

¿Qué cambiará en la selección del vacuno de leche Holstein en los próximos años?

En los últimos años se ha dado una rápida evolución de las técnicas de secuenciación del genoma y en 2008 ha comenzado a generalizarse su aplicación para la selección en el vacuno de leche Holstein, la era de la selección genómica ya ha comenzado.

¿EN QUE CONSISTE LA SECUENCIACIÓN DEL GENOMA?

El ADN está compuesto por dos cadenas paralelas de bases químicas, llamadas nucleótidos. Estas pueden ser de cuatro tipos: Adenina (A), Guanina (G), Citosina (C) y Timina (T). Cuando en el año 2000 se secuenció el genoma humano, lo que se había conseguido era conocer el orden exacto en el que aparecían estos cuatro nucleótidos en los cromosomas de unos pocos individuos (ver figura 1). Eso significa que se identificaron 3 billones de pares de bases presentes en cada uno de estos genomas. La importancia de conocer la secuencia de estas cuatro letras a lo largo del ADN es debida a que esta secuencia determina las proteínas que se sintetizan en las células y éstas a su vez intervienen en todos los procesos biológicos que se dan en los seres vivos y determinan diferencias observables entre los individuos.

¿QUÉ SON LOS SNP?

La trascendencia de haber secuenciado el genoma de varios individuos es debida a que la mayor parte del genoma es igual en todas las personas. Las variaciones son debidas a los cambios que se han ido dando en el ADN, pero representan una pequeña parte del total de bases del genoma. Un SNP ("polimorfismo de un solo nucleótido") es una **variación en la secuencia de ADN que afecta a una sola base** (ver figura 2). Se les denomina "snips" y son el tipo de variación más frecuente en el ADN, constituyen hasta el 90% de todas las variaciones genómicas humanas. Cualquier base del ADN podría ser sustituida por cualquier otra, pero, en la práctica, la mayoría de los SNP solo tienen dos variantes o alelos. Un aspecto importante es que, para que una variación en una posición se considere que es un SNP, la variante menos frecuente debe estar presente al menos en el 1% de la población. Los SNP presentan una distribución aproximadamente uniforme en el genoma y pueden estar o no asociados con diferencias observables entre los individuos. En general el efecto asociado a cada SNP será pequeño.

La gran cantidad de dinero público que se ha destinado a la secuenciación del genoma y al estudio de los SNP es debido a su utilidad para la investigación médica en torno a los riesgos de desarrollar determinadas enfermedades o sobre la eficiencia de los fármacos y la tolerancia a éstos.

Juan Pena

Departamento Técnico de CONAFE

Figura 1. Estructura del ADN (National Cancer Institute, 2002)

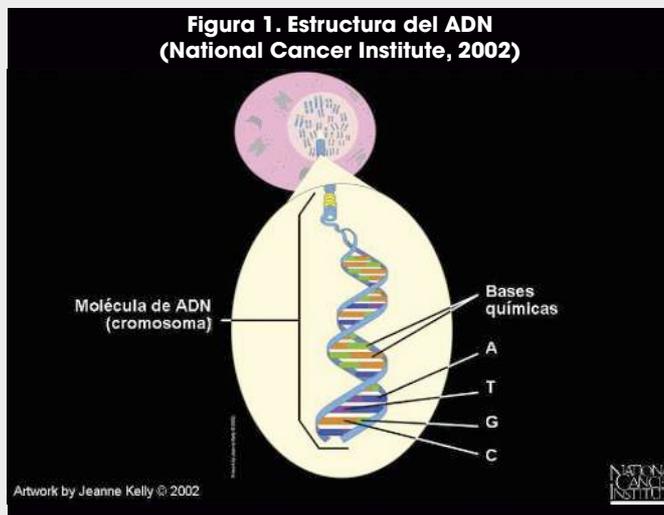
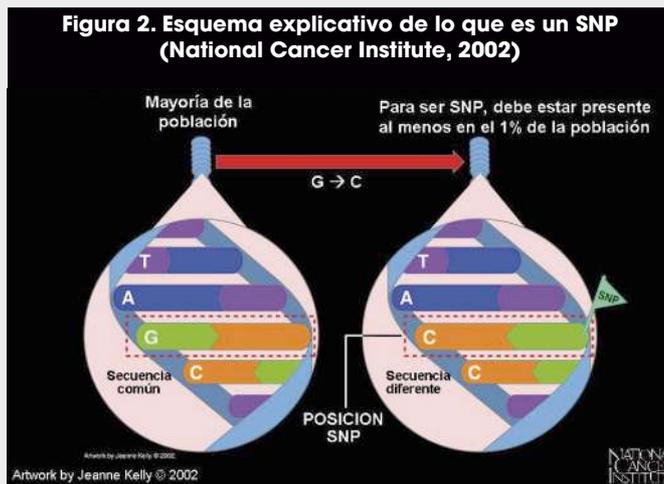


