ACTAS

DE LA

VII REUNIÓN NACIONAL

SOBRE

MEJORA GENÉTICA ANIMAL

LLEIDA 21 - 22 ABRIL 1994

ACTAS

DE LA

VII REUNIÓN NACIONAL

SOBRE

MEJORA GENÉTICA ANIMAL

LLEIDA 21 - 22 ABRIL 1994

Este volumen recoge el programa de la VII Reunión Nacional de Mejora Genética, así como las comunicaciones que se recibieron por escrito.

La organización de esta reunión fue posible gracias a la generosa ayuda de los siguientes organismos:

- Ministerio de Educación y Ciencia,
- Universitat de Lleida,
- Institut de Recerca i Tecnologia Agroalimentàries,
- Institut d'Estudis Ilerdencs.

A todos ellos, así como a los participantes, el Comité Organizador desea expresar su más sincero agradecimiento.

J.L. Noguera,
J. Estany,
M. Pérez-Enciso,
D. Babot,
L. Alfonso.
UdL-IRTA, 25006 Lleida

Programa

Día 21 de abril de 1994. Mañana

9'00 - 9'30 Inscripción.

Sesión : La Economía en la Mejora Genética

9'30 - 10'15 A. Blasco y A. Muñoz (Univ. Politécnica de Valencia, Univ. de Murcia).
Algunos aspectos económicos de los programas de mejora genética §.

10'15 -10'30 Discusión.

10'30 -11'00 A. Álvarez Pinilla (Centro de Experimentación Agraria, Asturias).
Nuevas perspectivas de la aplicación de la economía en la mejora animal §.

11'00 -11'15 Discusión.

11'15 -11'30 Pausa. Café.

11'30 -12'00 J. Pena y R. Alenda (CONAFE, Univ. Politécnica de Madrid).
Aspectos económicos en los programas de mejora genética en el ganado vacuno de leche §.

12'00 -12'15 Discusión.

12'15 -12'45 J. Cañón, J. Piedrafita y J. Altarriba (Univ. Complutense de Madrid, Univ. de Zaragoza, Univ. Autónoma de Barcelona).
Objetivos de selección en vacuno de carne §.

12'45 -13'10 Discusión de la comunicación. Discusión general de la sesión.

13'30 -15'15 Pausa. Comida.

Día 21 de abril de 1994. Tarde

Sesión : Comunicaciones libres

15'30 -15'45 D. Sorensen (National Institute of Animal Science, Dinamarca).
Inferencia bayesiana usando muestreo de Gibbs.

15'45 -16'00 L. Varona, C. Moreno, L. A. García-Cortés, J. Altarriba (Univ. de Zaragoza).
Análisis de sensibilidad en la estimación multicarácter de componentes de (co)varianza en vacuno lechero mediante muestreo de Gibbs.

16'00 -16'15 J. A. Baro, J. A. Carriedo y F. San Primitivo (Univ. de León).
Modelos matemáticos para la estimación de parámetros genéticos de la producción láctea y el recuento de células somáticas (SCC) en ganado ovino.

16'15 -16'30 M. Serrano, M. D. Pérez-Guzmán, V. Montoro y J.J. Jurado (INIA, Madrid; CERSYRA, Valdepeñas).
Mejora del modelo de evaluación genética en el esquema de selección de la raza ovina manchega.

16'30 -16'45 E. Díaz Carrillo, C. Angulo Heras, I. Jiménez Gamero, F. Baena Manzano, A. Alonso Moraga, A. Muñoz Serrano y J. M. Serradilla Manrique (Univ. de Córdoba).
Variación cuantitativa y cualitativa de las caseínas de la leche de cabra.

16'45-17'00 Pausa. Café.

17'00-17'15 E. Ugarte, E. Urarte, F. Arrese, J. Arranz e I. Beltrán de Heredia (CIMA, Vitoria).
Coste económico del programa de mejora genética y selección de las ovejas de raza Latxa y Carranzana en la Comunidad Autónoma del País Vasco.

17'15 -17'30 M. Analla, A. Muñoz-Serrano, A. Sánchez y J. M. Serradilla (Univ. de Córdoba).
Desarrollo y perspectivas del esquema de selección de la raza ovina Segureña.

17'30 -17'45 J. Tibau, J. Soler, N. Trilla (IRTA, Monells).
Aplicaciones del BLUP-Multicarácter a la evaluación oficial de reproductores porcinos del Libro Genealógico en Catalunya.

17'45 -18'00 P. Zaragoza, A. Romero y E. García-Muro (Univ. de Zaragoza).
Detección de portadores de la deficiencia de adhesión

leucocitaria (BLAD/LAD Bovine) en reproductores utilizados en España.

18'00-18'15 C. Elduque, S. Marcos e I. Martín (Univ. de Zaragoza). Contribución al conocimiento del mapa bovino (BOVMAP Project): Primeros resultados obtenidos en la búsqueda de nuevos marcadores.

18'15-18'30 C. Rodellar, F. Vázquez, R. Osta e I. Zarazaga (Univ. de Zaragoza).
Grupos sanguíneos vs. polimorfismos del DNA en la identificación del ganado vacuno: Una controversia en vías de solución.

18'30 -19'00 Discusión general de la sesión. Preparación de la VIII Reunión Nacional sobre Mejora Genética Animal.

22 de abril de 1994. Mañana

Sesión : Relación entre marcadores genéticos, genes mayores y caracteres cuantitativos

9'00 -9'45 J. M. Elsen (INRA-SAGA, Castanet-Tolosan, Francia). Investigación e identificación de genes de caracteres cuantitativos en animales domésticos (en inglés) §.

9'45 -10'00 Discusión.

10'00 -10'15 Pausa. Café.

10'15 -10'45 M. A. Toro y L. Silió (INIA, Madrid). Recientes progresos en el rastreo de genes en especies domésticas : Implicaciones en mejora animal §.

10'45 -11'00 Discusión.

11'00 -11'30 C. López-Fanjul (Univ. Complutense, Madrid). La descripción genética de los caracteres cuantitativos §.

11'30 -11'45 Discusión.

11'45 -12'00 Pausa. Café.

12'00 -12'30 E. Carbonell (IVIA, Valencia). Modelos estadísticos para la detección de genes que controlan la expresión de un carácter cuantitativo §.

12'30 -12'55 Discusión de la comunicación. Discusión general de la sesión.

12'55 -13'15 Clausura de la VII Reunión Nacional sobre Mejora Genética Animal.

13'15 - Comida de clausura.

---

§ Ponencia invitada.

Índice de Comunicaciones Escritas

A. Blasco y A. Muñoz. Algunos aspectos económicos de los programas de mejora genética. 1

J. M. Elsen. Recherche e identification de gènes à effets quantitatifs chez les animaux domestiques. 64

M. A. Toro y L. Silió. Recientes progresos en el rastreo de genes en especies domésticas: Implicaciones en mejora animal. 89

C. López-Fanjul de Argüelles. La descripción genética de los caracteres cuantitativos. 98

E. A. Carbonell y M. J. Asíns. Statistical models for the detection of genes controlling quantitative trait loci expression. 107

L. Varona, C. Moreno, L. A. García-Cortés, J. Altarriba. Análisis de sensibilidad en la estimación multicarácter de componentes de (co)varianza en vacuno lechero mediante muestreo de Gibbs. 134

J. A. Baro, J. A. Carriedo y F. San Primitivo. Modelos matemáticos para la estimación de parámetros genéticos de la producción láctea y el recuento de células somáticas (SCC) en ganado ovino. 141

M. Serrano, M. D. Pérez-Guzmán, V. Montoro y J. J. Jurado. Mejora del modelo de evaluación genética en el esquema de selección de la raza ovina Manchega. 145

E. Díaz Carrillo, C. Angulo Heras, I. Jiménez Gamero, F. Baena Manzano, A. Alonso Moraga, A. Muñoz Serrano y J. M. Serradilla Manrique. Variación cuantitativa y cualitativa de las caseínas de la leche de cabra. 148

E. Ugarte, E. Urarte, F. Arrese, J. Arranz e I. Beltrán de Heredia. Coste económico del programa de mejora genética y selección de las ovejas de raza Latxa y Carranzana en la Comunidad Autónoma del País Vasco. 151

M. Analla, A. Muñoz-Serrano, A. Sánchez y J. M. Serradilla. Desarrollo y perspectivas del esquema de selección de la raza ovina Segureña. 156

P. Zaragoza, A. Romero y E. García-Muro. Detección de portadores de la deficiencia de adhesión leucocitaria (BLAD/LAD Bovine) en reproductores utilizados en España. 159

C. Elduque, S. Marcos e I. Martín. Contribución al conocimiento del mapa bovino (BOVMAP Project): Primeros resultados obtenidos en la

búsqueda de nuevos marcadores. 162

C. Rodellar, F. Vázquez, R. Osta e I. Zarazaga. Grupos sanguíneos vs. polimorfismos del DNA en la identificación del ganado vacuno: Una controversia en vías de solución. 165

J. Cañón, J. Piedrafita y J. Altarriba. Objetivos de selección en vacuno de carne. 168

VII REUNION NACIONAL SOBRE MEJORA GENETICA ANIMAL
Lleida, 21-22 de Abril de 1994

ALGUNOS ASPECTOS ECONOMICOS DE LOS PROGRAMAS DE MEJORA GENETICA

I. PARTE TEORICA
Agustín Blasco
Universidad Politécnica de Valencia

II. PARTE APLICADA
Antonio Muñoz
Universidad de Murcia

I. PARTE TEORICA

A. Blasco
Departamento de Ciencia Animal. Universidad Politécnica de Valencia.

Cuando José Luis Noguera me propuso preparar una ponencia para nuestra reunión bianual, mi primera intención fue negarme, naturalmente. El tema que me proponía, la economía de los programas de mejora, me es hasta cierto punto ajeno, y por otra parte uno ya anda, como todos, bastante liado con sus asuntos cotidianos. Pero ya sabéis lo difícil que es negarse a los requerimientos de José Luis. Además, el ambiente amistoso y poco formal en que estas reuniones se desenvuelven, hace que uno espere más indulgencia de la que en realidad merece por meterse en campos que no le son propios. Porque en realidad mi interés en la economía de la mejora no pasa de ser un interés de usuario: en Valencia desarrollamos programas que tienen una incidencia directa entre los cunicultores del entorno. Por otra parte debemos explicar en clase estos temas, lo que nos obliga a un cierto enciclopedismo autodidacta que, si bien es divertido, nos procura más de una preocupación; ya sabéis que los autodidactas solemos pasar de una ignorancia absoluta, y a costa de un gran esfuerzo, a la mayor de las confusiones.

En el caso de los temas relacionados con la economía de la mejora este autodidactismo es más notorio que en otros temas, puesto que la información relevante anda dispersa y confundida con la irrelevante. A ello se une nuestro desconocimiento u olvido, en general, de la jerga económica (temible en manos de jóvenes doctores con ganas de progresar) y también el hecho de que parece un tema alejado de las preocupaciones que debiera tener un honrado genetista que trabaje para la Administración -y creo que de estos somos aquí mayoría aplastante-. Tal vez por eso sean tan populares los artículos en los que los pesos económicos son susceptibles de grandes errores sin que a los programas de selección les ocurra nada, y tal vez por eso nuestros libros de texto -escritos por genetistas ajenos a las empresas- son tan parcos en estos temas.

Y sin embargo son asuntos importantes. En primer lugar porque la economía de la mejora es algo más que el cálculo de los pesos económicos de un índice y en segundo lugar porque no es tan cierto que errores en estos temas sean más o menos irrelevantes para el resultado de un programa de selección, particularmente cuando se trata de seleccionar los objetivos de mejora o de realizar evaluaciones a largo plazo. Es frecuente citar el artículo de Charlie Smith (1983) como prueba de esta irrelevancia, olvidando que el propio Charlie indica en ese artículo que hay un buen número de casos en los que errores en el cálculo de pesos económicos conducen a pérdidas moderadas o considerables. Recientemente (1994) Amer, Fox y el propio Smith han señalado que, dependiendo de cómo se calcule la función de beneficios, los pesos económicos y la evaluación económica del progreso genético diferirán notoriamente. Por lo que respecta a los temas de la economía relacionados con la mejora genética, determinar cuáles son estos temas es algo bastante arbitrario. La relación que doy a continuación se basa -con algunos añadidos- en la de un curso que John Gibson impartió en Dinamarca en el 92 y amplió en Guelph en el 93. En realidad se basa más en las publicaciones de quienes

se han interesado por los temas que en los temas que tal vez podían ser de interés para la mejora. Recuerdo que en una conferencia sobre antropología en la Royal Society of Edinburgh, el conferenciante nos decía que el mapa de fósiles encontrados se correspondía muy bien con el de buscadores de fósiles.

Los temas son los siguientes:

1. Pesos económicos en índices de selección (incluyendo asuntos como pesos económicos no lineales, situaciones en que un producto no es comercializable si su valor es inferior a una cierta cantidad, pesos económicos en situación de cuotas y pesos económicos para caracteres categóricos).

2. Funciones de beneficios derivadas usando teoría económica estándar (lo que, supongo que para aumentar el prestigio de la técnica, llaman aplicación de la teoría neoclásica de la firma comercial ([1]).

3. Técnicas de programación lineal aplicadas a problemas relacionados con la mejora genética.

4. Técnicas de programación dinámica (árboles de decisión, por ejemplo) aplicadas a problemas relacionados con la mejora genética.

5. Modelos bio-económicos (usualmente emplean técnicas de simulación).

6. Técnicas de flujo de genes para predecir progresos genéticos en esquemas de mejora. Van unidos a tasas de descuento para evaluar económicamente programas a largo plazo (me refiero al descuento producido no sólo por la inflación sino por el coste de oportunidad del capital).

7. Funciones de utilidad (incorporación del riesgo de error al predecir la respuesta y otros riesgos en la evaluación económica de programas de mejora).

8. Mercado compartido y actitud de una empresa de mejora frente a sus competidores.

Es obvio que no es posible hablar de todo esto ahora. A lo más que puedo aspirar es a hacer una pequeña revisión sobre los temas que me parecen más interesantes. En este punto debo involucrar a John Gibson, pues una parte de lo que os voy a contar fue discutido en su curso, aunque advierto que estas notas son producto de mi propia revisión y contienen lógicamente mi visión del asunto, que no quisiera yo entremeter a John en errores y desaciertos a los que es ajeno.

---

[1] No sé lo que quieren decir con "the neoclasical theory of the competitive firm". La teoría neoclásica en economía data de cuando Alfred Marshall publicó en 1890 su libro 'Principles of economics', nada precisamente de última hora. Por otra parte los economistas que hablan de situaciones de competencia perfecta se refieren al 'Nuevo modelo clásico'.

De los temas que he citado, el primero me parece el más relevante, puesto que afecta directamente a los programas de mejora. El tema del uso de la teoría económica estándar es atractivo, aunque no parece decisivo a la hora de plantear pesos económicos o elegir entre varios esquemas posibles de mejora. Las técnicas de programación lineal y programación dinámica se han aplicado a decisiones de usuario de programas: qué comprar según las expectativas que ofrece el producto y su precio, por ejemplo, o a toma de decisiones en cuanto a eliminación de animales (*culling*), aunque son herramientas tan generales que supongo que pueden ser aplicadas a cualquier campo (por ejemplo, una forma de derivar índices con restricciones es usar técnicas de programación lineal). Me niego a comentar los aburridísimos artículos del grupo de Dickerson sobre modelos bioeconómicos. Mucho me temo que estas aproximaciones, si están bien, no dan resultados diferentes a las aproximaciones con funciones de beneficios, y si están mal, los errores son mucho más difíciles de detectar.

Las técnicas de flujo de genes y tasas de descuento son interesantes y se pueden aplicar a comparaciones de esquemas de mejora diferentes. No las trataré aquí por limitaciones de espacio y de tiempo, pero para vuestro alivio os diré que el artículo base de estos métodos es uno de los pocos artículos de Bill Hill relativamente sencillo de leer.

Sobre la incorporación de riesgo a la evaluación de programas de mejora han trabajado John Woolliams y Theo Meuwissen (1993) desde una perspectiva bayesiana . El tema parece interesante: consiste en esencia en seleccionar sobre una función de beneficios que incorpora la incertidumbre producida por la diferente cantidad de información de los candidatos, lo que es particularmente relevante en vacuno de leche. El centro del problema reside en que la selección sobre las estimas de valores aditivos maximiza la respuesta en una generación, pero cuando se consideran más generaciones depende de la variabilidad en la cantidad de información de los candidatos.

