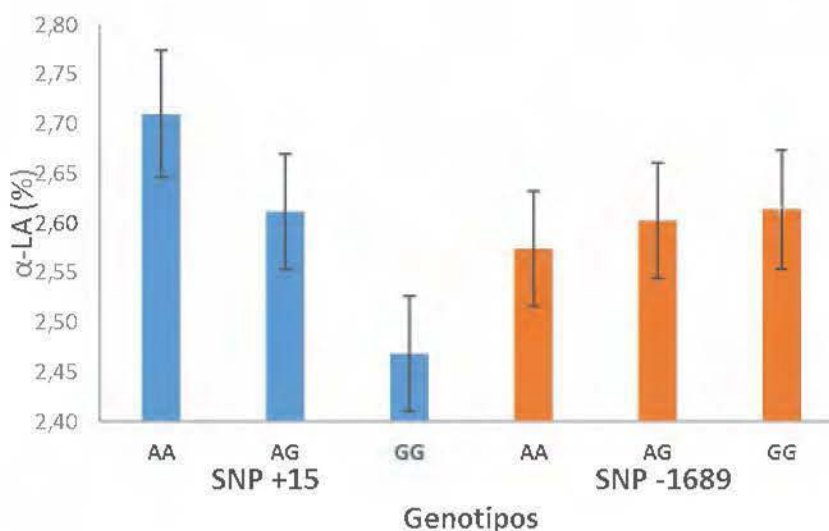
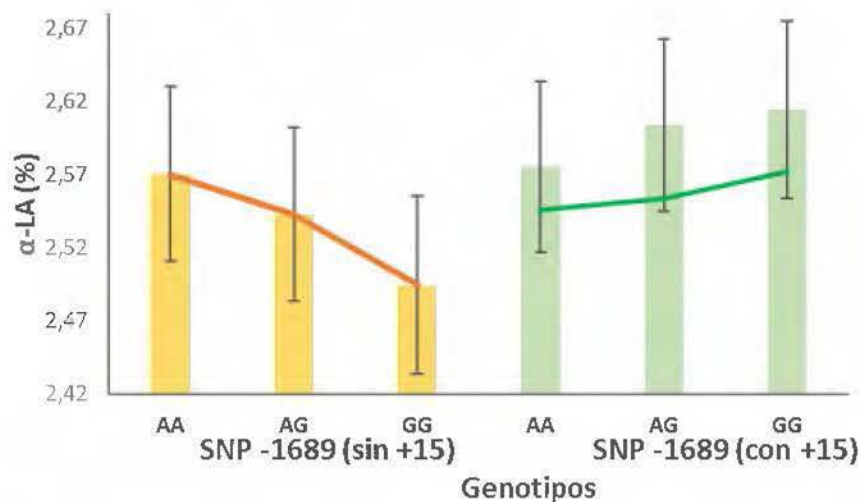


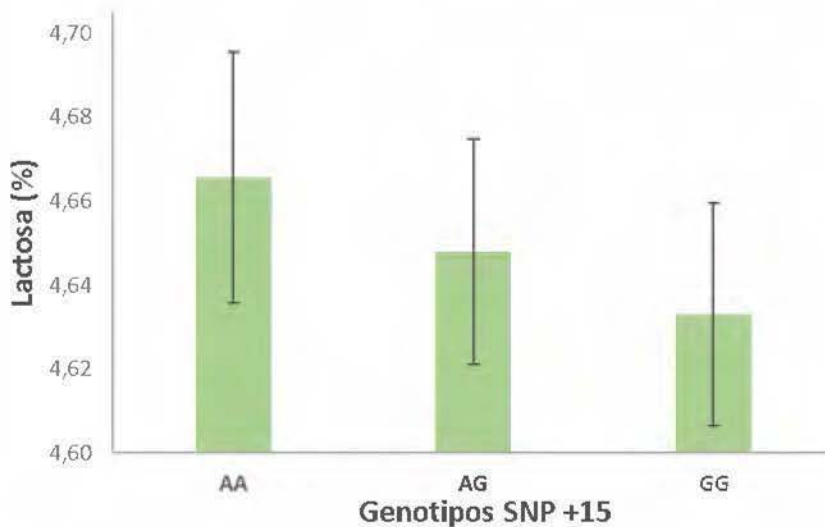
Fig. 4. Valores de concentración de  $\alpha$ -LA según el genotipo para el SNP +15 (barras azules) y SNP -1689 (barras anaranjadas), utilizando modelo que incluye efecto del SNP +15.

Los valores de  $\alpha$ -LA para el SNP -1689 presentan un cambio perceptible al analizar el modelo con y sin el efecto del SNP +15 (Fig. 5.)



**Fig. 5.** Comparación de los valores por genotipo de  $\alpha$ -LA para el SNP -1689 utilizando ambos modelos.

A pesar de que ninguno de los polimorfismos analizados estuvo significativamente asociado con el porcentaje de lactosa, el SNP +15 ( $p=0.073$ ) presenta tendencia a valores más altos con el genotipo AA (Fig 6.).



**Fig. 6.** Valores de concentración de lactosa según el genotipo para el SNP +15.

El porcentaje de caseínas totales presenta una asociación significativa con el SNP +15 (Cuadro 11), donde el genotipo GG presenta un promedio más alto (Fig. 7.).