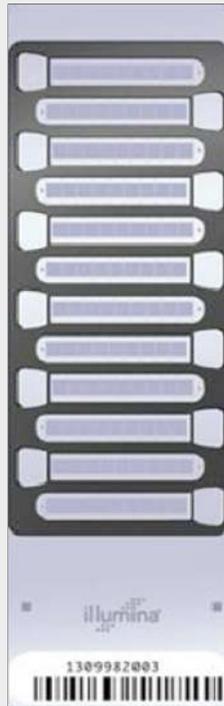
Figura 2. Esquema explicativo de lo que es un SNP (National Cancer Institute, 2002)



CHIP 50K SNP

A continuación del genoma humano, en 2004 se secuenció el genoma bovino, financiado también en su mayor parte por el gobierno de EEUU y cuyos resultados eran y son de acceso libre. A partir de esta información se han desarrollado aplicaciones comerciales para conocer a un precio asequible la información exacta que presenta cada toro o vaca en un elevado número de SNP. En 2005 era ya posible a partir de una muestra de sangre, semen o pelo conocer por un precio asequible (menos de 400 dólares) la información existente en 10.000 SNP. Pero, desde diciembre de 2007 ya se pueden cono-

Figura 3. Foto del chip de Illumina utilizado para la secuenciación de 58.000 SNP



cer las "letras" que presenta cada individuo en más de 50.000 SNP, por unos 250 dólares. Y es con esta última aplicación comercial con la que se ha producido el disparo de salida para esta carrera de la selección genómica (ver foto en figura 3). El precio actual en Europa está alrededor de los 200 euros.

Cuando hablamos del genotipo de un toro en este contexto de selección genómica, nos referimos a la información que presenta el toro en los 58.000 SNP integrados en el chip comercial, aunque en realidad el término genotipo engloba el total del ADN del animal.

¿QUÉ EFECTO TIENEN LOS SNP SOBRE LOS FENOTIPOS?

Los SNP no tienen porque formar parte de los genes. Pero, si están bien distribuidos a lo largo del genoma, muchos de ellos estarán próximos a zonas del ADN responsables de caracteres de interés, es decir, estarán asociados a genes. El objetivo será identificar estas asociaciones entre los SNP y las distintas características que interesa estudiar en los individuos. Por eso se dice que los SNP son "marcadores genéticos". En la figura 4 se representa un ejemplo del efecto de un SNP en el mérito genético de kg de leche. Podemos considerar que el "efecto" de este SNP es debido a que se transmite conjuntamente con un determinado gen que afecta la producción de kg de leche.

EVALUACIONES GENÓMICAS

Las evaluaciones genómicas consisten en establecer unas "fórmulas" mediante las que, a partir de la información que tiene un animal determinado en cada uno de los miles de SNP considerados, se predice su mérito genético en cada uno de los caracteres. Estas pruebas se denominan **valor genómico "directo"**, porque no se basan nada más que en los genotipos de los toros.