El último de los temas listados supone un planteamiento nuevo a la hora de calcular los pesos económicos, por lo que lo he desgajado del apartado 1. En un artículo aparecido en 1989, de Vries plantea algo que en el mundo empresarial era perfectamente conocido: los intereses de la empresa de mejora son la venta de genes, y esto no está necesariamente relacionado con la evaluación económica del progreso genético sino con lo que los consumidores están dispuestos a aceptar: es posible que una híbrida poco prolífica compense económicamente su baja prolificidad con buenos resultados en otros caracteres, pero seguramente los granjeros preferirán comprar sus hembras a otras empresas que les den más prolificidad. Si la cantidad de hembras que se vende es fija, la empresa debe considerar la mejora de sus caracteres en relación con sus competidores y con lo que el comprador desea. De Vries hace una primera aproximación al problema, tal vez algo elemental, pero me da la impresión de que es un campo abierto sugestivo para futuras investigaciones.

En esta revisión me ceñiré al tema de pesos económicos incluyendo asuntos relacionados con el uso de teoría económica estándar y la competencia entre empresas de mejora.

==========

Quiere nuestra tradición que atribuyamos al oscuro Farfiled Smith la invención de los índices de selección con sus pesos económicos incluidos, luego utilizados en animales por Hazel en 1943. La verdad es que los índices -sin pesos económicos- fueron ya propuestos (¡como no!) por Pearson en 1903. Por lo que respecta a los índices para varios caracteres, es cierto que en el citado -y por supuesto escasamente leído- artículo de 1936 Smith propone estos índices para una aplicación concreta en plantas. En el apartado de agradecimientos, señala que la parte teórica de su artículo es poco más que una transcripción literal de las ideas que Ronald Fisher le ha transmitido para atacar su problema -esto explica mejor quién pudo resolver un asunto de tal relevancia-.

La definición de pesos económicos es sencilla: el incremento de beneficio debido al incremento *genético* de un carácter. Esto se suele calcular computando la diferencia de beneficios entre la situación actual y la situación en la que un carácter aumenta una unidad -esto es, manteniendo los otros caracteres constantes-([2]). La causa, además de ser de un cierto sentido común, la podemos deducir de la propia formulación de la función de beneficios. El beneficio es una función de un conjunto de caracteres **x**, un conjunto de precios de los productos **v**, los costes variables **c**, y los costes fijos **k**

$$B = f(x_1, x_2, \ldots, x_n ; v_1, v_2, \ldots, v_m ; c_1, c_2, \ldots, c_p ; k_1, k_2, \ldots, k_q)$$

habitualmente se considera que los precios y los costes (tanto fijos como variables) son constantes sea cual sea el nivel de producción (de *output*). La función de beneficios es entonces una función sólo de los caracteres

$$B = f(x_1, x_2, \ldots, x_n)$$

Podemos linealizarla mediante una serie de Taylor (supongamos, para abreviar, que trabajamos con valores centrados),

$$B \approx \left[\frac{\partial f}{\partial x_1}\right]_{x_i = x_i, \forall i} \cdot x_1 + \left[\frac{\partial f}{\partial x_2}\right]_{x_i = x_i, \forall i} \cdot x_2 + \cdots + \left[\frac{\partial f}{\partial x_n}\right]_{x_i = x_i, \forall i} \cdot x_n$$

por tanto parece 'lógico' (o al menos me lo parece a mí) llamar pesos económicos a las cantidades que multiplican a los caracteres ([3]). Queda el problema de los términos de orden superior. Habitualmente se supone que como la mejora genética progresa

---

[2] Hay otra forma de aproximarse al problema que no la trato porque da prácticamente el mismo resultado y es más complicada de computar. Si se dispone del valor económico de un número suficientemente grande de animales, se puede calcular el peso económico como la regresión de un carácter sobre los beneficios (ver, p. ej. Schlote, 1977).

[3] El asunto es algo más complejo, si se quiere ser estricto. Para una descripción rigurosa ver, por ejemplo, Elsen et al., 1986

lentamente estos términos son despreciables; es decir, la aproximación lineal es razonable en el punto concreto en el que se aplica -en ese caso el cociente entre incremento de beneficio y de carácter al que nos referimos antes sería también una aproximación-. Esto es cierto, pero queda aún la duda de qué pasa a largo plazo. A largo plazo puede ocurrir que la relación entre producto y beneficio sea lineal (en ese caso las derivadas de orden superior son nulas), pero es poco probable por varias razones, entre ellas la previsible saturación del mercado. Si no es lineal lo que ocurre es que los valores *medios* que se usan para calcular los términos entre corchetes (los pesos económicos) cambian, y esto tiene que ser tenido en cuenta por el programa de mejora, puesto que entonces lo pesos económicos serán diferentes. En ocasiones este asunto puede dar lugar a serias dificultades, particularmente debido a la poca seguridad con que se prevé la relación producto-beneficio en el futuro o a la falta de información en la literatura científica para encontrar respuestas a situaciones en las que la mejora genética cambia la relación producto-beneficio (en la figura 1, al final del texto, se pone un ejemplo).

Para complicar algo más las cosas, Bright (1991), Amer y Fox (1992) y Amer, Fox y Smith (1994) hacen notar que esa no es la forma en la que los economistas se comportan para calcular el peso económico de los caracteres. No deseo discutir aquí la Historia Natural de los economistas ni su comportamiento ante diversas situaciones, pero puede ser interesante examinar sus procedimientos, habida cuenta de que son más generales que los que ahora comento. Una descripción de estas propuestas alternativas para calcular pesos económicos la haré más adelante.

Antes de continuar, quiero resaltar la importancia de calcular bien la función de beneficios. Es importante en primer lugar porque describe la situación económica de la producción: si hay o no puntos singulares, los orígenes de los costes y de los ingresos, etc. Esta función va a ser decisiva en el paso más importante del programa: la selección de los objetivos y de los criterios de selección, y aquí un error sí que puede tener consecuencias trascendentales para el éxito del programa. Finalmente hay que evitar "ocultar" caracteres bajo otros caracteres; quiero decir que en ocasiones un carácter es función de otros, y el no hacerlo explícito conduce a errores en el cálculo de las derivadas y a confusiones a la hora de determinar los objetivos. Por ejemplo, si se venden canales a pesos muy diferentes, en el apartado "costes" de un programa no puede aparecer una variable "coste de la alimentación por individuo" que sea independiente del peso de la canal , puesto que este coste depende del peso de la canal.

Pese a la sencillez del procedimiento se plantean al menos tres interrogantes:

1.- ¿Es posible que el incremento del beneficio debido al incremento *genético* sea diferente del debido al incremento fenotípico?

2.- Si al aumentar un carácter debido a la selección se produce un aumento correlativo en otro carácter ¿se debe tener en cuenta esta correlación al calcular los pesos económicos?¿qué media utilizamos para calcular el peso económico, la previa o la posterior a la selección?

3.- Parece que estamos suponiendo que si a un incremento de carácter corresponde un incremento de beneficio, al doble de incremento de carácter corresponderia el doble de incremento de beneficio (es decir, que las relaciones entre beneficios y caracteres es lineal). Es obvio que en muchas situaciones no es así (cuotas, umbrales de precios, etc.) ¿qué ocurre en esos casos? Es más, a largo plazo no es probable que esta linealidad se mantenga (por ejemplo, por saturación del mercado). Dado que los programas de mejora sólo tienen efectos apreciables a largo plazo ¿hasta qué punto es lógico mantener esta linealidad?

Puestos a plantear interrogantes, hay tres preguntas más que inciden directamente en la aplicación de los pesos a los programas de mejora:

4.- ¿Qué ocurre cuando en los esquemas de mejora hay intereses contrapuestos? Quiero decir que al ganadero puede importarle poco el rendimiento a la canal si no se lo pagan, pero no al matadero (por ejemplo). El contenido en carne interesa al matadero, pero no al consumidor, que compra lonchas. Si la calidad de carne no se paga no le interesará al matadero, pero sí al propietario del restaurante y al consumidor. Cabe pensar que los pesos económicos dependerán entonces de las unidades en las que se expresen (por animal, por kg de producto, etc.). Además, pueden haber funciones económicas diferentes para el mismo tipo de empresario que tenga su negocio en regiones diferentes o en zonas en las que la mano de obra u otro coste sean distintos.

5.- El beneficio se puede expresar como ingresos menos costes, pero también como ingresos dividido por costes (o por su inversa, costes divididos por ingresos). Esto da lugar a pesos económicos diferentes (más adelante veremos un ejemplo sencillo) ¿cuáles se deben usar?

6.- Si una empresa se encuentra en el óptimo de producción, la mejora genética puede desplazarle de ese óptimo ¿Es posible encontrar pesos económicos que muevan a un nuevo óptimo productivo?

Aunque la siguiente pregunta no comprometa al honrado funcionario, es importante para quienes vayan a trabajar para la empresa privada:

7.- ¿Son los pesos económicos calculados como antes hemos expuesto los que maximizan el beneficio de la empresa de mejora? Hay que tener en cuenta que, en situaciones de competencia, la mejora de un carácter que está muy por debajo de los competidores tiene una importancia muy superior a cuando tiene un nivel similar al de otras empresas, por ejemplo.

Las dos primeras preguntas son reflexiones elementales y fáciles de resolver. La tercera incluye una variedad bastante grande de casos que hay que tratar por separado. Las dos últimas preguntas parecían resueltas en el tan famoso como oscuro artículo de Smith, James y Brascamp de 1986 (raramente citado, y en algunas ocasiones incorrectamente), cuando Amer y Fox propusieron lo que consideraban una mejor solución al problema, planteando al mismo tiempo la pregunta 6. No estoy seguro de que Amer y Fox aporten

mucho a la solución de las preguntas 4 y 5, aunque otra cosa es que debamos replantearnos estos asuntos a la luz (al parecer particularmente brillante) de la *teoría neoclásica de la firma comercial*, que es la que Amer y Fox utilizan para iluminarnos. La séptima pregunta ha sido tratada en el cuasi-célebre artículo de De Vries (1989), pero involucra también asuntos como los índices para ganancias deseadas, puesto que la actitud de las compañías de mejora suele ser la de sobrecargar los pesos de los caracteres cuyo valor consideran que no están a la altura de la competencia.

¿Son estas cuestiones meros asuntos de interés académico? En parte tal vez esto sea así, aunque no despreciaría yo la posibilidad de hacer lo mismo de antes pero con las ideas más claras. Por ejemplo, en algunas ocasiones hemos comentado en estas reuniones que el BLUP más que eficacia aporta claridad: un índice con correcciones no da resultados muy diferentes pero se comprenden mejor los que provienen de un modelo y una forma de estimar como la del BLUP. Tal vez el examen detallado de las cuestiones propuestas y el uso de teoría económica estándar no den resultados muy diferentes a los que hoy manejamos, pero tal vez comprendamos mejor lo que hacemos. Sobre si estos asuntos son tarea de los economistas, imagino que lo son del mismo modo que aplicar modelos lineales a la mejora es tarea de estadísticos: lo importante no es quién lo haga sino cuál sea su resultado.

La utilidad de un refinamiento en las técnicas que se usen para calcular pesos económicos depende, obviamente, del tipo de programa al que se apliquen y de la situación del mercado. En la práctica de las empresas es frecuente cambiar los pesos totales del índice si algún carácter interesante no mejora mucho. Esto no implica corregir errores en los pesos económicos sino en el conjunto parámetros genéticos - pesos económicos (habitualmente los errores en los primeros son mayores). Pero como estoy seguro de que no aceptaríais este argumento para dejar de calcular los parámetros genéticos (a esto sí que se dedican los honrados funcionarios), pasaré a exponeros los temas relacionados con los pesos económicos con más detalle.

1.- ¿Es posible que el incremento del beneficio debido al incremento *genético* sea diferente del debido al incremento fenotípico?

Las respuesta a la pregunta primera es sencilla: no siempre coinciden. Si tenemos una función de beneficios fenotípica, un índice fenotípico I_E (útil para valorar a un animal, pero no a su descendencia)

$$I_E = a_1 x_1 + a_2 x_2 + \ldots$$

donde a_1 y a_2 son los pesos económicos, y tenemos un genotipo agregado A_E (yo prefiero llamarle valor aditivo económico, la expresión genotipo agregado me suena rara),

$$A_E = w_1 A_1 + w_2 A_2 + \ldots$$

donde los pesos económicos son w_1 y w_2 , es obvio que, para que las ecuaciones de dimensiones respectivas cuadren, a_i se expresa en $/valor fenotipico mientras que w_i se expresa en $/valor aditivo. En realidad la forma de calcular ambos pesos será habitualmente la misma, por ejemplo

a_i = ΔBeneficio / Δfenotípico del carácter

w_i = ΔBeneficio / Δgenético del carácter

Ocurre que en algunos casos estos valores no coincidirán. Se me han ocurrido tres ejemplos, pero seguro que hay muchos más:

a) En el caso del tamaño de camada en porcino, el incremento de producción de un lechón da un beneficio determinado (así se calcula a_i). Sin embargo, si este incremento es genético, podemos argüir que disponemos de una línea hiperprolífica (hoy en día todo el mundo tiene una línea hiperprolífica), con lo que el incremento de beneficio sería superior, puesto que este tipo de hembras está sobrevalorado (los directores de comercialización de las empresas dicen que no lo está, puesto que el valor de algo es lo que el comprador está dispuesto a pagar por ese algo).

b) Si un carácter no se hereda, obviamente tiene un peso a_i , pero es absurdo tratar de un peso w_i puesto que numerador y denominador son cero.

c) En el caso de que un carácter se encuentre en el óptimo, cualquier incremento del carácter produce un decremento del beneficio. Sin embargo es posible que la selección cambie ese óptimo, por lo que ambos pesos no coincidirían (en la figura 1, al final de este documento, lo intento explicar algo mejor).

2.- Si al aumentar un carácter debido a la selección se produce un aumento correlativo en otro carácter ¿se debe tener en cuenta esta correlación al calcular los pesos económicos?¿qué media utilizamos para calcular el peso económico, la previa o la posterior a la selección?

La pregunta 2 tiene un matiz. En la primera generación de selección, como los pesos económicos w_i se aplican sobre valores aditivos, las correlaciones entre los caracteres ya han actuado. Quiero decir que como la leche y su contenido en grasa están relacionados negativamente, animales con valor aditivo alto para un carácter tenderán a tenerlo bajo para el otro. Esto implica que las medias de los caracteres que se usan para calcular los pesos económicos deberían ser las medias *después de la selección*, pero como para calcular estas medias hace falta conocer los pesos económicos, habría que actuar de forma iterativa. Afortunadamente el progreso que los genetistas conseguimos en una generación es tan minúsculo que este asunto carece de importancia. Otro asunto es qué ocurre a la larga. Aquí las opiniones son controvertidas, desde quien opina que las pequeñas desviaciones producidas por el cambio de las medias debido a la selección no es de esperar que aumenten la ineficacia habitual del proceso (Smith, James y

Brascamp en su artículo de 1986), hasta quien advierte de notables diferencias en la estimación de la respuesta y de los pesos económicos según se tenga en cuenta o no la falta de linealidad a la que hace referencia la pregunta 3 (por ejemplo, Amer, Fox y el propio Smith, 1994).

3.- Parece que estamos suponiendo que las relaciones entre beneficios y caracteres es lineal. Es obvio que en muchas situaciones no es así ¿qué ocurre en esos casos? Dado que los programas de mejora sólo tienen efectos apreciables a largo plazo ¿hasta qué punto es lógico mantener esta linealidad?

La pregunta 3 trata un tema más complejo y que incluye situaciones distintas. las respuestas que se han dado al problema han sido de tres tipos:

1) *Usar índices no lineales* (cuadráticos, cúbicos, etc.). Tienen muchos problemas. En el caso de función de beneficios lineal $B = w_1A_1 + w_2A_2 + \ldots$, el índice óptimo es $I = w_1\hat{A}_1 + w_2\hat{A}_2 + \ldots$, y el problema se reduce a encontrar las estimas $\hat{A}_1$, $\hat{A}_2$, ..., mientras que si la función de beneficios es $B=f(A_1 , A_2 , \ldots)$, el índice no lineal óptimo *no* es $f(\hat{A}_1 , \hat{A}_2, \ldots)$. Además no maximizan la respuesta a la selección truncada. Esto se puede ver fácilmente con un ejemplo propuesto por Goddard (1983): si el índice fuera

$$I = x^2$$

donde x es el valor observado de un carácter, entonces se seleccionarían los individuos con valores extremos (los de alto **x** y los de muy negativo **-x**), con lo que la media en la siguiente generación sería cero. Como es natural, Goddard (cuyo artículo es muy bonito y fácil de entender, os lo recomiendo) trata casos más generales concluyendo que ciertos índices lineales aventajan a los no lineales para optimizar el beneficio. A estos índices nos referiremos a continuación. Entre nosotros me parece recordar que José Luis Campo y su grupo han trabajado experimentalmente en este asunto encontrando resultados en el mismo sentido.