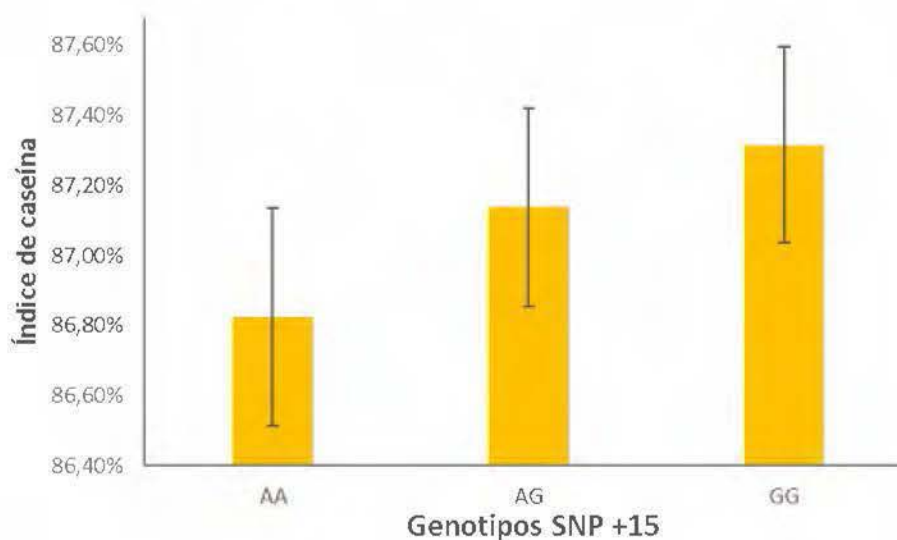


Fig. 7. Valores de índice de caseínas totales según el genotipo para el SNP +15.

#### CUADRO 11

Valores de significancia con respecto a índice de caseínas totales utilizando modelo con SNP +15

Fuente de variación	g.l.	Denominador g.l.	F incremental	F condicional	p <sup>1</sup>
Intercepto	1	1624.1	120000.00	477.90	4.5x10 <sup>-93</sup> ***
Días de lactancia (lactst)	1	1568.6	0.71	0.08	0.779
Días de lactancia – exponencial (e <sup>lac</sup> )	1	1580.7	0.43	0.17	0.678
Edad de primer parto (ca)	1	1608.8	1.7x10 <sup>-3</sup>	2.69	0.101
Edad de primer parto – cuadrático (ca <sup>2</sup> )	1	1601.8	2.69	2.73	0.098
Temporada de parto (season)	2	1587.4	0.79	0.86	0.424
Semental (scode)	1	310.5	2.56	2.35	0.126
SNP +15	2	1581.4	5.19	5.25	5.3x10 <sup>-3</sup> **
SNP -1689	2	1587.7	0.95	0.95	0.387

1. Significancia: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Ninguna de las demás variables analizadas presentó asociación significativa con los polimorfismos -1689 o +15 (Cuadro 12).

CUADRO 12

Valores de significancia con respecto a otras variables analizadas para los SNP +15 y -1689 y días de lactancia

Variable	Valor de p <sup>1</sup>		
	SNP +15	SNP -1689	Días de lactancia (lactst)
Porcentaje de grasa	0.503	0.817	1.3x10 <sup>-6</sup> ***
Porcentaje de proteínas totales	0.656	0.679	4.5x10 <sup>-7</sup> ***
Índice grasa / proteínas totales	0.169	0.789	0.005 **
Porcentaje de lactosa	0.073 .	0.342	3.6x10 <sup>-5</sup> ***
Producción de grasa	0.026 *	0.042 *	0.002 **
Producción de proteínas totales	0.556	0.168	5.7x10 <sup>-6</sup> ***
Producción de lactosa	0.374	0.087 .	6.3x10 <sup>-12</sup> ***
Concentración de urea	0.425	0.270	0.625
Producción de leche	0.535	0.107	8.5x10 <sup>-11</sup> ***
Porcentaje de β-Lactoglobulina	0.386	0.427	0.332
Porcentaje de caseína α1	0.950	0.694	0.002 **
Porcentaje de caseína α2	0.746	0.175	0.022 *
Porcentaje de caseína κ	0.527	0.409	9.1x10 <sup>-5</sup> ***
Porcentaje de caseína β	0.054 .	0.858	0.104
Porcentaje de caseínas totales	0.179	0.599	0.056 .
Producción de caseínas totales	0.655	0.147	1.9x10 <sup>-6</sup> ***

1. Significancia: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 '' 1

## Heredabilidad

Únicamente producción de grasa, porcentaje de  $\alpha$ -LA e índice de caseína estuvieron significativamente asociados con alguno de los SNP estudiados y por lo tanto la heredabilidad se calculó sólo para esas variables (Cuadro 13).

CUADRO 13

Valores de heredabilidad para las variables asociadas significativamente con los SNP analizados

Fenotipo	Heredabilidad	
	$h^2$	$h^2_{hBRD}$
Producción de grasa	0.26	0.35
Porcentaje de $\alpha$ -Lactalbúmina	0.54	0.67
Índice de caseína	0.62	0.68



## DISCUSIÓN

### **Estandarización de PCR y purificación**

Inicialmente, en la etapa de estandarización se presentó una banda inespecífica a la altura de los 900 pb. Esta debía eliminarse, de manera que no generara interferencia en la secuenciación y genotipificación al introducir secuencias ajenas a las de interés.