Actualmente, el valor genético de cualquier animal recién nacido se calcula como la media de las pruebas de sus padres y su fiabilidad está en torno al 30%. Pero, si se envía una muestra de sangre o pelo al laboratorio, se puede conocer qué información tiene ese animal en cada uno de los 58.000 SNP. Y, aplicando las fórmulas disponibles, se le calculará un valor genómico "directo". Y esta información está disponible al nacimiento del animal. Al combinar el índice de pedigrí y el valor genómico directo, la fiabilidad aumenta en media un 30% res-

Figura 4. Ejemplo explicativo del efecto de un SNP

	Posición SNP	Efecto del SNP en el animal
Animal 1		
Cromosoma del padre	----cgtaagat G taggct----	-1
Cromosoma de la madre	----cgtaagat A taggct----	+3
		+2 kg
Animal 2		
Cromosoma del padre	----cgtaagat G taggct----	-1
Cromosoma de la madre	----cgtaagat G taggct----	-1
		-2 kg
Animal 3		
Cromosoma del padre	----cgtaagat A taggct----	+3
Cromosoma de la madre	----cgtaagat A taggct----	+3
		+6 kg

pecto a la fiabilidad del índice de pedigrí. Si estamos en el caso de un toro recién nacido (ver figura 5), esta fiabilidad permitirá pre-seleccionar mejor qué toros jóvenes poner en testaje. Pero, además, a los 5 años de edad, cuando llegue su primera tanda de hijas, el toro alcanzará un 83% de fiabilidad, que aumentará al 88% gracias a la información del valor genómico directo. Es de especial importancia este aumento de la fiabilidad en caracteres que, como longevidad y fertilidad, tardan en conocerse y las pruebas de progenie tardan en alcanzar fiabilidades altas.

En el caso de una vaca (ver figura 6), si no se dispusiera de su genotipo en su segunda lactación, se estaría alcanzando una fiabilidad del 51%, inferior al 60% que se alcanza al nacimiento cuando el animal se ha genotipado. Esto hace que cambie radicalmente la información disponible para seleccionar madres de sementales.

Los valores citados para las fiabilidades se refieren al LPI canadiense según datos de Canadian Dairy Network (2009), pero nos dan una idea del impacto de las evaluaciones genómicas en las fiabilidades de los toros y de las vacas.

¿CÓMO SE CALCULAN LAS EVALUACIONES GENÓMICAS?

Para poder comenzar a calcular evaluaciones genómicas en un país es necesario disponer de los genotipos de los toros que tienen prueba en ese país y asociar las pruebas a los genotipos, para obtener las "fórmulas" que predicen los méritos genéticos a partir de los genotipos. Ese grupo de toros con el que se calculan las "fórmulas" de predicción del valor genómico directo, se llama "población de referencia".

El valor genómico directo para un carácter es la suma de los efectos asociados a todos los SNP en ese carácter. De hecho, en las evaluaciones genómicas el resultado que se busca es esa suma. Pero el efecto individual asociado a cada SNP, según las "letras" que haya en él, es de gran interés. A partir de ellos, se pueden seleccionar los SNP más informativos para diseñar nuevos chips con un número reducido de SNP y un coste mucho más bajo.

Figura 5. Evolución de la fiabilidad de un toro cuando se dispone de evaluaciones genómicas

AL NACIMIENTO

Valor genético= Índice Pedigrí
Fiabilidad=30%



Genotipo

SI EL TORO SE GENOTIPA

Valor genético= Índice Pedigrí + Valor genómico "DIRECTO"
Fiabilidad=60%

A LOS 5 AÑOS, con 100 hijas

Valor genético= Índice Pedigrí + Valor genómico "directo" + Hijas
Fiabilidad=88% (frente al 83% sin genotipar)

Figura 6. Evolución de la fiabilidad de una vaca cuando se dispone de evaluaciones genómicas

AL NACIMIENTO

Valor genético= Índice Pedigrí
Fiabilidad=30%



Genotipo

SI LA TERNERA SE GENOTIPA

Valor genético= Índice Pedigrí + Valor genómico "DIRECTO"
Fiabilidad=60%

A LA SEGUNDA LACTACIÓN

Valor genético= Índice Pedigrí + Valor genómico "directo" + Datos
Fiabilidad=68% (frente al 51% sin genotipar)

Estas fórmulas hay que estimarlas para cada población porque las asociaciones entre SNP y genes no tienen porque ser iguales en distintas poblaciones. Además hay que re-estimarlas periódicamente.