2) *Usar ciertos índices lineales* . Esta solución fue propuesta por primera vez para dos caracteres y de forma gráfica por Moav y Hill nada menos que en 1966 ([4]). No deja de sorprenderme que, siendo una solución correcta, no se haya realizado más investigación hasta 1988, año en que Itoh y Yamada dieron con una solución analítica al planteamiento gráfico de Moav y Hill. Pasternak y Weller propusieron en 1993 otra solución analítica también aplicada a varios caracteres. Me pesa no tener espacio para exponer ambas soluciones, que requieren un cierto tiempo de explicación, pero remito al artículo de Goddard para obtener una visión rápida e intuitiva de la metodología propuesta. La solución analítica temo que no es sencilla, aunque creo que es interesante investigar sobre sus posibles aplicaciones.

---

[4] En realidad fue Goddard (1983) quien indicó que el ejemplo gráfico de Moav y Hill servía para calcular los pesos económicos

3) *Usar índices con restricciones*. La idea es que para ciertos caracteres en los que hay umbrales (por ejemplo en peso de huevo, en el que a partir de cierto peso el precio es el mismo y por debajo de cierto peso no los pagan) puede hacerse mejora con la condición de que cierto carácter no modifique su valor medio. Hace unos años, en una reunión como ésta, en Madrid, discutí con Angel Bosch y José Luis Campo si los índices con restricciones podían ser óptimos (esto es: conducir a beneficios máximos). Mi opinión es la misma que entonces: NO. Intentaré exponer de nuevo los argumentos que utilicé -con poco éxito-. Supongamos que tenemos dos caracteres fuertemente correlacionados negativamente y decidimos imponer la restricción a uno de ellos para que no cambie. En ese caso sólo queda una pequeña parte de variación disponible para seleccionar al otro carácter. Si aplicamos un índice sin restricciones obviamente se producirán algunas pérdidas al aumentar el porcentaje de individuos rechazados, pero se producirá también una mejora más sustancial en el otro carácter. Si las pérdidas por el aumento de individuos rechazados compensan o no con las ganancias producidas por la mejora del otro carácter depende del peso económico de los caracteres, óptimo que varía con la selección.

De todas formas los índices con restricciones son objeto de abundantes críticas (Gibson y Kennedy 1993, Goddard 1983, James 1982) y no seré yo quien incida más en un asunto cuya polémica me parece superflua. Como hace notar Pasternak y Weller, nadie ha pretendido que son índices óptimos, sino solamente índices que hacen los que se proponen: fijar un carácter permitiendo la mejora en otros.

4.- ¿Qué ocurre cuando en los esquemas de mejora hay intereses contrapuestos?

La pregunta 4 tiene dos partes. La última parte la planteó ya Hazel en 1943: cada ganadero tiene su función de beneficio. Esta cuestión es, sin embargo, de difícil solución: la mejora actúa sobre grandes grupos de ganaderos, por lo que un programa de selección no puede dar satisfacción a un gran número de intereses particulares. En vacuno la ventaja de disponer de un catálogo de semen hace que un ganadero pueda construir funciones de beneficios propias para su caso particular, y de alguna forma es lo que las compañías de aves o cerdos intentan diversificando sus líneas (cuello desnudo o ponedoras enanas en aves, cruzamientos algo más rústicos o machos terminales más conformados en cerdo, por ejemplo), pero es claro que pese a todo la acción de un programa de mejora se ejerce sobre un número de ganaderos muy grande.

La otra parte de la pregunta, la existencia de grandes bloques de intereses contrapuestos, es más delicada. A este interrogante respondieron primero Brascamp, Smith y Guy en 1985 diciendo que si se incluyen los beneficios como un coste de producción entonces los pesos económicos son los mismos desde cualquier perspectiva. Naturalmente que esto plantea la pregunta de por qué deben incluirse estos beneficios como costes, y se puede argumentar de muchas formas. Por ejemplo, se puede decir que en situaciones de competencia perfecta el mercado actúa de tal forma que el trabajo del granjero es realmente lo que se considera beneficio, por lo que debe incluirse como un coste. O se puede argüir que debido a la competencia todos los granjeros tienen beneficios similares,

por lo que el hecho de tener beneficio empresarial no es muy diferente al hecho de tener costes de mano de obra. Este beneficio, que no es otro sino el requerido por el productor para estar en el negocio a largo plazo, es lo que al parecer los economistas llaman *beneficio normal*, y lo incluyen en el apartado de costes. Aunque toda esta argumentación es razonable, e incluso un tanto abrumadora cuando requiere al prestigio de los profesionales de la economía, yo no dejo de tener la impresión de que estos argumentos son *a posteriori*; es decir, que no se argumenta primero que los beneficios deben incluirse como parte de los costes y *consecuentemente* las distintas perspectivas económicas a las que hace referencia la pregunta 4 se equiparan, sino que al descubrir que la inclusión de los beneficios como costes resuelve este problema, uno busca a continuación argumentos que justifiquen esta inclusión. Esto, naturalmente, no invalida el razonamiento, pero tal vez haga que algunos de nosotros lo tomemos con una cierta precaución. Veamos como funciona en un ejemplo sencillo:

Supongamos que una empresa produce **n** individuos por hembra (los aficionados a las matemáticas que no se inquieten, el uso de negrilla no implica que sea un vector, es para distinguirlo mejor del texto), a peso comercial **w**, y que el coste *diario* de mantener un individuo es $\mathbf{c_I}$ mientras que el coste *anual* de mantener a la hembra es $\mathbf{c_H}$. Si **d** es el número de días que vive un individuo en la explotación y **v** el precio por kg vivo de los individuos producidos, y si consideramos que la función de beneficios es ingresos menos costes **B = R-C** (utilizo para ingresos la nomenclatura en inglés R=returns para evitar usar I, letra habitual en índices),

Por hembra

$$R = nwv \quad ; \quad C = nc_Id + c_H \quad ; \quad B = nwv - nc_Id - c_H$$

$$\frac{\partial B}{\partial n} = \bar{w}v - c_I\bar{d}$$

Por individuo

$$R = wv \quad ; \quad C = c_Id + (c_H/n) \quad ; \quad B = wv - c_Id - (c_H/n)$$

$$\frac{\partial B}{\partial n} = \frac{c_H}{\bar{n}^2}$$

el peso económicos del tamaño de camada **n** es, pues, distinto según se calcule por hembra o por individuo, y los mismo le ocurre a los pesos de **w** y **d**, tanto si se calculan de forma absoluta como relativa (es decir, tomando valores respecto a uno de ellos). Sin embargo, si B=0 las dos ecuaciones de beneficios pasan a ser la misma y, por tanto,

también los pesos económicos. Brascamp et al. ofrecen, por supuesto, una demostración general para este asunto, demostración sencilla que no vale la pena copiar aquí.

Puede resultar poco atractivo para el ganadero el que le digan que se trabaja sobre la base de beneficio cero (esto es, que el beneficio de la mejora va a parar al consumidor), y una empresa de mejora insistirá en aprovechar las pequeñas variaciones del mercado que la alejan de la competencia perfecta (por ejemplo, ofertando productos nuevos: cerdas hiperprolíficas, antes de que todo el mundo tenga una cerda hiperprolífica), pero hay un argumento simple que persuadirá al granjero a invertir en mejora genética: si no lo hace, sus competidores -que sí que invierten en mejora- acabarán por arrojarle fuera del mercado.

5.- El beneficio se puede expresar como ingresos menos costes, pero también como ingresos dividido por costes. Esto da lugar a pesos económicos diferentes ¿cuáles se deben usar?

La solución dada a la pregunta 5 incluyó también una respuesta a la pregunta 4, y fueron sus autores Smith, James y Brascamp. Descubrieron que, admitiendo ciertos supuestos, no sólo el cálculo de los pesos económicos era el mismo independientemente de qué perspectiva se tomara, sino que también daban el mismo resultado si se consideraba el beneficio como Ingreso/Coste en lugar de como Ingreso-Coste. Este último asunto es menos trivial de lo que parece, y pondré un ejemplo sencillo del propio James (1983) para aclararlo.

Supongamos que tenemos una función de beneficios muy simple: los ingresos vienen de la venta de un producto **W**, por ejemplo kg de canal, a precio **a** por unidad (por kg). Los ingresos son, por tanto

$$R = a \cdot W$$

El coste por unidad de producto (p. ej., alimentación, instalaciones, etc. , por kg de canal) es **F** y el precio por unidad es **b**.

$$C = b \cdot F$$

Si la función de beneficios es **B = R - C**, los pesos económicos de **F** y **W** son

$$w_W = (\delta B/\delta W) = a$$

$$w_F = (\delta B/\delta F) = - b$$

o, si se quiere de forma relativa, $\mathbf{w_W} = 1$, $\mathbf{w_F} = \mathbf{-b/a}$. Sin embargo, si la función de beneficios es **B = C/R**, los pesos económicos son

$$w_W = \frac{\partial B}{\partial W} = \frac{-b \cdot \overline{F}}{a \cdot \overline{W}^2}$$

$$w_F = \frac{\partial B}{\partial F} = \frac{b}{a \cdot \overline{W}}$$

y, de forma relativa,

$$w_W = 1$$

$$w_F = -\frac{\overline{F}}{\overline{W}}$$

que no sólo difieren tanto en términos absolutos como relativos sino que en el primer caso son dependientes de los precios y en el segundo no.

Este asunto de expresar los beneficios como diferencia o como cociente es, probablemente, un tema clásico en economía, pero la primera vez que yo lo he visto discutido en una revista de producción animal es en un artículo de Harris de 1970. Dickerson tiene también un artículo en el mismo año en el que propone el cociente **C/R** como medida de la eficiencia. La razón que habitualmente se da (p. ej., James, 1982) es que se debe favorecer la disminución de costes en lugar del aumento de la producción para evitar saturar el mercado y que los precios bajen. El cociente **C/R** es una forma aproximada de disminuir costes a nivel de retornos constante. Aunque es cierto que el aumento de la producción puede saturar el mercado, lo es sólo en parte, primero porque cada vez hay más gente y cada vez la gente que hay consume más. En un artículo magistral de Cunnigham publicado en un congreso sobre vacuno de carne se recoge la evolución del consumo de carne de varias especies ganaderas, y hay un aumento global de la carne consumida en los últimos años, aumento más pronunciado en carnes baratas como pollo y cerdo. Expresar los pesos como cociente es mejorar la eficacia del sistema, pero es discutible que coincida con el interés a corto plazo del ganadero.

Pasaré ahora a exponer los argumentos de Smith et al. (1986). Las condiciones para que los pesos económicos sean los mismos independientemente de la perspectiva tomada son, tal y como ellos lo indican en su artículo, expresar los costes fijos por unidad de producto y cambiar el tamaño de la empresa para que su producción equipare los cambios en producción obtenidos genéticamente. Así dicho queda como algo más bien arbitrario. Voy a reformular las condiciones para intentar aclararlas:

1. *Los costes tradicionalmente considerados como 'fijos' (amortización de instalaciones, mano de obra, financiación, maquinaria, etc.) dependen ahora del nivel de producción; es decir, pasan a ser costes variables.*

El argumento para aceptar esto es la escala de tiempo en que se mueve la Mejora

Genética. A largo plazo la maquinaria e instalaciones se habrán amortizado, y los costes laborales cambiarán con el tamaño de la empresa, que será a su vez modificado según sea más o menos rentable. Si no se admite esta forma de razonar, se puede considerar que la mejora genética tiene un efecto muy amplio, de forma que los programas de mejora afectan a un conjunto grande de ganaderos. Admitiendo esto, se tiene, pues, a un conjunto de ganaderos en continua inversión y modificación de factores fijos; no cada uno individualmente, pero sí en conjunto, debido a que siempre hay alguien que entra o sale del negocio, cambia el tamaño de la empresa, etc. Finalmente, se puede argumentar que la mejora genética no trabaja para cualquiera sino para gente eficiente; es decir, si un ganadero no tiene su empresa en el máximo de eficacia debe modificarla para aprovechar los beneficios de la mejora, y si ya está en ese máximo qué duda cabe que un incremento de su producción debido a la mejora le forzará a modificar sus costes fijos para llegar al nuevo máximo de eficacia ([5]).

La causa del interés en librarse de los efectos fijos es que influyen en la forma de estimar los pesos económicos. Volviendo al ejemplo de James de antes, si consideramos costes fijos k, los costes serían

$$C = b \cdot F + k$$

Si la función de beneficios es **B = R - C**, los pesos económicos de **F** y **W** son los mismos de antes, **a** y **-b** , sin que intervengan para nada los costes fijos. Sin embargo, si la función de beneficios es **B = C/R**, los pesos económicos son

$$w_W = \frac{\partial B}{\partial W} = \frac{-b \cdot \bar{F} - k}{a \cdot \bar{W}^2}$$

$$w_F = \frac{\partial B}{\partial F} = \frac{b}{a \cdot \bar{W}}$$

en los que intervienen los costes fijos.

2. *Los beneficios obtenidos deben considerarse nulos, tal y como se expone en la solución a la pregunta 4, pero debe incluirse en el cálculo los que hubieran podido ser obtenidos cambiando el tamaño de la empresa.*

La razón de querer considerar los beneficios nulos es la expuesta antes: Nuevamente el

---

[5] Por artificioso que parezca, está admitido por los economistas que los costes fijos desaparecen a largo plazo. Una buena introducción para refrescar la economía que la mayor parte de nosotros aprendimos y nos apresuramos a olvidar es el libro de Lipsey "Introducción a la economía positiva". Amer y Fox, los de la 'teoría neoclásica' usan la edición americana *Economics*, publicada en colaboración con Purvis y Steiner.

argumento es la escala de tiempo en que se mueve la mejora genética. Por ejemplo, si el óptimo de producción en vacuno de leche se encuentra en granjas de 40 vacas, lográndose así una producción media de 5.000 kg/año (no sé nada de vacas, considerad que es sólo un ejemplo), resultaría absurdo hacer programas de mejora dirigidos a granjas de 4 vacas, que producen 4.000 Kg/año de media, para que produjeran más. Además, en situaciones de competencia perfecta los granjeros de 4 vacas irán (al menos a largo plazo) desapareciendo.

Esta condición produce el mismo efecto que la de considerar al beneficio como un coste (redu p l cirlo a cero, como hicimos antes), iguala los pesos económicos independientemente de que el interés que se defienda sea el del ganadero de engorde o de ciclo completo, pero iguala además las perspectivas de considerar la función de beneficio como una diferencia entre ingresos y costes o como un cociente.

Veamos un ejemplo sencillo. Tomemos la función de beneficios que hemos usado para responder a la pregunta 4, pero sin costes fijos

$$B = R - C \; ; \qquad R = nwv \; ; \qquad C = nc$$

$$B = nwv - nc = n(wv - c)$$

Aquí el factor de escala es **n**. Un beneficio dB se obtiene al incrementarse el producto en un dw. Si **n** y **c** son constantes,

$$dB = n \cdot v \cdot dw$$

Sin embargo, un beneficio extra se puede obtener también cambiando el tamaño de la empresa (en ese caso, **n** no es constante),

$$dB = (\delta B/\delta w)dw + (\delta B/\delta n)dn = nv\, dw + (vw - c)\, dn$$

Smith et al. consideran qué ocurre si se impone la condición de que el beneficio obtenido por el incremento del carácter se iguale al obtenido al cambiar de escala (dB=0), y también cuando esta equiparación se produce sólo en ingresos (dR=0) o sólo en costes (dC=0). Concluyen que en los tres casos los pesos económicos relativos son proporcionales a los derivados tomando una función de beneficios del tipo B=R/C ([6]).

Los pesos económicos calculados de esta forma no varían sustancialmente de los calculados de forma tradicional. Por ejemplo, cuando equiparan los ingresos producidos por el cambio en un carácter con los producidos por el cambio de escala, la diferencia reside en que los ingresos están ponderados por un factor de eficiencia **C/R**. En las

---

[6] Raúl Ponzoni (1988) ha calculado los pesos económicos para un caso de merino en Australia considerando el beneficio como **R-C, R/C y C/R**, y sus conclusiones confirman las del trabajo de Smith et al., aunque es crítico a la hora de aceptar la desaparición de los costes fijos.

empresas ganaderas la cantidad C/R es próxima a 1 (por ejemplo, un ganadero que lleve 100 cerdas en ciclo completo ingresa 30 millones de pesetas anuales, pero gasta 28 -si tiene suerte-). Esto hace que el efecto de cambio de escala sea pequeño a corto plazo. A largo plazo las cosas pueden ser bastante diferentes, como hicieron notar Amer, Fox y Smith, 1994.