Dicha banda pudo generarse por un apareamiento de los imprimadores con algún pseudogen de *LALBA* o alguna lisozima, los cuales poseen una alta homología con el gen de  $\alpha$ -LA (Irwin *et al.* 2011), facilitado por la baja temperatura de alineamiento inicial (56 °C). Es por eso que se subió gradualmente la temperatura hasta que únicamente se observara el fragmento de 429 pb (Fig. 1.).

Las extracciones de las 1857 muestras de sangre de ganado Holstein-Friesian tuvieron que ser purificadas debido a que no amplificaban para ningún PCR. Inicialmente se pensó que el ADN podía estar degradado ya que dichas extracciones se realizaron desde el 2005, sin embargo, cuando se tomaba un número de estas muestras sin amplificar y se corrían en un gel de agarosa al 0.7%, se podía observar en todas ellas una banda pesada a la altura esperable del ADN genómico (no se muestran datos), por lo que la razón de que los PCR no funcionaran parecía ser inhibición.

Al estar almacenadas tanto tiempo pudo ocurrir que se desarrollaran microorganismos que eviten la acción de la polimerasa, que el pH cambiara a valores fuera de lo óptimo para el PCR u otro sinnúmero de factores (Wilson 1997, Kemp *et al.* 2006). Una vez que se purificaron, se probaron 15 muestras al azar y todas amplificaron el fragmento de 429 pb, por lo que se procedió a hacer la tipificación de todas las muestras.

## Secuenciación

Al secuenciar las muestras de los 72 toros se encontró el polimorfismo esperado en la posición -1689 (SNP\_ U63109\_263) y otro en la posición -1691 (SNP\_ U63109\_261).

Lundén & Lindersson (1998) reportaron el polimorfismo -1689 con una frecuencia alélica aproximada del 15% en 100 animales analizados, lo cual no coincide con los datos experimentales obtenidos. Por su parte, Voelker *et al.* (1997) y Martins *et al.* (2012) reportan una frecuencia aproximada del 35% en ganado Holstein, más similar a la frecuencia obtenida (Cuadro 5).

El polimorfismo -1691 ya ha sido reportado por Kaminski *et al.* (2005), pero no se han realizado estudios de asociación de este polimorfismo con características de la leche.

## Haplotipos y genotipificación

Para genotipificar el SNP -1689 se hizo uso de la enzima SduI, la cual reconoce la secuencia 5'-GDGCH↓C-3' siendo la posición de dicho polimorfismo la G resaltada. El ensayo de tipificación consistió en que las muestras con G en la posición -1689 eran digeridas y las que tenían la A no.

El SNP -1691 generó un inconveniente ya que este corresponde a la primera G de la secuencia de reconocimiento y existía el riesgo de que las muestras con la variante C en esa posición no pudieran ser genotipificadas por RFLP. Esto debido a que la enzima no iba a reconocer el sitio de corte en esos casos y por lo tanto la información que se infiriera no iba a ser precisa.

Sin embargo, al construir los haplotipos (Cuadro 7), se observó que la C del polimorfismo -1691 sólo ocurría en la variante A del SNP -1689, precisamente en la que no tiene que ocurrir digestión, por lo que no se invalidaba la genotipificación por RFLP. Asimismo, el análisis de ligamiento estimó un D' de 1 y un LOD de 2.84 (Fig. 3), lo cual evidencia un claro desequilibrio de ligamiento. El hecho de que apenas haya un nucleótido de distancia entre ambos polimorfismos hace que la probabilidad de recombinación entre ellos sea muy baja (Khatkar *et al.* 2007)

Esto se verificó comparando la información de secuenciación y genotipificación de las muestras de toros, en la cual hubo un 100% de coincidencia. También se secuenciaron 95 muestras de las 302 vacas Holstein-Friesian que habían dado un genotipo AA mediante el ensayo de tipificación y se obtuvo los mismos resultados con ambas técnicas.

Al construir los haplotipos en las 1857 muestras con los SNP -1689 y +15, se observa que la frecuencia de la combinación GA es de alrededor del 0.1% (Cuadro 8), los valores de  $D'$  y LOD para -1689 y +15 fue de 0.992 y 131.72 respectivamente, lo cual sugiere que no han ocurrido eventos de recombinación entre ambos. El polimorfismo +15A siempre se va a encontrar en el polimorfismo -1689A.

### **Análisis de asociación**

Para determinar la posible asociación de polimorfismos con características de la leche, se utilizó un modelo mixto descrito por Schopen *et al.* (2011) que contempla una serie de variables importantes con efectos distintos que pueden influenciar la composición de la leche.

Dentro de las variables se incluye la edad al primer parto, ya que esta influye en la producción de leche y concentración de grasa y proteína. Sin embargo, sólo hubo una leve significancia ( $p = 0.02$ ) con el porcentaje de proteínas. Se ha observado que a una edad de 20 a 22 meses, aumenta el porcentaje de proteínas totales, pero disminuye el porcentaje de grasa y la producción de leche (Pirlo *et al.* 2000, Mohd Nor *et al.* 2013). La mayoría de animales estudiados tenían una edad mayor a los 24 meses y quizá al haber pasado ese periodo crítico, los valores no variaron mucho con respecto a esta variable.