A MÁS TOROS MÁS FIABILIDAD

En la tabla 1 se puede observar la ganancia en fiabilidad en cada carácter en los toros jóvenes respecto al índice de pedigrí cuando éste se combina con el valor genómico directo, obtenido a partir del genotipo. Según estos datos de EEUU, la ganancia media en fiabilidad está en torno al 30%. Este incremento cobra un valor especial en caracteres como longevidad o fertilidad, que se conocen tarde en la vida de las hijas de los toros. Solo por esta razón las pruebas tardarían mucho en tener fiabilidades altas. Pero además, estos caracteres presentan baja heredabilidad y, debido a ello, necesitan añadir mucho mayor número de hijas que los caracteres de producción para alcanzar la misma fiabilidad. Es por eso que un aumento del 30% en la fiabilidad de estos caracteres equivale a un número mayor de hijas.

Un aspecto clave es que cuanto mayor sea el número de toros genotipados para el estudio de asociación entre pruebas y SNP, mayor fiabilidad tendrán las evaluaciones genómicas. Es por eso que las fiabilidades obtenidas para las valoraciones genómicas han ido aumentando a medida que se ha ido aumentando el número de toros incluidos en los estudios. De hecho, los incrementos en fiabilidad en EEUU mostrados en la tabla 1 están referidos a ecuaciones de predicción obtenidas con cerca de 4.500 toros y verificadas con unos 2.000 toros, pero los últimos resultados obtenidos en abril de 2009 con casi el doble de toros muestran que el aumento de fiabilidad es aún mayor (ver tabla 2). Si con 4.500 toros la ganancia en fiabilidad del Net Merit era de 24 puntos ahora es de 34.

Aunque en EEUU disponen de cerca de 3.000 vacas genotipadas, no han observado mejoras en la fiabilidad cuando se incluyen sus pruebas y genotipos en los cálculos de las ecuaciones de predicción. Probablemente será debido a la baja fiabilidad de sus pruebas. Lo que sí se ha detectado una vez más es que en las evaluaciones genéticas tradicionales, las madres de sementales tienden a estar sobrevaloradas, probablemente debido a tratamientos preferenciales.

Una manera de aumentar el número de toros a considerar en la población de referencia es incluir toros sin prueba nacional pero con prueba MACE. En la tabla 3 se puede observar como cuando Canadá utilizaba solo 1.095 toros para obtener sus ecuaciones de predicción, el aumento en la fiabilidad respecto al índice de pedigrí era de 10 puntos en los caracteres de producción. Pero, si consideraban 3.000 toros adicionales de EEUU con sus pruebas MACE en Canadá, la fiabilidad aumentaba en 25 puntos, 15 puntos más que cuando no se utilizaban las pruebas MACE.

Tabla 3. Aumento de la fiabilidad respecto al índice de pedigrí al incluir toros con prueba MACE (Schenkel, 2009)

	Solo Pruebas Nacionales Canadá	Pruebas Nacionales Canadá +Pruebas MACE en Canadá
Toros utilizados	1095	4121
Toros para verificar	524	524
Aumento fiab. production	+10%	+25%
Aumento fiab. RCS	+6%	+17%
Aumento fiab. tipo (20)	+12%	+18%

PAÍSES

En abril de 2008, EEUU comienza a calcular las evaluaciones genómicas y en enero 2009 las evaluaciones genómicas se incorporan a sus evaluaciones genéticas oficiales. Canadá comparte los genotipos con EEUU y comenzará en Agosto 2009 a publicar sus evaluaciones genómicas. Nueva Zelanda y Holanda están en situaciones parecidas a la de EEUU, pero en Holanda las pruebas genómicas son calculadas por Holland Genetics y no las comparte con la unidad de evaluación ge-

Tabla 1. Aumento de la fiabilidad respecto al índice de pedigrí en las evaluaciones de genómicas de EEUU para toros jóvenes sin hijas cuando se dispone de los genotipos (VanRaden, 2009).