La parte matemática del artículo de Smith et al. está expuesta de forma bastante oscura, particularmente la fórmula general que figura en el apéndice. En el curso de J. Gibson al que asistí nadie acabó con una idea muy clara sobre la forma en que deducen los pesos, y el propio Charlie Smith me comentó que tampoco entendía muy bien el apéndice, escrito por James al parecer durante un viaje en tren (un tren nocturno, imagino). Parte del problema viene de la poca claridad con que se ofrecen las definiciones y parte del uso de incrementos y derivadas parciales. Posteriormente encontré un modo de explicar este asunto de forma más sencilla usando diferenciales; os lo incluyo en un apéndice a estas notas por si puede ayudaros a entender el artículo.

6. Si una empresa se encuentra en el óptimo de producción, la mejora genética puede desplazarle de ese óptimo ¿Es posible encontrar pesos económicos que muevan a un nuevo óptimo productivo?

Hay una aparente contradicción en el procedimiento de descontar los beneficios debidos a un cambio de tamaño de la empresa. Si la empresa se encuentra en su óptimo de producción, la mejora genética alterará este óptimo, y si no se encuentra en el óptimo parece que la sugerencia es: vaya usted primero al óptimo y mejore luego. El asunto es particularmente obvio en el caso de la existencia de cuotas a la producción. De todas formas se puede argüir que el cambio en la cantidad de producto debido a la mejora genética debería tenerse en cuenta, sobretodo si se cree que hay un óptimo productivo fuera del cual cambios en la cantidad de producto conducen a disminuciones en el beneficio.

Para el caso de cuotas en vacuno de leche, John Gibson propone una modificación al procedimiento de Smith et al. de forma que tiene en cuenta que, si **y** es la producción de leche y **n** el número de vacas,

$$(n - dn)(y + dy) = n \cdot y$$

es decir, que se mantenga la producción total a costa de reducir el número de vacas. La situación de cuota es delicada porque depende de una política agraria que no necesariamente va a ser la misma a medio o largo plazo, pero el tema de reoptimizar la producción ha sido tratado con mayor amplitud por Amer y Fox utilizando teoría económica estándar.

Amer y Fox (ahora con Smith, en un artículo de 1994) se sienten sorprendidos por la forma de comportarse de los mejoradores genéticos ante el problema de los pesos económicos. Al parecer un economista agrario lo primero que haría es plantear la

función de producción del granjero y encontrar para qué niveles de *inputs* (alimentación, número de animales, etc.) se encuentra el beneficio óptimo, y seguidamente averiguaría dónde se encuentra el nuevo beneficio óptimo para un incremento de un carácter determinado, siendo el peso económico del carácter el incremento entre ambos beneficios dividido por el incremento del carácter. Esto se trata en la primera parte del artículo, que es prácticamente una copia literal de un artículo de Bright publicado en 1991. Voy a exponer el procedimiento con más detalle.

Al parecer, si a un economista agrario se le deja suelto frente al problema de la evaluación del incremento del beneficio que se produce al incrementarse un carácter, se comporta de la siguiente manera:

1.- *Decide qué aspecto tiene la función de producción.*

Una función muy popular entre economistas agrarios es la de Cobb y Douglas (es un artículo antiguo, 1928, y está citado por Bright, 1991 y por Amer et al., 1994. Esta función se encuentra también junto a otras en un libro de Dillon, 1977, pero he mirado el libro sólo por encima y no sé qué relevancia tiene escoger una función u otra). Para el caso del ejemplo que hemos usado con Smith et al., una empresa de producción de carne tendría una función del tipo

$$y = k \cdot w \cdot C_1^a \cdot C_2^b$$

donde **y** es la cantidad de producto producido (p. ej. total de kg de canal vendidos, los *outputs*), k es una constante, **w** es el peso individual de una canal, $\mathbf{C_1}$, $\mathbf{C_2}$, ..., son los factores (los *inputs*) asociados a la producción (alimentación mano de obra, etc.), y los coeficientes **a, b,** ..., son lo que los economistas llaman *elasticidades parciales de producción*, y representan la proporción relativa en la que el producto **y** aumenta cuando aumenta el factor correspondiente ([7]). Normalmente los coeficientes **a, b,**... varían entre 0 y 1. Un ejemplo que pone Bright para ovino es el siguiente: Si se producen ovejas a un peso medio de 20 kg,

$$y = 23.467 \cdot 20 \cdot N^{0.34} L^{0.24} F^{0.15}$$

donde **N** es el número de ovejas, **L** la superficie de pasto y **F** el alimento consumido (no importa en qué unidades se den, pero debe cuadrar la ecuación de dimensiones: al final se cuentan kg de canal producidos en total). Naturalmente la función de Cobb y Douglas no es la única función que puede usarse, y en el libro de Dillon como he dicho antes se discuten otras.

---

[7] Creo que quieren decir lo siguiente: considerando constante todo menos $\mathbf{C_1}$, tenemos que log **y** = cte. + **a** log $\mathbf{C_1}$, por tanto $dy / y = a \cdot dC_1 / C_1$

2.*Construye la función de beneficios y calcula el valor de los factores que la maximizan*

En este caso la función de beneficios tiene un aspecto muy simple:

$$B = y \cdot p_y - (C_1 \cdot p_{C1} + C_2 \cdot p_{C2})$$

donde las **p** son los precios del producto y de los factores de producción. A continuación se deriva respecto a $\mathbf{C_1}$ y $\mathbf{C_2}$, se iguala a cero y se calculan los valores de $\mathbf{C_1}$ y $\mathbf{C_2}$ que producen el máximo beneficio, a los que llamaremos $\mathbf{C^m_1}$ y $\mathbf{C^m_2}$. Esto factores también producen una cantidad de producto óptima, $\mathbf{y^m}$ calculada a partir de la función de Cobb y Douglas. Hasta el momento no se ha producido ningún incremento en los caracteres. La función de beneficios es, en su máximo,

$$B^m = y^m \cdot p_y - (C^m_1 \cdot p_{C1} + C^m_2 \cdot p_{C2})$$

3.*Calcula la función de beneficios derivada de un cambio en un carácter y la maximiza.*

Supongamos que se produce un incremento en peso d canal, d**w**. La nueva función de producción será

$$y_1 = k\,(w + dw)\,C^a_1 \cdot C^b_2$$

que da lugar a una nueva función de beneficios $\mathbf{B_1}$ con $\mathbf{y_1}$ en lugar de **y**. Esta nueva función tiene otro máximo para otros valores de $\mathbf{C_1}$ y $\mathbf{C_2}$, que se pueden calcular análogamente, derivando e igualando a cero las derivadas.

El peso económico del carácter **w** es la diferencia entre los dos beneficios óptimos, el calculado antes de mejorar el carácter y el calculado después: $\mathbf{B^m_1 - B^m}$, dividido por el incremento del carácter d**w**.

Aunque a corto plazo esto puede parecer un refinamiento innecesario, a medio y largo plazo o cuando los cambios genéticos son moderados o grandes las diferencias entre aproximaciones es notable, particularmente si se utiliza la técnica de Smith et al. de descontar los efectos debidos a cambio del tamaño de la empresa (Amer, Fox y Smith, 1994).

Otro asunto que sorprende a estos señores (Amer y Fox, 1992) es la linealidad de las funciones de producción. En su modelo, el incremento de coste al aumentar una unidad de producto (el llamado coste marginal) decrece hasta llegar a un mínimo y luego aumenta al continuar aumentando la producción (de lo contrario no habría un nivel de producción óptimo). Sin embargo consideran que en una pequeña empresa el aumento de una unidad de producto siempre lleva al mismo aumento en ingresos, supongo que porque una pequeña empresa no puede saturar el mercado y producir una caída de precios. El óptimo de la función de beneficios se produce cuando se cortan las curvas

de coste marginal e ingreso marginal (es un resultado clásico en economía, aunque Amer y Fox lo demuestran en su artículo). Un incremento genético modifica las curvas y se llega a nuevos óptimos. En la figura 2, al final del texto, se representan en línea sólida la situación de partida, y en línea punteada la situación después de que actúe la selección. La curva de costes marginales inicial (MC_0) corta a la recta de ingresos marginales (MR) en el punto **d** ([8]), por tanto la situación óptima de partida es aquélla en la que la empresa produce una cantidad $\mathbf{Y}^*_0$. Tras actuar la mejora genética, se obtiene una nueva curva de costes marginales (MC_n) que corta a la de ingresos marginales (que como hemos comentado es la misma, MR) en el punto **g**, dando lugar a un nuevo nivel óptimo de producción, $\mathbf{Y}^*_n$. Mientras tanto, los costes medios de la empresa han pasado de la curva (AC_0) a (AC_n) ([9]).

Aunque, sorprendentemente, Amer y Fox no lo hacen explícito en su trabajo, el peso económico debiera ser, pues, el incremento de beneficio entre dos situaciones óptimas al aumentar una unidad el carácter. El beneficio inicial de la empresa era el área **abcd** de la figura; es decir: ingresos menos costes para una cantidad de producto $\mathbf{Y}^*_0$. Tras la selección el beneficio es el área **aefg** . El incremento de beneficio se debe, pues, a dos causas: la disminución del coste medio (más eficacia) es una y el aumento de producción la otra:

$$dB = Y^*_n\, dAC + (MC_0\text{-}AC_0)\, dY^*$$

correspondiente a las dos áreas rayadas de la figura.

Si utilizamos pesos económicos lineales, la curva de costes marginales pasa a ser una recta paralela a MR, con lo que no se alcanza nunca el óptimo (o se alcanza a tamaño infinito de la producción), lo que es una inconsistencia que ya hemos comentado antes en otro contexto. Esto les hace decir que el cambio de escala de Smith et al. es innecesario e inconsistente con la teoría neoclásica de la firma comercial. Aquí creo que introducen un error y pienso que deliberadamente. Smith et al. ponen un ejemplo de pesos económicos calculados con funciones de producción lineales, pero en ningún momento la teoría se restringe a funciones lineales (ver apéndice I). Además, como hacen notar Brascamp y de Vries (1992), el problema se presenta cuando el área inferior rayada es grande; esto es, cuando un incremento del carácter conduce a cambios notables en la curva de costes, lo que no es, en principio, esperable (yo añadiría que a corto plazo; a largo plazo se repiten las críticas hechas antes a funciones de beneficios no lineales, pero desde otra perspectiva).

El aporte principal de esta teoría es, probablemente, exponer con más claridad y realismo

---

[8] La figura está tomada del artículo de Amer y Fox, pero allí está mal dibujada y no se ve bien el que *d* es el punto de corte entre ingresos y costes marginales.

[9] $AC_0 = C_0/Y_0$ donde C_0 es el coste de producir la cantidad de producto Y_0 .
$MC_0 = \delta AC_0/\delta Y_0$

la situación económica de las empresas cuyo beneficio se pretende maximizar. Creo que tiene la ventaja de que los economistas han desarrollado un cuerpo de conocimiento que puede ser útil para estos menesteres. Requieren un conocimiento de la situación del mercado más profunda que la que exigen los métodos tradicionales y esto tal vez sea posible de adquirir a través de estudios privados o estudios realizados por funcionarios de la Administración, pero ignoro en estos momentos hasta qué punto es realista desarrollar los pesos económicos de la manera que proponen Amer y Fox.

7.- ¿Son los pesos económicos calculados como antes hemos expuesto los que maximizan el beneficio de la empresa de mejora?

El comportamiento de un mejorador de empresa ante la selección está influido por lo que cree que puede afectar a las ventas de su producto. Si un carácter flojea y eso puede ser determinante a la hora de recibir pedidos, aumentará el peso económico del carácter. Esto ha sido tratado formalmente como índices para ganancias deseadas (ver Brascamp, 1984 para una revisión), que han sido criticados por sus sensibilidad a errores en los parámetros genéticos (ver los trabajos de José Luis Campo y equipo, por ejemplo), pero no estoy seguro que sea allí donde reside el problema. Aparte del obvio componente aleatorio, causa razonable de que lo esperado diste de lo obtenido, los parámetros genéticos que se usan en los índices no están calculados, habitualmente, en la población sometida a selección. Además, cuando lo están y difieren de lo que se sospecha que pueden ser los valores verdaderos, tampoco se utilizan (dando lugar a un curioso bayesianismo de madera, como los sonetos de Borges). Puede ocurrir que la corrección del peso del índice sea realmente una corrección del parámetro genético errado. El problema es cómo formalizar todo esto para que carezca de arbitrariedad el peso que se aplica.

Un primer intento ha sido el de De Vries (1989). Supongamos que todos los clientes tienen acceso a los productos de todas las empresas. Consideremos que un carácter tiene un nivel mínimo de aceptabilidad por el comprador (primera hipótesis fuerte). Cada comprador tiene su nivel de aceptabilidad de forma que entre todos se distribuyen de forma Normal para cada carácter i (segunda hipótesis fuerte). La desviación típica de esa distribución, **s**, tiene que ser estimada por estudios de mercado. En la figura 3 se representa una Normal indicando la proporción $\mathbf{p_i}$ de clientes que aceptarían distintos niveles de un carácter. Si una empresa tiene un nivel de carácter $\mathbf{x_i}$, el área rayada es la proporción de clientes que aceptarían comprarles. El resto requiere niveles más altos del carácter para decidirse a comprar. La cuota de mercado de una empresa es la fracción de clientes que les compran, que si se fijan en tres caracteres, por ejemplo, será

$$Q = c \cdot p_1 \cdot p_2 \cdot p_3$$

donde c es una constante, supongo que para que al final todas las cuotas sumen el 100%. El peso económico de un carácter podría ser el incremento de cuota de mercado que se produce al incrementarse el valor del carácter

$$a_i = \delta Q / \delta x_i$$

El argumento puede hacerse más complejo introduciendo el precio de venta del producto (de la híbrida, por ejemplo) x_P . Puede calcularse cuánta cuota de mercado se gana al reducir el precio (aunque no siempre ocurre así: Picasso decía que incrementaba el precio de sus cuadros para no defraudar a su clientela).

$$a'_{iP} = \delta Q / \delta x_P$$

Finalmente de Vries propone como peso económico al aumento de la cuota de mercado producida al incrementarse el carácter, respecto al aumento que se produce al reducir el precio:

$$w_i = a_i / a'_{iP}$$

La deducción de estos pesos económicos es sencilla, pero como no está totalmente explicitada en el artículo, os ofrezco una versión detallada en el Apéndice 2.

No tengo una opinión definida sobre este método, y habrá que esperar a ver cómo funciona. Asuntos como el averiguar por estudios de mercado la desviación típica de la aceptabilidad o el factor de sustitución del que se habla en el apéndice 2 me son ajenos y no sé si presentan dificultades especiales, o si su precisión una vez estimados es elevada o no. Me da la impresión de que sobrevalora la cuota de mercado, al menos en especies de intervalo generacional corto, ya que si el consumidor paga suficientemente por un carácter, la reacción de las empresas es inmediata, lo que en su terminología vendría a querer decir que la desviación típica s es muy pequeña, o bien que los niveles de las empresas andan siempre cerca del nivel medio de aceptación. En otro contexto y en vacuno de leche, Jack Dekkers ha usado un modelo de simulación basado en las técnicas de flujo de genes que os comenté antes, para comparar programas alternativos de empresas de inseminación artificial. Aunque el problema que se plantea (cuál es la combinación óptima de toros a probar y de hijas por toro para optimizar el beneficio de la empresa) no es la misma que aquí tratamos, concluye que los principales beneficios, que pueden ser notables, provienen de las cuotas de mercado que las empresas acaparan, siendo asuntos como el precio del semen o las tasas de descuento más bien secundarios. En el artículo de Amer y Fox de 1992, proponen para vacuno de carne un modelo en el que la curva de demanda es muy elástica cuando se considera a una empresa relativamente pequeña o a un grupo de ganaderos reducido (esto es: pequeñas variaciones en el precio hacen que los consumidores prefieran ese producto), pero muy inelástica cuando se considera la demanda total. Esto hace que el beneficio de la mejora genética vaya en general a los consumidores, pero que las empresas extraigan sus beneficios de abarcar más cuota de mercado. En un artículo reciente, Brascamp y de Vries comentan que Olivier les ha dicho que, haciendo una simulación y comparando este método con el de ganancias deseadas, funciona mejor el nuevo método. Por otro lado, al aplicarlo un estudiante de doctorado, de Boer, a un programa de selección de aves de puesta, el excesivo énfasis que la solución daba al número de huevos ponía en

peligro las ganancias, debido a pérdidas en el peso de huevo que rebasaban los umbrales de aceptabilidad. Esto último sólo quiere decir que este método -como cualquier otro- debe usarse con sentido común, e intentando ser conscientes de qué es lo que no está incluido en el modelo que uno aplica.