La estación es otra variable importante, ya que tiene un efecto muy marcado sobre los porcentajes de grasa y proteína. Durante los meses más calientes (Junio-Agosto en el hemisferio norte) estos porcentajes disminuyen. Conforme va avanzando el otoño (Setiembre-Noviembre) vuelven a subir hasta alcanzar el máximo durante el invierno (Diciembre-Febrero) y comienzan a bajar nuevamente durante la primavera (Marzo-Mayo) (Heck *et al.* 2009a). Para los datos

analizados tampoco se observó diferencias significativas en esta variable, lo cual concuerda con lo reportado por Stoop *et al.* 2009. Esto puede deberse a que la variación observada entre las muestras de diferentes estaciones corresponda más a cambios en el manejo y la dieta, los cuales van contemplados en la variable hato (herd).

Los días de lactancia son determinantes ya que la mayor cantidad de grasa y proteína en la leche se produce en el calostro justo después del parto. Además, el punto más alto de producción de leche se alcanza entre los 25 y 50 días luego del parto, la cual comienza a descender con el paso de los meses (Stoop *et al.* 2009), esto se confirmó en los datos mediante la correlación negativa en características de proteínas y grasa. Esta variable fue significativa para casi todos los rasgos analizados, salvo índice de caseína,  $\beta$ -caseína, caseínas totales,  $\beta$ -Lactoglobulina y urea.

La variable semental se incluye en el modelo para medir cualquier variación que pudiera generarse de usar sementales jóvenes y probados. Los jóvenes al no tener aún hijas demostradas con buen rendimiento podrían generar valores significativamente menores (De Roos *et al.* 2011). En los datos no se observó significancia para esta variable.

El efecto genético aditivo del animal, obtenido mediante la genealogía, permite calcular cuánto de la variación que se observa en el fenotipo se debe a la variación en el genotipo (Heck *et al.* 2009b, Bilal *et al.* 2012). En el caso del efecto hato, permite estimar la variación introducida por las diferentes fincas, ya sea manejo, alimentación y cualquier otro elemento que pudiera afectar (Heck *et al.* 2012). Teniendo estas variables, se puede estimar la heredabilidad y la heredabilidad intrahato.

Agregando el efecto fijo del polimorfismo en el modelo se puede establecer si hay alguna relación entre este y la variable dependiente.

#### *$\alpha$ -Lactalbúmina*

Al analizar el SNP +15 se encontró una asociación significativa a este componente tal y como ha sido reportado previamente (Schopen *et al.* 2011, Visker *et al.* 2011), con el genotipo AA dando

un porcentaje más alto. Sin embargo, el SNP -1689 no presentó asociación alguna (Fig. 4.), esto coincide con lo reportado por Voelker *et al.* (1998).

Inicialmente se probó el modelo descrito sin incluir el efecto del SNP +15, dando una falsa asociación del polimorfismo -1689 con la concentración de  $\alpha$ -LA (Cuadro 9).

El polimorfismo +15A siempre se encuentra asociado al polimorfismo -1689A, haciendo que el efecto del SNP +15 que eleva el promedio de  $\alpha$ -LA, esté contenido dentro de la variante A del SNP -1689.

De esta manera, si no se toma en cuenta el efecto del SNP +15 a la hora de analizar el polimorfismo -1689, este va a tener un efecto “diluido” de asociación a  $\alpha$ -LA, por el desequilibrio de ligamiento con +15. En la Fig. 4 se observa el efecto del SNP +15, donde el genotipo AA posee el promedio más alto y el GG el más bajo, así mismo, en la Fig. 4 el SNP -1689 sin tomar en cuenta el SNP +15, parece tener la misma tendencia; los genotipos AA y AG poseen el promedio más alto mientras que GG tiene el más bajo.

El efecto del SNP +15 sobre la concentración de  $\alpha$ -LA podría estar a nivel de traducción. Dicho nucleótido se encuentra en la región 5'UTR que logra transcribirse en ARNm. A esta región se le atribuyen diferentes funciones como transporte y eficiencia de traducción (Kim 2013). Este polimorfismo podría ser clave en la unión del ARN con los ribosomas, pero para determinar el efecto específico necesitan realizarse estudios a nivel bioquímico donde se establezca su afinidad con la variación en la secuencia.

### *Caseínas totales*

Al igual que lo reportado por Schopen *et al.* (2011) y Visker *et al.* (2011) el SNP +15 mostró una asociación significativa con el porcentaje de caseínas totales (Cuadro 11), pero al contrario que con  $\alpha$ -LA, el genotipo GG presenta el promedio más alto, mientras que AA el más bajo.

Esto es congruente con lo esperable, ya que existe una correlación inversa entre la concentración de lactosa y la de cantidad de caseína en la leche. Si un genotipo aumenta la

concentración de  $\alpha$ -LA, va a haber una mayor síntesis de lactosa, ya que es el componente enzimático que junto con la UDP-galactosil transferasa forma la lactosa sintetasa. Este carbohidrato lleva agua hacia las vesículas del aparato de Golgi (porque allí es donde se sintetiza) y afecta el volumen de la leche y la concentración de caseína, la cual es empacada en estas vesículas (Bleck *et al.* 2009, Kuhn & White 2009).