Carácter	Ganancia en fiabilidad
Net merit	+24%
Leche	+26%
Grasa	+32%
Proteína	+24%
% Grasa	+50%
% Proteína	+38%
Vida productiva	+32%
Recuento células somáticas	+23%
Tasa preñez de las hijas	+28%
Calificación Final	+20%
Profundidad de ubre	+37%
Angulo Podal	+25%
Media de 27 caracteres	+29%

Tabla 2. Aumento de la fiabilidad respecto al índice de pedigrí para toros jóvenes sin hijas al aumentar el número de toros utilizados en las ecuaciones de predicción (VanRaden, 2009)

Número de toros		Aumento de la fiabilidad respecto al índice de pedigrí	
Predictores	Predichos	NM\$	Media 27 car.
2130	261	+13%	+17%
3576	1759	+23%	+23%
4422	2035	+24%	+29%
7600	9690	+34%	+33%



nética holandesa, con lo cual no se pueden integrar en las evaluaciones genéticas oficiales de ese país. Francia y Alemania tampoco integran por el momento sus evaluaciones genómicas con las oficiales, aunque en ambos países los cálculos son realizados por las unidades de evaluación genética. Otros países que están ya trabajando con las evaluaciones genómicas son Irlanda, Dinamarca, Suecia y Finlandia, los tres últimos conjuntamente.

En España, CONAFE, el grupo de investigación INIA-UPM y los programas de testaje españoles han acordado un plan para la puesta en marcha de las evaluaciones genómicas. Con ello se obtendrán unas ecuaciones de predicción para estimar el valor genómico directo de los animales a partir de la información de los SNP. Posteriormente se integrarán en las evaluaciones genéticas oficiales actuales basadas en datos y genealogías. Se genotiparán inicialmente cerca de 2000 toros. Para poder disponer de los genotipos de estos toros, es decir, para conocer la información que tienen en cada SNP, es necesario disponer de una muestra

de sangre, semen o pelo. Por eso, CO-NAFE ha estado localizando todas las dosis de semen de toros viejos existentes en los centros de inseminación artificial. En una reunión de Interbull celebrada en Enero de 2009, un investigador de EEUU presentó un estudio según el cual las fiabilidades de los toros jóvenes al nacimiento podrían alcanzar el 81% si se trabajara conjuntamente entre todos los países para constituir una población de referencia de 40.000 toros. Ese valor de fiabilidad es el mínimo que se exige en muchos países, por ejemplo en España, para aparecer en las listas de mejores toros. Pero, por el momento la colaboración entre países es limitada. Todo esto hace que, a día de hoy, las fiabilidades de las pruebas genómicas todavía sean bajas para competir con los toros con hijas para su uso por los ganaderos, aunque sean muy interesantes para los programas de testaje.



CHIP DE BAJO COSTE

Dado que el coste de genotipar un animal para el chip de 58.000 SNP es demasiado alto para su uso masivo, en EEUU se plantearon seleccionar los SNP que mejor predicen el Net Merit para incluirlos en un chip más pequeño, que pudiera comercializarse a un coste mucho más bajo. De esta forma se podría utilizar, por ejemplo, para preseleccionar madres de sementales o, incluso, para tomar decisiones de selección sobre las novillas de reposición. Finalmente, teniendo en cuenta los costes de cada uno de los chips y la capacidad de predecir el Net Merit, se decidió desarrollar un chip de solo 384 SNP. Por una parte, incluiría 100 SNP para poder realizar verificaciones de parentesco en todas las razas de vacuno. El resto de los SNP permitiría predecir el valor genómico directo para Net Merit, aumentando la fiabilidad del Net Merit respecto al índice de pedigrí en 10 puntos, frente 24 puntos con el total de SNP. Y ello con un coste de 25 dólares. Con 2000 SNP el aumento en fiabilidad era bastante mayor que con el chip de 384 SNP, pero el coste también. El inconveniente que se han encontrado con los 384 SNP es que toda la ganancia en fiabilidad del Net Merit es través de los caracteres de producción y que los caracteres funcionales no aumentan nada su fiabilidad. Actualmente todavía no ha comenzado su comercialización.