==========
@ @ @
==========

Con esto he terminado la discusión de los puntos que me proponía. Hay algunos asuntos que quedan en el tintero por no haber tenido incidencia en el tema, como la solución de Bulmer (1985) al uso de funciones de beneficio no lineales -sólo practicable en líneas puras de plantas, en animales requiere tener los datos de todos los posibles grupos de animales que podrían ser seleccionados- o el problema tratado por Melton et al (1989) en el probablemente más vapuleado articulo de la historia reciente -la inclusión de variables de manejo en la función de beneficios-. Como este último caso ha sido criticado conceptualmente (Goddard 1983), matemáticamente (Thompson, 1980) y desde un punto de vista de su aplicación práctica (James, 1982), además de haberse demostrado que no hay una solución general al problema (Gibson en su curso de 1993), no creo que valga la pena ni exponerlo.

Cuando expuse este articulo en un seminario que tuvo lugar en mi Departamento, una compañera me hizo notar que a largo plazo pueden pasar muchas más cosas (al decir de Keynes, a largo plazo todos estaremos muertos). Pueden aparecer nuevos objetivos de selección y desaparecer otros. Pueden aparecer situaciones como las de cuota, en el que hay que incluir restricciones a la producción y en las que los pesos económicos ya no son lineales. Pueden aparecer avances tecnológicos que hagan superflua la mejora genética o que produzcan un cambio radical en sus métodos (ejemplo clásico: vacunas en programas de resistencia a enfermedades, aunque no es el único ejemplo; piénsese en las posibilidades de la genética molecular o la clonación). Yo no tengo respuesta a estas interrogantes, que me parecen razonables. Un programa de mejora genética a largo plazo debería tener en cuenta estas posibilidades, e incluso la posibilidad de que existan a largo plazo inversiones mejores que las de la genética. En realidad yo no creo que estos nuevos objetivos o situaciones sean imprevisibles, más bien suelen aparecer muchos años antes de que los mejoradores decidan alterar sus programas de mejora. Las cuotas en producción de leche y mantequilla hacía muchos años que se veían venir antes de que finalmente se discutiera sobre eficacia de programas bajo cuota, y las deducciones de Avalos y Smith que revalorizaron la selección por tamaño de camada se basan en un examen atento de los índices de selección, no en nuevos descubrimientos. Entiendo que en ocasiones se produce alguna pequeña revolución (la aparición del merino Booroola, la del test del halotano, e incluso la irrupción de las cerdas chinas, aunque se encuentran descripciones de su producción al menos desde 1929), pero la base de los programas creo que puede ser calculada a largo plazo razonablemente. Por ejemplo, la introducción del carácter 'contenido en proteína' en la evaluación de la leche

de vacuno y ovino, o la introducción de objetivos relacionados con la calidad de la carne en porcino, no es previsible que desaparezcan en muchos años.

Una condición implícita que me aparece, sin embargo, poco dudosa, es el plazo de los programas de mejora. Se puede argüir que en especies como pollo o conejo puede haber programas a corto plazo, pero no lo creo así. Las heredabilidades realizadas de la mayor parte de caracteres de interés económico no permiten pensar que el plazo de un programa de mejora sea breve, ni siquiera en especies en las que el intervalo generacional es corto. La excepción podría estar en el contenido en carne en porcino, favorecido por una elevada heredabilidad, pero ya en límites en los que poca mejora se va a poder hacer en el futuro. Es precisamente este largo plazo de todos los programas de mejora genética lo que hace tan atractivos los atajos (el gen mayor cuya caza es la ilusión de todo genetista, la clonación de individuos extraordinarios, etc.). Sin embargo esta misma longitud en la escala temporal es la que permite a los programas "rectificar sobre la marcha", que es la forma habitual de trabajar en programas no científicos (en los programas científicos se trabaja más explicando por qué la mejora no ha tenido el éxito esperado). Esto plantea hasta qué punto es importante la precisión en el cálculo de los pesos económicos, la determinación de los objetivos de mejora y la elección de esquemas o programas alternativos (por ejemplo: MOET frente a inseminación artificial o a algún programa que use ambas técnicas) y cuál debe ser la actitud del genetista ante los métodos de cálculo.

Respecto a la precisión de los pesos ya he comentado antes mi opinión: son bastante robustos a errores que no sean desmesurados, aunque hay ciertas situaciones (caracteres relacionados negativamente, caracteres dominantes en un índice) que deben tomarse con mayor precaución (todo esto está expuesto con claridad por Vandepitte y Hazel y por Charlie Smith, por lo que no vale la pena que reproduzca aquí sus argumentos). Hay que tener en cuenta que, aunque los pesos del índice (los globales, incluyendo la parte de parámetros genéticos) puedan ser modificados, los resultados de la selección tardan en verse varias generaciones, por lo que no es bueno decidir con frivolidad confiando siempre en que en el peor de los casos se pueden realizar modificaciones posteriores.

En cuanto al método de cálculo, depende de para quién se trabaje. En esencia, la conclusión de Smith et al. es que los pesos económicos deben calcularse como Ingresos/Costos para que sea indiferente la perspectiva de cálculo. Sin embargo para una empresa de mejora *no* es indiferente esta perspectiva. Ellos venden tratando de sacar partido a las pequeñas variaciones de la situaciones de competencia imperfecta, y sus intereses son los de los compradores (habitualmente los ganaderos), no los del sistema en conjunto, por tanto andan escasamente interesados en maximizar la utilidad general del sistema (Ingresos/Costos), sino más bien en incrementar la cuenta de resultados (Ingresos - Costos) o en ocupar zonas de mercado que ahora ocupan sus competidores (sistema de de Vries, pregunta 7). Si se trabaja para una asociación de ganaderos y se dispone de información suficiente, la forma de trabajar de Bright y de Amer y Fox es sin duda atractiva. Si se trata de un programa nacional y abarca los intereses de todo el sector, puede pensarse en la solución de Smith et al. Si no es un programa nacional pero está llevado a cabo por la Administración, también puede pensarse en este tipo de

soluciones.

En un futuro inmediato preveo que habrá más investigación en planteamientos como los de De Vries, Bright y Amer y Fox. Tienen soluciones similares a corto y medio plazo, aunque diferentes a largo plazo, pero sobretodo abordan los problemas desde perspectivas metodológicamente atractivas. Ya veremos.

REFERENCIAS

Os cito aquí las referencias que he utilizado para este discurso. No me he fijado si están todas las referencias que cito o si hay de más, pero os aseguro que las que vienen a continuación son las esenciales para introducirse en este tema, a falta de una referencia que parece importante, citada por Amer y Fox (1992) para la crítica al problema del cambio de escala de Smith et al., 1986, que no me ha sido posible conseguir. Esta referencia es McArthur, A.T.G., 1987, Wheiting breeding objectives, an economic approach *Proc. of the 6th annual conf. of the Australian Assoc. of Animal Breed. and Genet., Perth.* pp. 187-197.

Como en las guías gastronómicas, me ha parecido oportuno clasificarlas según su importancia. En negrita cito las que creo esenciales. Con una estrella figuran las interesantes, aunque no esenciales, y el resto son más bien complementarias.

ALLAN, B. B.; FREDEEN, H. T.; WEISS, G. M. 1985. The sensitivty of a two-trait selection index to changes in economic weights and genetic parameter estimates. *Can. J. Anim. Sci.* 65: 21-29.

AMER, P.R.; FOX, G. C.; SMITH, C. 1994. Economic weights from profit equations: appraising their accuracy in the long run. *Anim. Prod.* 58:11-18.

AMER, P. R.; FOX, G. C. 1992. Estimation of economic weights in genetic improvement using neoclassical production theory: an alternative to rescaling. *Anim. Prod.* 54: 341-350.

*BRASCAMP, E. W. 1984. Selection indices with constraints. *Anim. Breed. Abstracts* 52: 645-654.

*BRASCAMP, E. W.; VRIES, A.G. 1992. Defining the breeding goals for pig improvement. *Pigs News and Information* 13 (1): 21N-26N.

BRASCAMP, E. W.; SMITH, C.; GUY, D. R. 1985. Derivation of economic weights from profit equations. *Anim. Prod.* 40: 175-180.

*BRIGHT, G. 1991. Economic weights from profit equations: appraising their accuracy. *Anim. Prod.* 53: 395-398.

BULMER, M.J. 1985. The mathematical theory of quantitative genetics. *Claredon Pess*

CUNNINGHAM, E. P. 1988. Meat production in the community. *In: Proc 3rd World Congress Sheep and Beef Cattle Breeding.* Paris 1988. vol 1, 3-21.

DE VRIES, A. G. 1989. A method to incorporate competitive position in the breeding goal. *Anim. Prod.* 48: 221-227.

DEKKERS, J. C. M.; SHOOK, G. E. 1990. Economic evaluation of alternative breeding programs for commercial artificial insemination firms. *J. Dairy Sci.* 73: 1902-1919.

DICKERSON, G. 1970. Efficiency of animal production-molding the biological components. *J. Anim. Sci.* 30 (6): 349-359.

DILLON, J.L. 1977. The analysis of response in crop and livestock production. *Pergamon Press*

ELSEN, J.M.; BIBE, B.; LANDAIS, E.; RICORDEN, G. 1986. Twenty remarks on economic evaluation of selection goals. vol 12, *Inc Proc 3rd World Congr. Genet. Appl. Livest Prod.* Lincoln, Nebraska pp 321-327.

*GIBSON, J. P.; KENNEDY, B. W. 1990. The use of constrained selection indexes in breeding for economic merit. *Theor. Appl. Genet.* 80: 801-805.

GIBSON, J. P. 1989. The effect of princing systems, economic weights, and population parameters on economic response to selection on milk components. *J. Dairy Sci.* 72. 3314-3326.

*GIBSON, J. P. 1989. Selection on the major components of milk: alternative methods of deriving economic weights. *J. Dairy Sci.* 72: 3176-3189.

GODDARD, M. E. 1983. Selection indices for non-linear profit functions. *Theor. Appl. Genet.* 64: 339-344.

HARRIS, D. L. 1970. Breeding for efficiency in livestock production: defining the economic objetives. *J. Anim. Sci.* 30: 360-365.

HARRIS, D.L.; STEWARD, T. S.; ARBOLEDA, C. R. 1984. Animal breeding programs: systematic approach to their desing. *Advances in Agricultural Tecnology* NC-8. Agricultural Research Service. U.S. Departament of Agriculture. February 1984.

HAZEL, L. N. 1943. The genetic basis for constructing selection indexes. *Genetics* 28: 476-490.

HILL, W. G. 1974. Prediction and evaluation of response to selection with overlapping generations. *Anim. Prod.* 18: 117-119.

HILL, W. G. 1971. Investment apparaisal for national breeding programmes. *Anim. Prod.* 13: 37-50.

*ITOH, Y.; YAMADA, Y. 1988. Linear selection indices for non-linear profit fuctions. *Theor. Appl. Genet.* 75: 553-560.

*JAMES, J. W. 1982. Economic aspects of developing breeding objetives: general considerations. *In: Future developments in the genetic improvement of animals.* Barker, J. S. F., Hammond, K., Mc Clintock, A. E. (Eds). Academic Press, Sydney, 107-108.

JAMES, J. W. 1980. Index selection for simultaneous improvement of several characters. *In: Proc. Xp XIV Inter. Cong. Genet.* Mir Publ. Moscow 1980. Vol 1, book 2: 221-229.

*McCLINTOCK, A. E.; CUNNINGHAM, E. P. 1974. Selection in dual purpose cattle populations: defining the breeding objetive. *Anim. Prod.* 18: 237-247.

MELTON, B. E.; HEADY, E. O.; WILLHAM, R. L. 1979. Estimation of economic values for selection indices. *Anim. Prod.* 28: 279-286.

MILLER, R. H.; PEARSON, R. E. 1979. Economic aspects of selection. *Anim. Breed. Abstracts* 47 (6): 281-290.

*MOAV, R. 1973. Economic evaluation of genetic differences. *In: Agricultural Genetics.* (R. Moav, ed). Wiley, New York. 319-352.

MOAV, R.; HILL, W. 1966. Specialised sire and dam lines. 4. Selection within lines. *Anim. Prod.* 8: 375-390.

MUÑOZ - LUNA, A. ; YADAV, S. B. S.; DEMPFLE, L. 1988. Derivation of econo mic weights for several traits for the Rubia-Gallega cattle en Spain. *J. Anim. Breed. Genet.* 105: 372-383.

*OLLIVIER, L.; GUEBLEZ, R.; WEEB, A. J.; Van der STEEN, H. A. M. 1990. Breeding goals for nationally and internationally operating pig breeding organisations. *Proc. 4th World Congr. Genet. Appl. Livest. Prod.* Edinburgh, vol XV, 383-394.

PASTERNAK, H.; WELLER, J.L., 1993. Optimum linear indices for non-linear profit functions. *Anim. Prod.* 55: 43-50.

PONZONI, R. W. 1985. Linear approximation of non-linear indices: an example with Australian Merino sheep. *Z. Tierzüchtg. Züchtgsbiol.* 102: 395-399.

PONZONI, R. W. 1988. The derivation of economic values combining income and expense in different ways: an example with Australian Merino sheep. *J. Anim. Breed. Genet.* 105: 143-153.

RØNNINGEN, K. 1971. Tables for estimating the loss in efficiency when selecting according to an index based on a false economic ratio between two traits. *Acta Agriculturæ Scandinavica* 21: 33-49.

SCHLOTE, W. 1977. Choix et pondération économique des caractères en sélection animale. *Ann. Génét. Sél. anim.* 9 (I): 63.-72.

SMITH, C.; JAMES, J. W.; BRASCAMP, E. W. 1986. On the derivation of economic weights in livestock improvement. *Anim. Prod.* 43: 545-551.

SMITH, C. 1983. Effects of changes in econimic weights on efficiency of index selection. *J Anim. Sci.* 56 (5): 1057-1064.

STRAIN, J. H.; NORDSKOG, A. W. 1962. Genetic aspect of the profit equation in a broiler enterprise. *Poultry Science* vol. XLI, 6: 1892-1902.

THOMPSON, R. 1980. A note on the selection of economic values for selection indices. *Anim. Prod.* 31: 115-117.

*VANDEPITTE, W. M.; HAZEL, L. N. 1977. The effect of errors in the economic weights on the accuracy of selection indexes. *Ann. Génét. Sél. anim.* 9 (I): 87-103.

*WOOLLIAMS, J.A.; MEUWISSEN, T.H.E. 1993. Decision rules and variance of response in breeding scheemes. *Anim. Prod.* 56:179-186.

Apéndice I

Voy a utilizar exactamente la notación de Smith et al.: P=Beneficios, R=Ingresos, C=Costes, y consideraremos nulos los ingresos en lugar de los beneficios, tal y como ocurre en el apéndice del artículo.

Consideremos

$$P = R - C$$

Si consideramos el beneficio obtenido por un sólo carácter **y**, considerando al resto constante, y teniendo en cuenta que hay un factor de escala **n** que también es variable,

$$dP = (\delta P/\delta y)\, dy + (\delta P/\delta n)\, dn = (\delta R/\delta y - \delta C/\delta y)\, dy + (\delta R/\delta n - \delta C/\delta n)\, dn \qquad (1)$$

Haciendo que los ingresos sean nulos, dR = 0, tenemos

$$dR = (\delta R/\delta y)\, dy + (\delta R/\delta n)\, dn = 0$$

$$(\delta R/\delta y)\, dy = -(\delta R/\delta n)\, dn \qquad (2)$$

Este resultado (2) aparece en el artículo sin ninguna justificación, como un acto "natural" derivado del cambio de escala (hay una diferencia en el signo, pero es irrelevante, puesto que el cambio de escala puede ser positivo o negativo). El caso es que no hay ninguna definición explícita de lo que es un factor de escala en el artículo de Smith et al., lo que no deja de sorprender, dado que toda la argumentación gira precisamente en torno al cambio de escala de la empresa. Por lo que yo deduzco, un factor de escala es un parámetro que afecta igualmente a ingresos y costes, lo que quiere decir que

$$R = f(n) \cdot g(y_1, y_2, \ldots, y_n)$$
$$C = f(n) \cdot h(y_1, y_2, \ldots, y_n)$$

donde $g(y_1, y_2, \ldots, y_n)$ y $h(y_1, y_2, \ldots, y_n)$ son funciones independientes de $f(n)$. Un ejemplo citado antes puede ser el número de animales de la granja. En cualquier caso, con esa definición se cumple que

$$(\delta R/\delta n) = f' \cdot g$$
$$(\delta C/\delta n) = f' \cdot h$$

$$(\delta R/\delta n) / R = (\delta C/\delta n) / C \qquad (3)$$

La fórmula (3) aparece en el texto de Smith et al. como la definición de factor de escala. Puede que sea cuestión de gustos o, como me hacía notar Marcos Rico, una forma habitual y precisa de definir ciertos factores en ecuaciones diferenciales. Yo no sé si la definición que propongo es demasiado restrictiva, pero me gustaría que si encontráis alguna otra solución más genérica que permitiera definir al factor de escala de forma explícita, me la comunicárais.