A pesar de que no hubo una asociación significativa entre los polimorfismos analizados y la concentración de lactosa, en la Fig. 6 se observa como para el SNP +15 el genotipo AA genera un promedio más alto y GG uno más bajo, al igual que con  $\alpha$ -LA. De acuerdo a lo explicado en el párrafo anterior esto es esperable ya que a mayor concentración de  $\alpha$ -LA, mayor síntesis de lactosa.

### **Heredabilidad**

El porcentaje de  $\alpha$ -LA y el Índice de caseína poseen una heredabilidad intrahato bastante alta (Cuadro 13) lo que coincide con lo reportado por Schopen *et al.* (2009). Esto sugiere que sí es factible modificar estos parámetros de la leche mediante selección.

En el caso de producción de grasa, no fue un valor tan alto, pero podría tomarse en cuenta si se hiciera selección con estos marcadores.

## CONCLUSIONES

- El polimorfismo -1689 sí está presente en la población estudiada de ganado Holstein-Friesian de países bajos, pero no está asociado a la composición proteínica de la leche.
- Se corroboró que el polimorfismo +15 sí está asociado a la concentración de  $\alpha$ -LA y caseínas totales, tal y como se había reportado previamente, por lo que dicho marcador puede incluirse dentro de los polimorfismos a analizar dentro de los programas de mejoramiento asistido.
- Los SNP +15 y -1689 están en desequilibrio de ligamiento lo que genera una falsa asociación del -1689 a  $\alpha$ -LA cuando no se toma en cuenta el efecto del +15.
- El desequilibrio de ligamiento que existe entre los polimorfismos -1691 y -1689 permitió la genotipificación del SNP -1689 con la enzima SduI.

## RECOMENDACIONES

- Debe analizarse el polimorfismo encontrado en la posición -1691 y determinar si presenta asociación con alguna característica de la leche.
- Para determinar el mecanismo mediante el cual el SNP+15 afecta la concentración de  $\alpha$ -LA, necesitan realizarse estudios a nivel bioquímico donde se establezca su efecto sobre la traducción de dicha proteína.



## BIBLIOGRAFÍA

- Benton, M.J. 1999. Early origins of modern birds and mammals: molecules vs morphology. *BioEssays*. 21:1043-51.
- Berman, A. 2011. Invited review: are adaptations present to support dairy cattle productivity in warm climates? *J. Dairy Sci.* 94: 2147-2158.
- Bilal, G., R.I. Cue, A.F. Mustafa & J.F. Hayes. 2012. Short communication: Estimates of heritabilities and genetic correlations among milk fatty acid unsaturation indices in Canadian Holsteins. *J. Dairy Sci.* 95:7367-7371.
- Bleck, G.T. & R.D. Bremel. 1993a. Sequence and single-base polymorphisms of the bovine  $\alpha$ -lactalbumin 5'-flanking region. *Gene*. 126:213-218.
- Bleck, G.T. & R.D. Bremel. 1993b. Correlation of the  $\alpha$ -Lactalbumin (+15) polymorphism to milk production and milk composition of Holsteins. *J. Dairy Sci.* 76:2292-2298.
- Bleck, G.T., M.B. Wheeler, L.B. Hansen, H. Chester-Jones & D.J. Miller. 2009. Lactose synthase components in milk: concentrations of  $\alpha$ -lactalbumin and  $\beta$ 1,4-galactosyltransferase in milk of cows from several breeds at various stages of lactation. *Repro. Dom. Anim.* 44:241-247.
- Boston, W.S., G.T. Bleck, J.C. Conroy, M.B. Wheeler & D.J. Miller. 2001. Short communication: effects of increased expression of  $\alpha$ -lactalbumin in transgenic mice on milk yield and pup growth. *J. Dairy Sci.* 84:620-622.
- Brew, K. 2003.  $\alpha$ -Lactalbumin. p. 387-419. *In* P.F. Fox & P.L.H. McSweeney (eds.). *Advanced Dairy Chemistry: Proteins*. Kluwer Academic/Plenum Publishers, New York, EEUU.
- Butler, D., B.R. Cullis, A.R. Gilmour & B.J. Gogel. 2006. *ASReml-R reference manual*. NSW Department of Primary Industries and Queensland Government, Department of Primary Industries and Fisheries.
- Barrett, J.C., B. Fry, J. Maller & M.J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 21:263-265.