¿QUÉ CAMBIARÁ EN LOS PRÓXIMOS AÑOS?

Los próximos años veremos como surgen nuevos métodos estadísticos que permitirán aumentar las fiabilidades obtenidas actualmente y se desarrollaran chips de bajo coste para permitir el uso masivo de estas nuevas tecnologías. Los principales países comenzarán a realizar evaluaciones genómicas y las integrarán con sus evaluaciones genéticas en base a datos y genealogías. En el corto plazo, INTERBULL debe validar las evaluaciones genómicas nacionales y se debe perfeccionar y armonizar el cálculo de las fiabilidades. Asimismo, todavía está pendiente que la Unión Europea autorice la comercialización de toros jóvenes basados en sus pruebas genómicas, ya que actualmente se exige que las pruebas estén basadas en hijas. Por otra parte, el aumento de las fiabilidades de los caracteres funcionales permitirá aumentar su peso relativo en los índices de selección, como es el caso del ICO.

En cuanto a los programas de testaje, se generalizará la preselección de los toros jóvenes al nacimiento aumentando el valor genético medio de los toros que finalmente sean testados. En algunos programas de testaje podría disminuir el número de toros testados anualmente, especialmente en aquellos con más toros. Un aspecto que va a cambiar radicalmente es la selección de madres de sementales, ya que al nacimiento de una ternera ya se puede disponer de un índice genético que tiene mayor fiabilidad que el que se obtiene

ahora en la segunda lactación, y además la mayor fiabilidad en los caracteres funcionales permitirá tomar mejores decisiones de selección. Pero además de todo esto, se podrá evitar el problema de los tratamientos preferenciales. Todo esto ampliará el abanico de familias que contribuyan con madres de sementales. Adicionalmente, gracias a los SNP, se puede establecer con mayor exactitud la fracción de ADN que comparten los toros a testar en un mismo centro de inseminación artificial y se deberá tener en cuenta para controlar mejor el parentesco entre ellos.

Los ganaderos van a sufrir una gran presión comercial para la utilización de grupos de toros con prueba genómica pero sin hijas, con el argumento de que utilizando grupos de toros se minimiza el riesgo de la todavía baja fiabilidad de las pruebas genómicas predichas mediante el índice de pedigrí y los genotipos. Pero parece difícil que estos toros consigan una cuota de mercado importante mientras sus fiabilidades no alcanzan valores entre el 80 y el 90%. El aumento de la fiabilidad al nacimiento es muy interesante para los centros de IA en la preselección de los toros jóvenes y madres de sementales, pero son todavía bajas para competir con los toros con hijas. Sin embargo, una ventaja especialmente interesante para los caracteres funcionales de baja fiabilidad es que, como se ha explicado anteriormente, los toros podrán alcanzar mayores fiabilidades con su primera tanda de hijas y eso permitirá que las pruebas fluctúen menos cuando llega la segunda tanda tras su uso masivo. También sería posible estimar el valor genético directo de animales que no están ni en control lechero ni en el libro genealógico, aunque las fiabilidades serían bajas, en torno al 30%. Y en un futuro, cuando el coste del genotipado se reduzca mucho, será posible realizar programas de acoplamiento en base a los genotipos sin utilizar genealogías.

Las organizaciones de recogida de datos desarrollarán nuevos servicios de recogida de muestras, banco de ADN y genotipado. La información de los SNP se incorporará a las bases de datos nacionales como una información adicional para las evaluaciones genéticas. Las pruebas de progenie y la recogida de datos deberán mantenerse para poder obtener pruebas con fiabilidades altas y para poder re-estimar las ecuaciones de predicción cada cierto tiempo. Al mismo tiempo será aún más importante la recogida de nuevos datos, como la incidencia de enfermedades, ya que sería posible establecer evaluaciones genómicas sobre una parte de la población y que pudieran luego aplicarse a cualquier animal genotipado. En un futuro las verificaciones de parentesco se realizarán en base a los SNP e incluso se podrán establecer "relaciones genómicas" entre animales sin genealogía, según la fracción de ADN que compartan.