Sustituyendo (2) y (3) en (1), tenemos

$$dP = (\delta R/\delta y - \delta C/\delta y)\, dy + [-\delta R/\delta y - (C/R)(-\delta R/\delta y)]\, dy =$$
$$= [(C/R)\delta R/\delta y - \delta C/\delta y\,]\, dy$$

Por tanto, el peso económico de **y** es

$$dP/dy = (C/R)\delta R/\delta y - \delta C/\delta y \qquad (4)$$

Si no hubiéramos tenido en cuenta el factor de escala (es decir, si n=cte.)

$$dP/dy = dR/dy - dC/dy$$

que es la forma habitual de deducir pesos económicos.

El hecho de considerar dR=0 iguala las perspectivas de varios sectores como vimos al resolver la pregunta 5 (el ejemplo estaba puesto para dB=0), falta aclarar que los pesos económicos así calculados son iguales a los que se derivan de la función

$$P = R/C.$$

Procediendo de forma análoga,

$$dP = [\delta(R/C) / \delta y\,]\, dy + [\delta(R/C) / \delta n\,]\, dn$$

Ahora bien, el segundo término es nulo por la definición de factor de escala. Si no se acepta la definición que he dado antes, es de todas formas fácil de ver que es nulo:

$$[\delta(R/C) / \delta n\,]\, dn = (1/C)\,(\delta R/\delta n)\, dn - (R/C^2)\,(\delta C/\delta n)\, dn$$

sustituyendo (2) y (3),

$$[\delta(R/C) / \delta n\,]\, dn = (-1/C)(\delta R/\delta y)\, dy - (R/C^2)\,(C/R)\,(\delta R/\delta n)dn =$$
$$= (-1/C)(\delta R/\delta y)\, dy - (1/C)\,(-\delta R/\delta y)\, dy = 0$$

Por tanto,

$$dP = [\delta(R/C) / \delta y\ ]\ dy = (1/C)\ (\delta R/\delta y)\ dy\ -\ (R/C^2)\ (\delta C/\delta y)\ dy\ =$$

$$= (R/C^2)\ [(C/R)\delta R/\delta y\ -\ \delta C/\delta y]\ dy$$

El nuevo peso económico calculado a partir de P=R/C sólo difiere del anterior, calculado a partir de P=R-C, en el factor (constante para todos los caracteres) R/C^2 . Por tanto, los pesos económicos relativos son los mismos se use una función de beneficios u otra. Este aspecto también está un tanto oculto en el artículo de Smith et al., y no figura en él esta deducción.

Según se tome como condición dP=0, dR=0 ó dC=0, la constante anterior varía, pero siempre salen los pesos proporcionales a $\delta(R/C) / \delta y$.

En nuestro ejemplito en el texto, en un caso es

$$dP/dw = (C/R)\delta R/\delta w - \delta C/\delta w = (nc/nwv) \cdot nv - 0 = nc/w.$$

y en el otro

$$dP/dw = dR/dw - dC/dw = nv - 0 = nv.$$

Apéndice II

Se trata de derivar

$$w_i = a_1 / a'_i = (\delta Q/\delta x_i) / (\delta Q/\delta x_p)$$

donde

$$Q = c \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_n$$

y donde

$$p_i = \int_{-\infty}^{x_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x_i - m_i)^2}{s^2} \cdot v_i^2} dx_i$$

Lo único raro de la expresión de $\mathbf{p_i}$ el término $\mathbf{v_i}$ (que en la nomenclatura del artículo de de Vries es e_i pero que lo cambio por no liar con el número e). Es un problema meramente de unidades de medida. $\mathbf{v_i}$ es el peso económico del carácter **i** calculado por cualquiera de los métodos expuestos antes. De Vries define el nivel de aceptación de un carácter ponderado por su peso económico (o el peso económico ponderado por su nivel de aceptación, que ambas interpretaciones serían válidas) y tipificado:

$$t_i = v_i \ (x_i - m_i) / s$$

Observad que si el nivel de aceptación del carácter es bajo, el peso $\mathbf{t_i}$ tenderá a ser grande en valor absoluto y negativo, mientras que si el carácter se sitúa en la media de aceptabilidad, $\mathbf{t_i}$ será cero.

Utilizando este nivel de aceptación $\mathbf{t_i}$, la expresión de $\mathbf{p_i}$ es

$$p_i = \int_{-\infty}^{t_i} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2} \cdot t_i^2} dt_i = \int_{-\infty}^{t_i} z_i dt_i$$

con lo que

$$\delta Q/\delta x_i = (\delta Q/\delta p_i) \ (\delta p_i/\delta t_i) \ (\delta t_i/\delta x_i)$$

$$\delta Q/\delta p_i = c \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_{i-1} \cdot p_{i+1} \cdot \ldots \cdot p_n = c \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_n \ (1/p_i) = Q/p_i$$

$$\delta p_i/\delta t_i = z_i$$

$\delta t_i/\delta x_i = v_i/s$

Así sale la fórmula (4) del artículo de de Vries:

$$\delta Q/\delta x_i = (v_i/s)\ (z_i/p_i)\ Q$$

Asumiremos que el precio del producto se encuentra en el nivel medio de aceptación ($x_P = m_P$), y que el "peso económico" v_P del precio del producto es **-1**, ya que un aumento de una unidad monetaria en el precio de compra de un animal reproductor x_P implica una reducción de una unidad monetaria en el beneficio del granjero.

Si $x_P = m_P$ entonces $p_P = 0.5$, $t_P = 0$,

$$\frac{z_P}{p_P} = \frac{1/\sqrt{2\pi}}{1/2} = \sqrt{\frac{2}{\pi}} = (2/\pi)^{1/2}$$

$$\delta Q/\delta x_P = (-1_i/s)\ (2/\pi)^{1/2}\ Q$$

Con lo que, finalmente el peso económico es

$$w_i = -\ (\delta Q/\delta x_i)\ /\ (\delta Q/\delta x_P) = v_i\ (z_i/p_i)\ (\pi/2)^{1/2}$$

Cuando el nivel de aceptación de un carácter coincide con el nivel de aceptación medio ($x_i = m_i$), entonces $t_i = 0$, $z_i/p_i = (2/\pi)^{1/2}$, y coinciden w_i con v_i .

Si introducimos un factor de compensación entre caracteres **c**, de forma que la falta de aceptación de un carácter se compensa con el exceso de aceptación de otros, la media de aceptación de un carácter pasa a ser, en unidades monetarias

$$m_i \cdot v_i = m_i^c \cdot v_i - c \cdot \sum_{i \neq j} (x_j - m_j^c) v_j$$

donde m^c_i es la media de aceptación del carácter **i** por los competidores. En la situación actual no hay grandes diferencias entre empresas, de forma que las ventajas y los inconvenientes de cada un ade ellas suelen estar compensados a la hora de proponer un producto a la venta razonable. Esto significa que

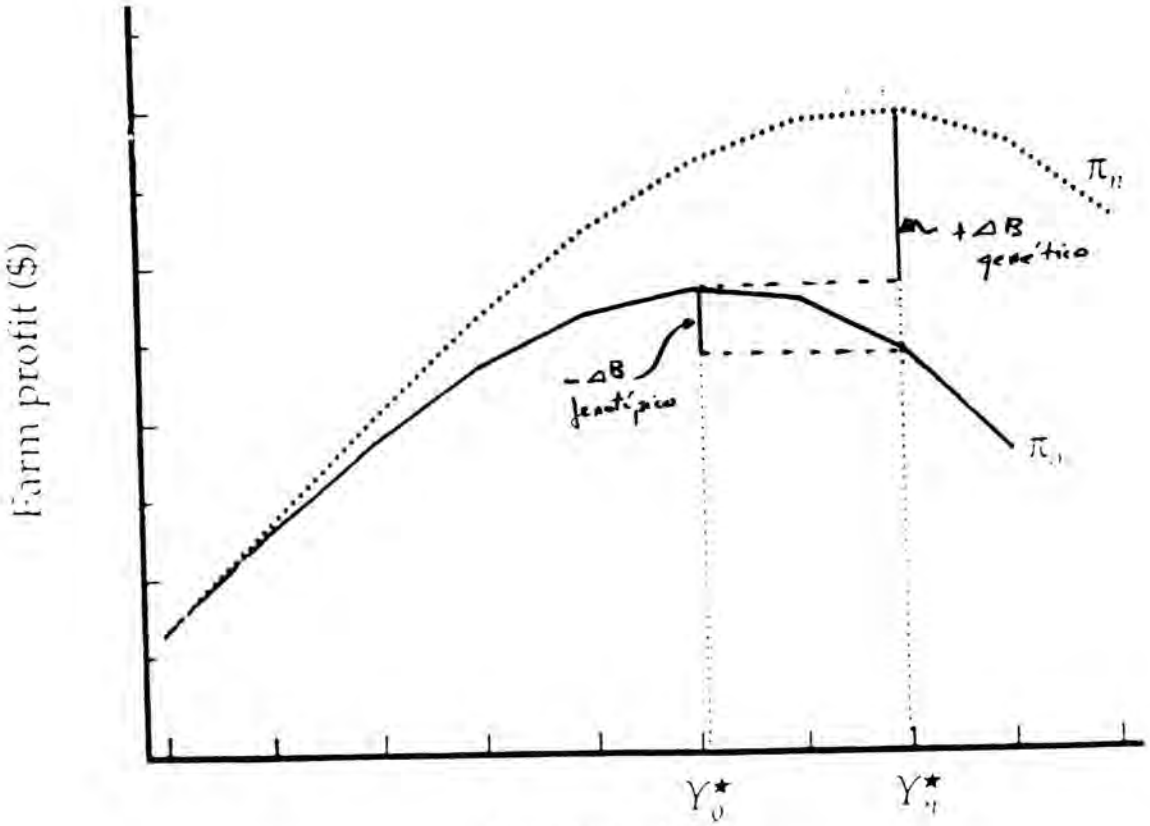
$$\sum_{i=1}^{i=n} (x_i - m_i^c) \cdot v_i = 0 \rightarrow \sum_{i \neq j} (x_j - m_j^c) \cdot v_j = -(x_i - m_i^c)$$

con lo que

$$m_i = m^c_i + (1/v_i)\, c\, (x_i - m^c_i)$$

y ya pueden calcularse $\mathbf{z}_i$ y $\mathbf{p}_i$ para el peso económico $\mathbf{w}_i$.

FIGURA 1



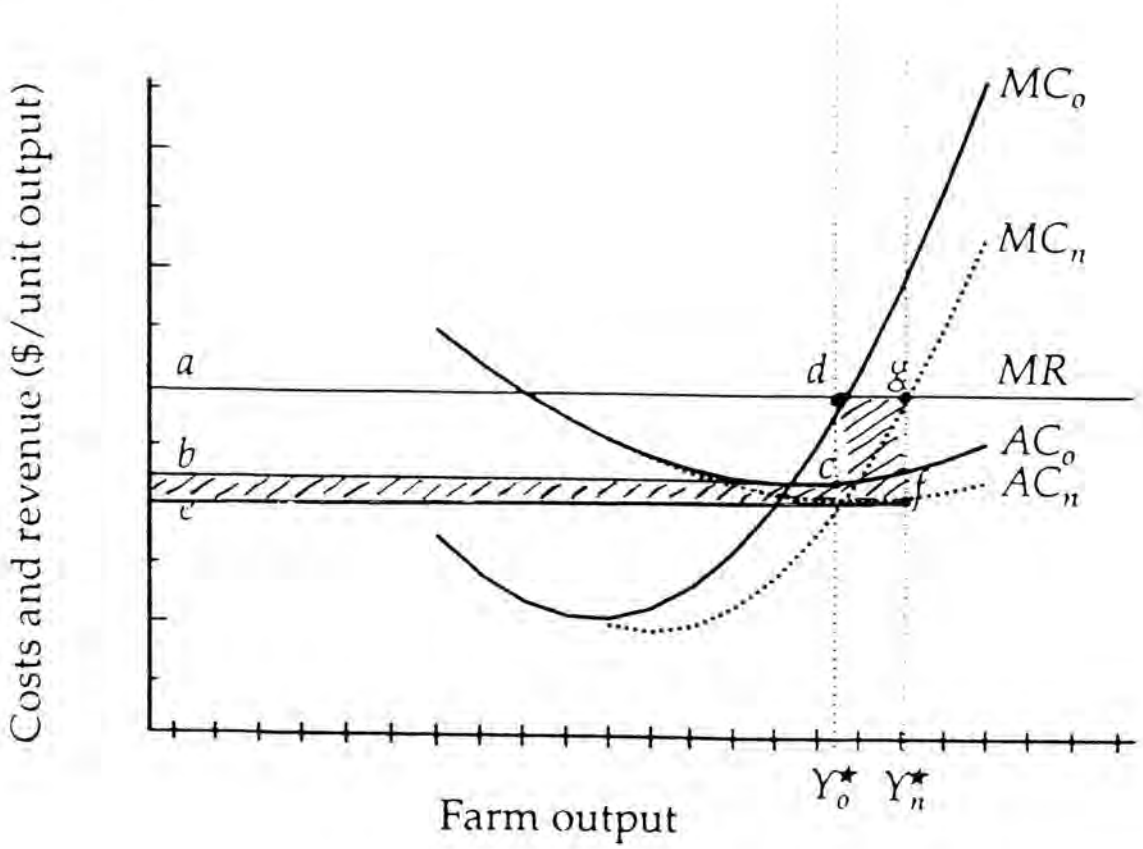
Change in a farm's profit function with genetic improvement.

La figura está tomada de Amer y Fox (1992) en un contexto diferente, pero es útil para ilustrar este punto.

Supongamos que la función de benficios es la indicada por la línea sólida. Supongamos que la empresa produce una cantidad de producto Y^*_0 ; esto es, se encuentra en el óptimo de producción. Cualquier incremento *fenotípico* del carácter (pasar a producir Y^*_n , por ejemplo) le conducirá a pérdidas.

Supongamos que la mejora genética cambia la función de producción y pasa ahora a ser la indicada en la línea de puntos. Un incremento *genético* en la producción le da lugar ahora a ganancias.