- Baumrucker, C.R. 2008. Hormones and biologically active compounds in milk. *In* P.L.H. McSweeney & P.F. Fox (eds.) *Advanced dairy chemistry: Lactose, water, salts and minor constituents*. Springer, New York, New York, EE.UU.
- De Marchi, M., G. Bittante, R. Dal Zotto, C. Dalvit & M. Cassandro. 2008. Effect of Holstein-Friesian and Brown Swiss breeds on quality of milk and cheese. *J. Dairy Sci.* 91:4092-4102.
- De Roos, A.P.W., C. Schrooten, R.F. Veerkamp & J.A.M. van Arendonk. 2011. Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls. *J. Dairy Sci.* 94:1559-1567.
- Evers, J. 2004. The milkfat globule membrane - compositional and structural changes post secretion by the mammary secretory cell. *Int. Dairy J.* 14:661-674.
- FAO Statistics Division. 2010. Cow milk (whole, fresh) production (tonnes). (Consultado: 27 mayo 2012, <http://faostat.fao.org/site/569/default.aspx#ancor>)
- Fox, P.F. & A. Brodtkorb. 2008. The casein micelle: historical aspects, current concepts and significance. *Int. Dairy J.* 18:677-684.
- Fox, P.F. 2009. Milk: an overview, p. 1-54. *In* A. Thompson, M. Boland & H. Singh (eds.). *Milk Proteins: from Expression to Food*. Academic Press, San Diego, California, EEUU.
- Goldman, A.S. 2002. Evolution of mammary gland defense system and the ontogeny of the immune system. *J. Mammary Gland Biol. Neoplasia.* 7:277-289.
- Grimble, R.F. & Y.K. Mansaray. 1987. Effects in Rats of Dietary Protein Inadequacy on Lactose Production, Milk Volume and Components of the Lactose Synthetase Complex (EC 2.4.1.22). *Ann. Nutr. Metab.* 31:179-184.
- Gillespie, J.R. & F.B. Flanders. 2009. *Modern Livestock and Poultry Production*. Delmar Cengage Learning, New York, New York, EE.UU.
- Ha, E. & M.B. Zemel. 2003. Functional properties of whey, whey components, and essential amino acids: mechanisms underlying health benefits for active people. *J. Nutr Biochem.* 14:251-258.

- Heck, J.M., H.J. van Valenberg, J. Dijkstra & A.C. van Hooijdonk. 2009a. Seasonal variation in the Dutch bovine raw milk composition. *J. Dairy. Sci.* 92:4745-4755.
- Heck, J.M., A. Schennink, H.J.F. van Valenberg, H. Bovenhuis, M.H.P.W. Visker, J.A.M. van Arendonk & A.C.M. van Hooijdonk. 2009b. Effects of milk protein variants on the protein composition of bovine milk. *J. Dairy Sci.* 92:1192-1202.
- Heck, J.M.L., H.J.F. van Valenberg, H. Bovenhuis, J. Dijkstra & A.C.M. van Hooijdonk. 2012. Characterization of milk fatty acids based on genetic and herd parameters. *J. Dairy Res.* 79:39-46.
- Ingram, C.J.E. & D. Swallow. 2007. Lactose malabsorption. *In* P.L.H. McSweeney & P.F. Fox (eds.) *Advanced dairy chemistry: Lactose, water, salts and minor constituents*. Springer, New York, New York, EE.UU.
- Irwin, D.M., J.M. Biegel & C.B. Stewart. 2011. Evolution of the mammalian lysozyme gene family. *BMC. Evol. Biol.* 11:166.
- Jensen, R.G. 1995. General Description of Milks, p. 2-3. *In* R.G. Jensen (ed.). *Handbook of Milk Composition*. Academic Press, San Diego, California, EE.UU.
- Johnson J.D, R.O. Christiansen & N. Kretchmer. 1972. Lactose synthetase in mammary gland of the California sea lion. *Biochem. Biophys. Res. Commun.* 47:393-397.
- Kaminski, S., A. Ahman, A. Rusc, E. Wojcik & T. Malewski. 2005. MilkProtChip - A microarray of SNPs in candidate genes associated with milk protein biosynthesis: Development and validation. *J. Appl. Genet.* 46:45-58.
- Kemp, B.M., C. Monroe & D.G. Smith. 2006. Repeat silica extraction: a simple technique for the removal of PCR inhibitors from DNA extracts. *J. Archaeol. Sci.* 33:1680-1689.
- Khatkar, M.S., K.R. Zenger, M. Hobbs, R.J. Hawken, J.A. Cavanagh, W. Barris, A.E. McClintock, S. McClintock, P.C. Thomson, B. Tier, F.W. Nicholas & H.W. Raadsma. 2007. A primary assembly of a bovine haplotype block map based on a 15,036-single-nucleotide polymorphism panel genotyped in Holstein-Friesian cattle. *Genetics.* 176:763-772.