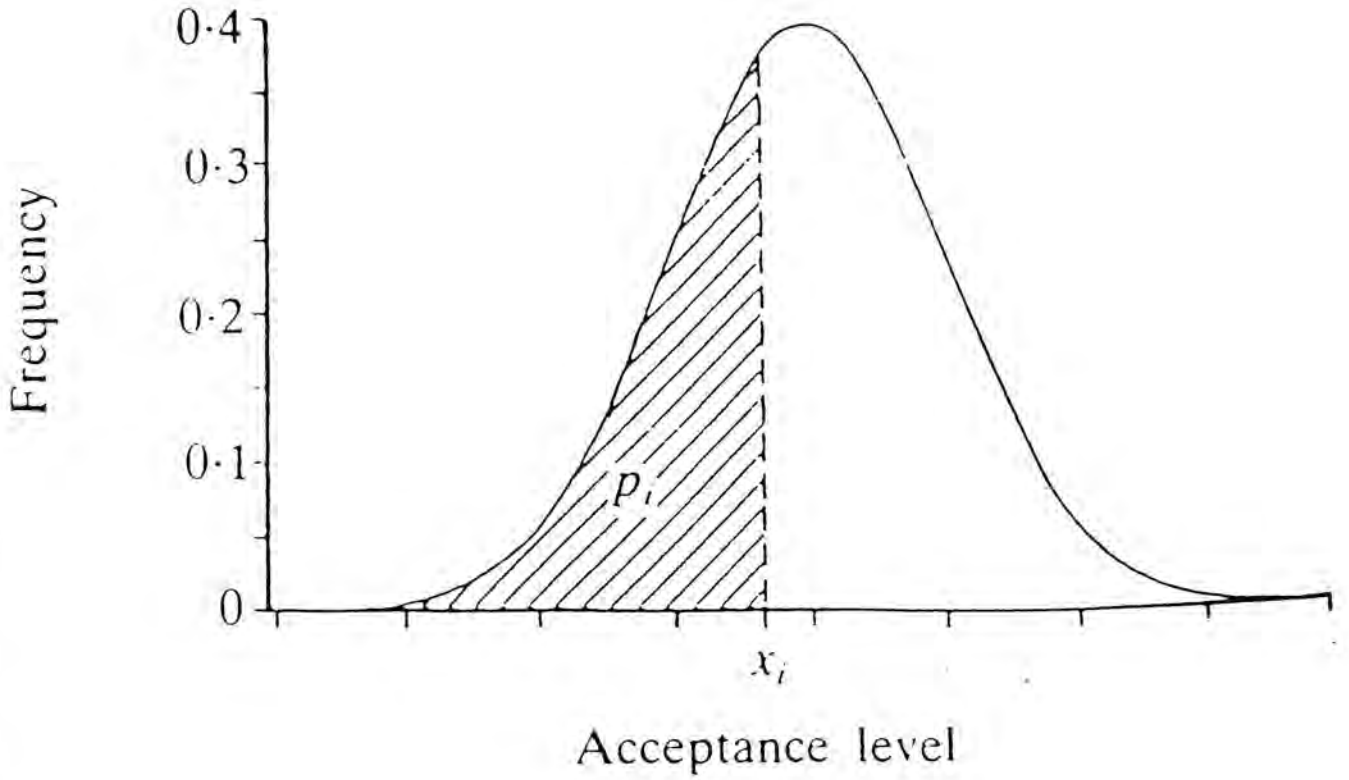
FIGURA 2



La figura se explica con detalle en el texto. Se presentan en línea sólida la situación de partida, y en línea punteada la situación después de que actúe la selección. La curva de costes marginales inicial (MC_0) corta a la recta de ingresos marginales (MR) en el punto **d**, por tanto la situación óptima de partida es aquélla en la que la empresa produce una cantidad $\mathbf{Y^*_0}$. Tras actuar la mejora genética, se obtiene una nueva curva de costes marginales (MC_n) que corta a la de ingresos marginales en el punto **g**, dando lugar a un nuevo nivel óptimo de producción, $\mathbf{Y^*_n}$. Mientras tanto, los costes medios de la empresa han pasado de la curva (AC_0) a (AC_n)

El beneficio inicial de la empresa era el área **abcd** de la figura. Tras la selección el beneficio es el área **aefg**. El incremento de beneficio corresponde a las dos áreas rayadas de la figura.

FIGURA 3



II. PARTE APLICADA: BALANCE ECONOMICO DE UNA EXPLOTACION DE GANADO PORCINO

Antonio Muñoz
Facultad de Veterinaria. Universidad de Murcia

Se incluyen en este apartado todos los vectores económicos que condicionan el capítulo de ingresos y gastos, y quedan reflejados en:

Ingresos (I) = Ingresos por la venta de lechones (IVL) + Ingresos por la venta de cerdos cebados (IVC) + Ingresos por la venta de cerdas de desecho (IVD) + Ingresos por la venta de purines.

Gastos (G) = Gastos generados por cerda (GC) + Gastos generados por la descendencia (GD).

Se tratan a continuación cada uno de los vectores económicos que conforman los ingresos y gastos desglosados en las diferentes coordenadas que contribuyen a su formación.

II.1.- INGRESOS

Dentro de la explotación porcina que nos ocupa se pueden recibir ingresos como consecuencia de la actividad que se desarrolla, por medio de cuatro vías: por la venta de lechones, por la venta de cerdos cebados, por la venta de reproductores de desvieje, y por la venta de purines.

1.1.- Ingresos por la venta de lechones

Los ingresos por la venta de lechones vienen determinados por la combinación de una serie de parámetros genéticos y ambientales que podemos expresar de la siguiente manera:

$$IVL = LVP \times NP \times (1-S_1) \times (1-S_2) \times W_2 \times PL \times PRLV \quad (1)$$

donde:

LVP = lechones vivos por parto
NP = número de partos
S_1 = grado de mortalidad del lechón (nac.-dest.)
S_2 = grado de mortalidad del lechón (dest.-inic.cebo)
W_2 = peso al inicio del cebo
PL = precio del lechón (ptas.)
PRLV = porcentaje de lechones vendidos

1.2.- Ingresos por la venta de cerdo cebados

Nos vienen determinados por los parámetros anteriores, más aquellos que desarrollan su influencia en la etapa 3.

$$IVC = LVP \times NP \times (1-S_1) \times (1-S_2) \times (1-S_3) \times W_3 \times CC \times P_3 \times (1-PRLV) \quad (2)$$

donde:

S_3 = grado de mortalidad del cerdo en cebo
W_3 = peso del cerdo al final del cebo (Kg.)
CC = coeficiente de clasificación de canales
P_3 = precio por Kg. de peso vivo de cerdo selecto (ptas.)

1.3.- Ingresos por la venta de reproductoras de desecho (IVD)

$$IVD = (F - S_4) \times W_4 \times P_4 \quad (3)$$

donde:

F = tasa de reposición de reproductoras
S_4 = grado de mortalidad anual de una cerda reproductora.
W_4 = peso de una cerda reproductora de desecho (Kg.)
P_4 = precio por Kg.de peso vivo de una cerda reproductora (ptas.)

1.4.- Ingresos por la venta de purines (IVP)

$$IVP = PRP \times POP \times [\ PPR + (TP_2 \times PPL) - (TP_3 \times PPC)\ ] \quad (3)$$

siendo:

TP_2 = LVP x NP x $(1-S_1)$ x $(1-S_2/2)$ (nº lechones etapa 2)
TP_3 = LVP x NP x $(1-S_1)$ x $(1-S_2)$ x $(1-S_3/2)$ (nº cerdos en la etapa 3)

donde:

PRP = precio por Kg. de estiércol líquido (ptas.)
POP = grado de aprovechamiento de los purines
PPR = producción de purines anual por reproductora (Kg.)
TP_2 = número de lechones en la etapa 2
TP_3 = número de lechones en la etapa 3.
PPC = producción de purines por cerdo (cebo) (kg.)

II.2.- GASTOS TOTALES (G)

Se han estudiado los gastos refiriéndolos a los gastos generados por cerda (GC), y a los gastos generados por la descendencia en cada una de las etapas.

2.1.-Gastos generados por cerda (GC)

Los gastos generados por cerda se han dividido en cuatro capítulos:

A/ GASTOS GENERALES (GG)

$$GG = GBA + (GEPP \times NP) + (GELD \times NP \times DML) \quad (5)$$

donde:

GG = gastos generales
GBA = gastos básicos anuales por cerda (ptas.)
GEPP = gastos extras no alimentarios por parto (ptas.)
GELD = gastos extras por lechón destetado (ptas.)
DML = duración media de la lactación (días)

B/ GASTOS DE INSEMINACION (GI)

$$GI = NIP \times NP \times A \quad (6)$$

donde:

GI = gastos de inseminación (ptas.)
NIP = número de inseminaciones por parto
A = precio de una dosis de semen (ptas.)

C/ GASTOS DE REPOSICION (GR)

$$GR = (F - PRR) / (1 - S_4) \quad (7)$$

donde:

GR = gastos de reposición (ptas.)
PRR = precio de una reproductora de reposición (ptas.)

D/ GASTOS DE ALMENTACION

$$GA = [RBA + ((LVP/1-S_0) \times NP \times REPLG) + (LVP \times (1-S_1) \times NP \times DML \times REPLD)] \times PRM \quad (8)$$

donde:

GA = gastos de alimentación (ptas.)
RBA = requerimientos básicos anuales (Kg.)
S_0 = grado de mortalidad embrionaria
REPLG = requerimientos extras para desarrollar un lechón nacido vivo por parto (Kg.)
REPLD = requerimientos extras durante la lactación por lechón destetado (Kg.)
PRM = precio por Kg. de pienso para madres (ptas.)

2.2.- Gastos generados por la descendencia

Estos gastos se han dividido en tres apartados según la etapa en que han sido generados.

A/ GASTOS GENERADOS POR LA DESCENDENCIA EN LA ETAPA 1.-

$$GDE_1 = LVP \times NP \times (1-S_1) \times [GELD + (CD_1 \times A_1 \times PR^1)] \quad (9)$$

$$\text{Siendo:} \quad A_1 = \frac{W_1 - W_0}{G_1}$$

donde:

GDE_1 = gastos generados por la descendencia en la etapa 1 (ptas.)
GELD = gastos extras por lechón destetado (ptas.)
CD_1 = requerimientos diarios de pienso en la etapa 1 (Kg.)
G_1 = ganancia media diaria en peso en la etapa 1 (g.)
PR_1 = precio por Kg. de pienso de iniciación de lechones (ptas.)

B/ GASTOS GENERADOS POR LA DESCENDENCIA EN LA ETAPA 2.-

$$GDE_2 = LVP \times NP \times (1-S_1) \times (1-S_2/2) \times [GDNA_2 + (CD_2 \times PR_2)] \times A_2 \quad (10)$$

$$\text{Siendo:} \quad A_2 = \frac{W_2 - W_1}{G_2}$$

donde:

GDE_2 = gastos generados por la descendencia en la etapa 2 (ptas.)
$GDNA_2$ = gastos diarios no alimentarios en la etapa 2 (ptas.)
CD_2 = requerimientos diarios de pienso en la etapa 2 (Kg.)
PR_2 = precio por Kg. de pienso para la recria de los lechones (ptas.)
G_2 = ganancia media diaria de peso en la etapa 2 (g.)

C/ GASTOS GENERADOS POR LA DESCENDENCIA EN LA ETAPA 3.-

$$GDE_3 = LVP \times NP \times (1-S_1) \times [(1-S_2) \times [GECC + (1-S_3/2) \times (GDNA_3 + (CD_3 \times PR_3)] \times A_3] \times (1 - PRLV) \quad (11)$$

$$\text{Siendo: } A_3 = \frac{W_3 - W_2}{G_3}$$

donde:

GDE_3 = gastos generados por la descendencia en la etapa 3 (ptas.)
GECC = gastos extras por cerdo cebado (ptas.)
$GDNA_3$ = gastos diarios no alimentarios en la etapa 3 (ptas.)
CD_3 = requerimientos diarios de pienso en la etapa 3 (Kg.)
PR_3 = precio por Kg. de pienso para cebo (ptas.)
G_3 = ganancia media diaria de pienso en la etapa 3 (g.)

II.3.- CARACTERES INCLUIDOS EN LA ECUACION DE BENEFICIOS

Los caracteres biológicos de importancia económica sobre el capítulo de ingresos y gastos, debido a la influencia que tienen sobre la producción porcina, y que a su vez son susceptibles de mejora genética son:

- LVP.- Lehones vivos por parto
- NP .- Número de partosç
- CC .- Coeficiente de clasificación de canales
- G_1 .- Ganancia media diaria en la etapa 1
- G_2 .- Ganancia media diaria en la etapa 2
- G_3 .- Ganancia media diaria en la etapa 3
- CD_2.- Consumo diario de pienso en etapa 2
- CD_3.- Consumo diario de pienso en etapa 3

El coeficiente de clasificación de canales (CC) se define mediante la siguiente ecuación:

$$CC = PAE + PA_1 \times PC_1 / P_3 + PA_2 \times PC_2 / P_3$$

Siendo:

PAE: % de descendientes que obtuvieron la clasificación de selecto.
PA_1: % de descendientes que obtuvieron la clasificación de normal.
PA_2: % de descendientes que obtuvieron la clasificación de graso.
PC_1: Precio del Kg. de p.v. del cerdo normal.
PC_2: Precio del Kg. de p.v. del cerdo graso.
PC_3: Precio del Kg. de p.v. del cerdo selecto.

II.4.- CALCULO DE PESOS ECONOMICOS

Se han calculado los pesos económicos de las diferentes características biológicas mediante el desarrollo propuesto por Harris (1970), calculando el peso económico mediante la derivada parcial de cada carácter manteniendo el resto de los caracteres al nivel de su valor medio.

Igualmente se han desarrollado las modificaciones propuestas por Smith et al.,(1986) a fin de encontrar unos pesos económicos relativos que fueran independientes de la base de evaluación tal como propone Muñoz et al. (1991)

Las fórmulas generales aplicadas para cada una de las alternativas propuestas, partiendo de la ecuación de beneficios general:

$$B = I - G$$

Los pesos económicos se obtienen a partir de la secuencia propuesta por estos autores:

1°) Peso económico del carácter X_i, cuando un pequeño cambio genético en X_i provoca un cambio en el beneficio:

$$\frac{\partial B_1}{\partial X_i} = \frac{\partial I}{\partial X_i} - \frac{\partial G}{\partial X_i}$$

2°) Peso económico del carácter X_i, cuando la empresa se reconvierte en una cantidad n, para variar los ingresos:

$$\frac{\partial B_2}{\partial X_i} = \frac{\partial I}{\partial X_i} - \frac{G}{I} \cdot \frac{\partial I}{\partial X_i}$$

3°) Pesos económicos netos del carácter X_i, es decir, el beneficio que genera el cambio genético menos la reconversión de la empresa:

$$\frac{\partial B_3}{\partial X_i} = \frac{G}{I} \cdot \frac{\partial I}{\partial X_i} - \frac{\partial G}{\partial X_i}$$

Siendo : $\Delta B_3 = \Delta B_1 - \Delta B_2$

4°) Factor de proporcionalidad que independiza el cálculo de los pesos económicos y que está siempre presente independientemente de la base de evaluación.

$$\frac{\partial \phi}{\partial X_i} \qquad (\phi = I / G)$$

Igualmente se estimaron los diferentes pesos económicos obtenidos en función del progreso económico y de las desviaciones típicas fenotípicas y genéticas aditivas (Muñoz et al.1988 a.); para lo que se utilizó una intensidad de selección de i=1 y los siguientes parámetros genéticos:

	LVP	NP	CC	G_1	G_2	G_3	CD_2	CD_3
Heredabilidad	0.12	0.15	0.14	0.35	0.3	0.3	0.3	0.25
σp	1.39	0.17	0.14	0.021	0.039	0.071	0.023	0.068

Se ha realizado un análisis de sensibilidad del modelo diseñado en el que todos los gastos fijos se multiplicaron por una secuencia que iba desde un aumento del 100 % hasta una reducción al 20 % del valor utilizado en el modelo.

Los gastos fijos sometidos a esta variación fueron:

- Gastos básicos anuales por cerda (GBA)
- Gastos extras no alimentarios por parto (GEPP)
- Gastos diarios extras durante la lactancia (GEDL)
- Gastos extras por lechón destetado (GELD)
- Gastos diarios no alimentarios en la etapa 2 ($GDNA_2$)
- Gastos diarios no alimentarios en la etapa 3 ($GDNA_3$)
- Gastos extras por cerdo cebado (GECC)

El sistema de cálculo y ecuaciones de las derivadas para cada carácter se ha realizado como sigue:

1.- Ecuacion B = I - G

$$B = IVL + IVC + IVD + IVP - GG - GI - GR - GA - GDE_1 - GDE_2 - GDE_3$$

Derivadas parciales para el carácter LVP

$$\partial B = \partial IVL/\partial LVP + \partial IVC/\partial LVP + \partial IVD/\partial LVP + \partial IVP/\partial LVP - \partial GG/\partial LVP - \partial GI/\partial LVP - \partial GR/\partial LVP - \partial GA/\partial LVP - \partial GDE_1/\partial LVP - \partial GDE_2/\partial LVP - \partial GDE_3/\partial LVP$$

$$\partial IVL/\partial LVP = NP \times (1-S_1) \times (1-S_2) \times W_2 \times PL \times PRLV$$

$$\partial IVC/\partial LVP = NP \times (1-S_1) \times (1-S_2) \times (1-S_3) \times W_3 \times CC \times P_3 \times (1-PRLV)$$

$$\partial IVP/\partial LVP = PRP \times POP \times [PPR + (NP \times (1-S_1) \times (1-S_2/2) \times PPL) - (NP \times (1-S_1) \times (1-S_2) \times (1-S_3/2))] \times PPC \times (1 - PRLV)$$

$$\partial GA/\partial LVP = [RBA + ((1/1-S_0) \times NP \times REPLG) + (1-S_1) \times NP \times DML \times REPLD)] \times PRM$$

$$\partial GDE_1/\partial LVP = NP \times (1-S_1) \times [GELD + (CD_1 \times A_1 \times PR_1)]$$

$$\partial GDE_2/\partial LVP = NP \times (1-S_1) \times (1-S_2/2) \times [GDNA_2 + (CD_2 \times PR_2)] \times A_2$$

$$\partial GDE_3/\partial LVP = NP \times (1-S_1) \times [(1-S_2) \times [GECC + (1-S_3/2) \times (GDNA_3 + (CD_3 \times PR_3)] \times A_3] \times (1 - PRLV)$$

Las derivadas de GG, GI, GR e IVD respecto al carácter LVP son iguales a cero.