- Kim, J.J. 2013. Translational Regulation of Bovine Casein. Tesis de Ph.D., Universidad de Guelph, Ontario, Canada.
- Kuhn, N.J. & A. White. 2009. The Topography of Lactose Synthesis. *J. Mammary Gland Biol. Neoplasia*. 14:213-220.
- Lamy, E., S. van Harten, E. Sales-Baptista, M.M. Mendes-Guerra & A.M. de Almeida. 2012. Factors Influencing Livestock Productivity. P. 19-51. *In V. Sejian, S.M.K. Naqvi, T. Ezeji, J. Lakritz & Rattan Laln (eds.). Environmental Stress and Amelioration in Livestock Production. Springer-Verlag, New York, EEUU.*
- Langer, P. 2009. Differences in the composition of colostrum and milk in eutherians reflect differences in immunoglobulin transfer. *J. Mammology*. 90:332-339.
- Lundén, A. & M. Lindersson. 1998.  $\alpha$ -Lactalbumin polymorphism in relation to milk lactose. *Proc. 6th World Congr. Genet. Appl. Livest. Prod.* 25:47-50.
- Martin, P., M. Szymanowska, L. Zwierzchowski & C. Leroux. 2002. The impact of genetic polymorphisms on the protein composition of ruminant milks. *Reprod. Nutr. Dev.* 42:433-459.
- Martins, L.F., M.P. Milazzotto, W.B. Feitosa, A.R.S. Coutinho, R. Simões, M.G. Marques, M.E. Assumpção & J.A. Visintin. 2008. Sequence variation of the  $\alpha$ -lactalbumin gene in Holstein and Nellore cows. *Anim. Biotechnol.* 19:196-199.
- McKenzie, H.A. & F.H. White. 1991. Lysozyme and  $\alpha$ -lactalbumin: structure, function, and interrelationships. *Advances Prot. Chem.* 41:173-315.
- Mohd Nor, N., W. Steeneveld, T. van Werven, M.C.M. Mourits & H. Hogeveen. 2013. First-calving age and first-lactation milk production on Dutch dairy farms. *J. Dairy Sci.* 96:981-992.
- Natale, M., C. Bisson, G. Monti, A. Peltran, L.P. Garoffo, S. Valentini, C. Fabris, E. Bertino, A. Coscia, A. Conti. 2004. Cow's milk allergens identification by two-dimensional immunoblotting and mass spectrometry. *Mol. Nutr. Food. Res.* 48:363-369.

- Pilson, M.E.Q. & A.L. Kelly. 1962. Composition of the milk from *Zalophus californianus*, the California sea lion. *Science*. 135:104-105.
- Pirlo, G., F. Miglior & M. Speroni. 2000. Effect of age at first calving on production traits and on difference between milk yield returns and rearing costs in Italian Holsteins. *J. Dairy Sci.* 83:603-608.
- Prendiville, R., K.M. Pierce & F. Buckley. 2010. A comparison between Holstein-Friesian and Jersey dairy cows and their F1 cross with regard to milk yield, somatic cell score, mastitis, and milking characteristics under grazing conditions. *J. Dairy Sci.* 93:2741-2750.
- R Development Core Team. 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. (Consultado: 15 abril 2012, [www.r-project.org](http://www.r-project.org))
- Schopen, G.C.B., J.M.L. Heck, H. Bovenhuis, M.H.P.W. Visker, H.J.F. van Valenberg & J.A.M. van Arendonk. 2009. Genetic parameters for major milk proteins in Dutch Holstein-Friesians. *J. Dairy Sci.* 92:1182-1191.
- Schopen, G.C.B., M.H.P.W. Visker, P.D. Koks, E. Mullaart, J.A.M. van Arendonk & H. Bovenhuis. 2011. Whole-genome association study for milk protein composition in dairy cattle. *J. Dairy Sci.* 94:3148-3158.
- Selo, I., G. Clement, H. Bernard, J. Chatel, C. Creminon, G. Peltre, J. Wal. 1999. Allergy to bovine  $\beta$ -lactoglobulin: specificity of human IgE to tryptic peptides. *Clin. Exp. Allergy* 29:1055-1063.
- Sheehy, P.A., P. Williamson, J.A. Sharp, K. Menzies, C. Lefevre, M. Digby, K.R. Nicholas & P.C. Wynn. 2009. Significance, origin and function of bovine milk proteins: the biological implications of manipulation or modification. p. 81-106. *In* A. Thompson, M. Boland & H. Singh (eds.). *Milk Proteins: from Expression to Food*. Academic Press, San Diego, California, EEUU.
- Solounias, N., J.C. Barry, R.L. Bernor, E.H. Lindsay & S. Mahmood Raza. 1995. The oldest bovid from the Siwaliks, Pakistan. *J. Vert. Paleont.* 15:806-814.

- Stephens, M. & P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *Am. J. Hum. Genet.* 76:449-462.
- Stoop, W.M., H. Bovenhuis & J.A.M. van Arendonk. 2007. Genetic parameters for milk urea nitrogen in relation to milk production traits. *J. Dairy Sci.* 90:1981-1986.
- Stoop, W.M., H. Bovenhuis, J.M.L. Heck & J.A.M. van Arendonk. 2009. Effect of lactation stage and energy status on milk fat composition of Holstein-Friesian cows. *J. Dairy Sci.* 92:1469-1478.
- Storey, J. D. & R. Tibshirani. 2003. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA.* 100:9440-9445.
- Van Knegsel, A.T.M., H. van den Brand, J. Dijkstra, W.M. van Straalen, M.J. Heetkamp, S. Tamminga & B. Kemp. 2007. Dietary energy source in dairy cows in early lactation: Energy partitioning and milk composition. *J. Dairy Sci.* 90:1467-1476.
- Vilotte, J., S. Soulier, J. Mercier, P. Gaye, D. Hue-Delahaie & J. Furet. 1987. Complete nucleotide sequence of bovine  $\alpha$ -lactalbumin gene: comparison with its rat counterpart. *Biochimie.* 69:609-620.
- Visker, M.H.P.W., A. Schennink, J.A.M. van Arendonk & H. Bovenhuis. 2011. Effects of the  $\alpha$ -lactalbumin +15 polymorphism on milk protein composition. EAAP, Stavanger. Agosto 2011.
- Visker, M.H.P.W., J.M.L. Heck, H.J.F. van Valenberg, J.A.M. van Arendonk & H. Bovenhuis. 2012. Short communication: A new bovine milk-protein variant:  $\alpha$ -lactalbumin variant D. *J. Dairy Sci.* 95:2165-2169.
- Voelker, G.R., G.T. Bleck & M.B. Wheeler. 1997. Single-base polymorphisms within the 5' flanking region of the bovine  $\alpha$ -lactalbumin gene. *J. Dairy Sci.* 80:194-197.
- Voelker, G.R., G.T. Bleck & M.B. Wheeler. 1998. Identification of variations in the  $\alpha$ -Lactalbumin gene in cattle and potential correlations to milk production traits. (Consultado: 27 febrero 2013, <http://www.livestocktrail.illinois.edu/dairy/paperDisplay.cfm?ContentID=281>)

Wilson, I.G. 1997. Inhibition and facilitation of nucleic acid amplification. *Appl. Environ. Microbiol.* 63:3741-3751.

Zenger, K.R., M.S. Khatkar, J.A. Cavanagh, R.J. Hawken & H.W. Raadsma. 2007. Genome-wide genetic diversity of Holstein Friesian cattle reveals new insights into Australian and global population variability, including impact of selection. *Anim. Genet.* 38:7-14.

## APÉNDICES

**Apéndice 1**

## Purificación con placa de 96 pozos MultiScreen HTS PCR

## Protocolo

1. Cargar la reacción de PCR en la placa MultiScreen HTS (usualmente entre 100 y 300  $\mu$ l de muestra por pozo). No necesariamente tiene que usar todos los 96 pocillos de una vez. Para una recuperación óptima de productos de PCR pequeños (<300 pb) diluir las reacciones a 300  $\mu$ l utilizando agua de grado biología molecular o solución amortiguadora TE.
2. Coloque la placa sobre el colector de vacío MultiScreen™.
3. Aplique vacío a 25 pulgadas de mercurio durante 5 a 12 minutos o hasta que los pozos estén secos. Volúmenes mayores a 100  $\mu$ l y presiones menores a 25, incrementarán los tiempos de filtración. Para una recuperación óptima de fragmentos pequeños (<300 pb) aplique vacío a una presión de 10 pulgadas de mercurio. El efecto combinado de la dilución y presión baja de vacío aumentan el tiempo de filtración.
4. Deje otros 30 segundos la placa en el vacío hasta que los pocillos se vean vacíos y todo el líquido se haya filtrado. Los filtros se ven brillantes aún después de que están secos.
5. Lave con 35  $\mu$ l de agua grado biología molecular para obtener una mayor pureza repitiendo los pasos 2 al 4.
6. Después de que la filtración por vacío está completa, remueva la placa del colector de vacío y seque la base de la placa con papel toalla y agregue un volumen de agua o TE igual al volumen original de la reacción de PCR con una micropipeta. Mezcle la muestra vigorosamente en un agitador de placas por 5 minutos. Para una recuperación óptima de productos pequeños de PCR, dejar preincubar la solución de 5 a 10 minutos en los pocillos antes de agitar y/o extienda el tiempo de agitación a 10 minutos.
7. Recoja el producto de PCR purificado de cada pozo con una micropipeta y cóloquela en un tubo o placa.



## Apéndice 2

### Purificación de reacción de secuenciación con NaAc/EDTA/Etanol

#### Protocolo

1. Agregar 1  $\mu$ l de EDTA 125 mM a cada pozo. Asegurarse que se disuelva en las muestras.
2. Agregar 1  $\mu$ l de NaAc 3 M a cada pocillo y asegurarse que se disuelva, mezclar con pipeteo. El acetato de sodio es muy volátil, no dejar abierto en el área de trabajo por periodos prolongados.
3. Agregar a cada pozo 38  $\mu$ l de etanol absoluto, enfriado previamente a -20 °C.
4. Reposar por 30 minutos sobre hielo.
5. Centrifugar durante 30 min en 3000 g a 4 °C.
6. Verter el contenido de la placa sobre un pedazo de papel toalla en un movimiento rápido y golpear suavemente en el papel toalla.
7. Centrifugar la placa boca abajo sobre un nuevo papel toalla por 1 min a 700 g.
8. Resuspender la muestra en 10  $\mu$ l de formamida.

**Apéndice 3****Cuantificación de 15 muestras purificadas con kitOneStep™ PCR Inhibitor Removal**

Muestra	ng/μl	260/280	260/230
1-1	139.2	1.9	1.81
1-2	122.64	1.88	1.76
1-3	122.38	1.88	1.75
1-4	116.97	1.91	1.63
1-5	140.18	1.92	1.85
1-6	135.15	1.93	1.79
1-7	120.69	1.93	1.7
1-8	142.54	1.89	1.72
1-9	120.04	1.93	1.53
1-10	138.38	1.88	1.78
2-1	40.61	1.94	1.08
2-2	138.68	1.93	1.78
2-3 E7	133.87	1.92	1.82
2-4 G9	139.19	1.89	1.8
2-5 A10	214.46	1.89	1.8