Derivadas parciales para el carácter NP

Partimos de la misma ecuación de beneficios, pero en este caso derivando para el carácter NP.

$$\partial \mathbf{IVL}/\partial \mathbf{NP} = LVP \times (1-S_1) \times (1-S_2) \times W_2 \times PL \times PRLV$$

$$\partial \mathbf{IVC}/\partial \mathbf{NP} = LVP \times (1-S_1) \times (1-S_2) \times (1-S_3) \times W_3 \times CC \times P_3 \times (1-PRLV)$$

$$\partial\ \mathbf{IVP}/\partial\ \mathbf{NP} = PRP \times POP\ X \times [PPR + (LVP \times (1-S_1) \times (1-S_2/2) \times PPL) - (LVP \times (1-S_1) \times (1-S_2) \times (1-S_3/2))] \times PPC \times (1 - PRLV)$$

$$\partial\ \mathbf{GG}/\partial\ \mathbf{NP} = GBA + GEPP + (GELD \times DML)$$

$$\partial\ \mathbf{GI}/\partial\ \mathbf{NP} = NIP\ X\ A$$

$$\partial\ \mathbf{GA}/\partial\ \mathbf{NP} = [RBA + ((LVP/1-S_0) \times REPLG) + (LVP \times (1-S_1) \times DML \times REPLD)] \times PRM$$

$$\partial\ \mathbf{GDE_1}/\partial NP = LVP \times (1-S_1) \times [GELD + (CD_1 \times A_1 \times PR_1)]$$

$$\partial\ \mathbf{GDE_2}/\partial NP = LVP \times (1-S_1) \times (1-S_2/2) \times [GDNA_2 + (CD_2 \times PR_2)] \times A_2$$

$$\partial\ \mathbf{GDE_3}/\partial NP = LVP \times (1-S_1) \times [(1-S_2) \times [GECC + (1-S_3/2) \times (GDNA_3 + (CD_3\ X\ PR_3)] \times A_3] \times (1 - PRLV)$$

Las derivadas de IVD y GR son iguales a cero.

Derivadas parciales para CC

Las derivadas parciales de IVL, IVD, IVP, GG, GI, GR, GA, GDE_1, GDE_2 y GDE^3 son iguales a cero.

$$\partial \mathbf{IVC}/\partial \mathbf{CC} = NP \times LVP \times (1-S_1) \times (1-S_2) \times (1-S_3) \times W_3 \times P_3 \times (1-PRLV)$$

Derivadas parciales para G_1

Son todas iguales a cero con excepción de:

$$\partial \mathbf{GDE_1}/\partial \mathbf{G_1} = NP \times LVP \times (1-S_1) \times [GELD + (CD_1 \times (W_1 - W_0/ G_1^2 \times PR_1)]$$

Derivadas parciales para G_2

Son todas iguales a cero con excepción de:

$$\partial \mathbf{GDE_2}/\partial \mathbf{G_2} = NP \times LVP \times (1-S_1) \times (1-S_2/2) \times [GDNA_2 + (CD_2 \times PR_2)] \times (W_2 - W_1)/ G^2$$

Derivadas parciales para G_3

Son todas iguales a cero con excepción de:

$$\partial \mathbf{GDE_3}/\partial \mathbf{G_3} = NP \times LVP \times (1-S_1) \times [(1-S_2) \times [GECC + (1- S_3/2) \times (GDNA_3 + (CD_3 \times PR_3)] \times (W_3 - W_2)/ G_3] \times (1 - PRLV)$$

2.- ECUACION B= I - G/I x I

$$\partial B/\partial X_i = \partial I/\partial X_i - G/I \times \partial I/\partial X_i$$

Aplicando las derivadas parciales para esta ecuación vamos a obtener los pesos económicos por reconversión de la empresa, obteniendo los mismos resultados para el término $\partial I/\partial X_i$ que hemos hallado en 6.1., pero multiplicados por el coeficiente G/I.

3.- ECUACION B= G/I x I - G

$$\partial B/\partial X_i = G/I \times \partial I/\partial X_i - \partial G/\partial X_i$$

Se obtienen los mismos valores que en 6.1. para las derivadas de $\partial I/\partial X_i$ y $\partial G/\partial X_i$ con cada uno de los caracteres.

En esta ecuación se obtienen los pesos económicos que proporciona la mejora genética menos la reconversión de la empresa.

4.- ECUACION $\Phi = I/G$

Derivadas parciales para el caracter LVP

$$\partial\Phi/\partial LVP = \frac{\left(\frac{\partial IVL}{\partial LVP} + \frac{\partial IVC}{\partial LVP} + \frac{\partial IVD}{\partial LVP} + \frac{\partial IVP}{\partial LVP}\right) \times (GT)}{G^2} -$$

$$- \frac{\left(\frac{\partial GC}{\partial LVP} + \frac{\partial GDE_1}{\partial LVP} + \frac{\partial GDE_2}{\partial LVP} + \frac{\partial GDE_3}{\partial LVP}\right) \times (IT)}{G^2}$$

Siendo: IT = Ingresos totales
GT = Gastos totales
GC = GG + GI + GR + GA

Las derivadas parciales respecto a LVP de IVD, GG, GI, GR son iguales a cero. Por tanto:

$$\partial\Phi/\partial LVP = (NP \times (1-S_1) \times (1-S_2) \times W_2 \times PL \times PRLV + NP \times (1-S_1) \times (1-S_2) \times (1-S_3) \times W_3 \times CC \times P_3 \times (1-PRLV) + PRP \times POP \times [PPR + (NP \times (1-S_1) \times (1-S_2/2) \times PPL) - (NP \times (1-S_1) \times (1-S_2) \times (1-S_3/2))] \times PPC \times (1 - PRLV)) \times GT - ([RBA + ((1/1-S_0) \times NP \times REPLG) + (1-S_1) \times NP \times DML \times REPLD)] \times PRM - NP \times (1-S_1) \times [GELD + (CD_1 \times A_1 \times PR^1)] - NP \times (1-S_1) \times (1-S_2/2) \times [GDNA_2 + (CD_2 \times PR_2)] \times A_2 - NP \times (1-S_1) \times [(1-S_2) \times [GECC + (1-S_3/2) \times (GDNA_3 + (CD_3 \times PR_3)] \times A_3] \times (1 - PRLV)) \times IT / G^2 .$$

Derivadas parciales para el carácter NP

$$\partial\Phi/\partial NP = \frac{\left(\frac{\partial IVL}{\partial NP} + \frac{\partial IVC}{\partial NP} + \frac{\partial IVD}{\partial NP} + \frac{\partial IVP}{\partial NP}\right) \times (GT)}{G^2} -$$

$$- \frac{\left(\frac{\partial GC}{\partial NP} + \frac{\partial GDE_1}{\partial NP} + \frac{\partial GDE_2}{\partial NP} + \frac{\partial GDE_3}{\partial NP}\right) \times (IT)}{G^2}$$

Las derivadas parciales respecto a NP sobre los caracteres IVD, GR son iguales a cero. Por tanto:

$\partial\Phi/\partial NP$ = (LVP x $(1-S_1)$ x $(1-S_2)$ x W_2 x PL x PRLV + LVP x $(1-S_1)$ x $(1-S_2)$ x $(1-S_3)$ x W_3 x CC x P_3 x (1-PRLV) + PRP x POP X x [PPR + (LVP x $(1-S_1)$ x $(1-S_2/2)$ x PPL) - (LVP x $(1-S_1)$ x $(1-S_2)$ x $(1-S_3/2)$)] x PPC x (1 - PRLV)) x GT - (GBA + GEPP + (GELD x DML) - NIP X A - [RBA + ((LVP/$1-S_0$) x REPLG) + (LVP x $(1-S_1)$ x DML x REPLD)] x PRM - LVP x $(1-S_1)$ x [GELD + (CD_1 x A_1 x PR^1)] - LVP x $(1-S_1)$ x $(1- S_2/2)$ x [$GDNA_2$ + (CD_2 x PR_2)] x A_2 - LVP x $(1-S_1)$ x [$(1-S_2)$ x [GECC + $(1-S_3/2)$ x ($GDNA_3$ + (CD_3 X PR_3)] x A_3] x (1 - PRLV)) x IT / G^2

Derivadas parciales para el carácter CC

$$\partial\Phi/\partial CC = \frac{\left(\frac{\partial IVL}{\partial CC} + \frac{\partial IVC}{\partial CC} + \frac{\partial IVD}{\partial CC} + \frac{\partial IVP}{\partial CC}\right) \times (GT)}{G^2} -$$

$$- \frac{\left(\frac{\partial GC}{\partial CC} + \frac{\partial GDE_1}{\partial CC} + \frac{\partial GDE_2}{\partial CC} + \frac{\partial GDE_3}{\partial CC}\right) \times (IT)}{G^2}$$

Las derivadas parciales respecto a CC de los caracteres IVL, IVD, IVP, GG, GI, GR, GA, GDE_1, GDE_2 y GDE_3 son iguales a cero, con lo que el único término que influye es IVC. Por tanto:

$$\partial\Phi/\partial CC = LVP \times NP \times (1-S_1) \times (1-S_2) \times (1-S_3) \times W_3 \times CC \times P_3 \times (1-PRLV) \times GT / G^2$$

Derivadas parciales para el carácter G1

$$\partial\Phi/\partial G_1 = \frac{\left(\frac{\partial IVL}{\partial G_1} + \frac{\partial IVC}{\partial G_1} + \frac{\partial IVD}{\partial G_1} + \frac{\partial IVP}{\partial G_1}\right) \times (GT)}{G^2} -$$

$$- \frac{\left(\frac{\partial GC}{\partial G_1} + \frac{\partial GDE_1}{\partial G_1} + \frac{\partial GDE_2}{\partial G_1} + \frac{\partial GDE_3}{\partial G_1}\right) \times (IT)}{G^2}$$

Las derivadas parciales respecto a G_1 de los caracteres IVL, IVC, IVD, IVP, GDE_2 y GDE_3 son iguales a cero. Por tanto:

$$\partial\Phi/\partial G_1 = - NP \times LVP \times (1-S_1) \times [GELD + (CD_1 \times (W_1 - W_0 / G_1^2 \times PR_1)] \times IT / G^2$$

Derivadas parciales para el carácter G2

$$\partial\Phi/\partial G_2 = \frac{\left(\frac{\partial IVL}{\partial G_2} + \frac{\partial IVC}{\partial G_2} + \frac{\partial IVD}{\partial G_2} + \frac{\partial IVP}{\partial G_2}\right) \times (GT)}{G^2} -$$

$$- \frac{\left(\frac{\partial GC}{\partial G_2} + \frac{\partial GDE_1}{\partial G_2} + \frac{\partial GDE_2}{\partial G_2} + \frac{\partial GDE_3}{\partial G_2}\right) \times (IT)}{G^2}$$

Las derivadas parciales con respecto a G_2 de los caracteres IVL, IVC, IVD, IVP, GDE_1 y GDE_3 son iguales a cero. Por tanto:

$$\partial\Phi/\partial G_2 = - NP \times LVP \times (1-S_1) \times (1-S_2/2) \times [GDNA_2 + (CD_2 \times PR_2)] \times (W_2 - W_1)/ G_2 \times IT / G^2$$

Derivadas parciales para el carácter G3

$$\partial\Phi/\partial G_3 = \frac{\left(\frac{\partial IVL}{\partial G_3} + \frac{\partial IVC}{\partial G_3} + \frac{\partial IVD}{\partial G_3} + \frac{\partial IVP}{\partial G_3}\right) \times (GT)}{G^2} -$$

$$- \frac{\left(\frac{\partial GC}{\partial G_3} + \frac{\partial GDE_1}{\partial G_3} + \frac{\partial GDE_2}{\partial G_3} + \frac{\partial GDE_3}{\partial G_3}\right) \times (IT)}{G^2}$$

Las derivadas parciales con respecto a G_3 de los caracteres IVL, IVC, IVD, IVP, GDE_1 y GDE_3 son iguales a cero.

Por tanto:

$$\partial\Phi/\partial G_3 = - NP \times LVP \times (1-S_1) \times [(1-S_2) \times [GECC + (1-S_3/2) \times (GDNA_3 + (CD_3 \times PR_3)] \times (W_3 - W_2)/G_3] \times (1 - PRLV) \times IT / G^2$$

Derivadas parciales para el carácter CD2

$$\partial\Phi/\partial CD_2 = \frac{\left(\frac{\partial IVL}{\partial CD_2} + \frac{\partial IVC}{\partial CD_2} + \frac{\partial IVD}{\partial CD_2} + \frac{\partial IVP}{\partial CD_2}\right) \times (GT)}{G^2} -$$

$$- \frac{\left(\frac{\partial GC}{\partial CD_2} + \frac{\partial GDE_1}{\partial CD_2} + \frac{\partial GDE_2}{\partial CD_2} + \frac{\partial GDE_3}{\partial CD_2}\right) \times (IT)}{G^2}$$

Las derivadas parciales con respecto a CD_2 de los caracteres IVL, IVC, IVD, IVP, GDE_1 y GDE_3 son iguales a cero. Por tanto:

$$\partial\Phi/\partial CD_2 = [\, LVP \times NP \times (1-S_1) \times (1-S_2/2) \times (GDNA_2 + PR_2) \times A_2 \,] \times (IVL + IVC + IVD + IVP)$$

Derivadas parciales para el carácter CD3

$$\partial\Phi/\partial CD_3 = \frac{\left(\frac{\partial IVL}{\partial CD_3} + \frac{\partial IVC}{\partial CD_3} + \frac{\partial IVD}{\partial CD_3} + \frac{\partial IVP}{\partial CD_3}\right) \times (GT)}{G^2} -$$

$$- \frac{\left(\frac{\partial GC}{\partial CD_3} + \frac{\partial GDE_1}{\partial CD_3} + \frac{\partial GDE_2}{\partial CD_3} + \frac{\partial GDE_3}{\partial CD_3}\right) \times (IT)}{G^2}$$

Las derivadas parciales con respecto a CD_3 para los caracteres IVL, IVC, IVD, IVP, GDE_1 y GDE_2 son iguales a cero.

$$\partial\Phi/\partial CD_3 = [\, LVP \times NP \times (1-S_1) \times (1-S_2) \times [GECC + [(1-S_3/2) \times (GDNA_3 + PR_3)] \times A_3 \times (1-PRLV)] \times [IVL + IVC + IVD + IVP]$$

-III. RESULTADOS Y DISCUSION DEL EJEMPLO PLANTEADO.-

El estudio de pesos económicos nos depara una variación significativa en el beneficio generado por la unidad de producción determinada para este estudio (cerda reproductora), a diferencia del apartado anterior, ya que en este caso hay que utilizar la unidad de producción como base de evaluación para calcular el potencial económico que los diferente caracteres biológicos tienen en la economía de la explotación, siendo estos muy variados según los años estudiados:

B_{1989} = 17.047 PTAS.
B_{1990} = 10.188 PTAS.
B_{1991} = 746 PTAS.

B_{medio} = 9.121 PTAS.

El estudio de los pesos económicos de los caracteres biológicos suceptibles de mejora genética durante estos tres años aparece detallado en la Tabla 1, y nos muestra la influencia que cada carácter tiene sobre la generación final de beneficio.

Igualmente, y siguiendo las recomendaciones de Smith et al.(1986) abordamos el estudio de los valores relativos y los factores de proporcionalidad que independizaban los resultados de la base de evaluación y de las posibles reconversiones que pudieran acontecer en la explotación, obteniédose los resultados que igualmentre aparecen en la Tabla 1.

Con objeto de estudiar la solidez que a lo largo del periodo estudiado mostraban los dos métodos, hemos calculado una serie de ratios que podrían establecerse entre los resultados obtenidos (Tabla 2).

Analizando la Tabla 2 observamos que el gardo de dispersión de los valores relativos de los pesos económicos con respecto a la media se encuentra entre 1-12%, mientras que en el caso de los factores de proporcionalidad varía entre 0-20%, lo que nos indica el interés de realizar reviones contínuas de estos parámetros, aunque sólo el hecho de un cambio de dirección de la selección tendría efectos negativos en el progreso genético global.

Por último sería interesante estudiar el efecto de estas variaciones en grandes estructuras ganaderas que acceden siempre a un mismo mercado y que estas revisiones podrían aplicarse en líneas de machos finalizadores y abuelas.

TABLA 1. RELACION DE VALORES ABSOLUTOS, VALORES RELATIVOS Y FACTORES DE PROPORCIONALIDAD DE LOS DIFERENTES PESOS ECONOMICOS POR AÑOS

PESOS ECONOMICOS

		1989			1990			1991			MEDIA	
	V.ABS.	V.REL.	F.PROP.	V.ABS.	V.REL.	F.PROP.	V.ABS.	V.REL.	F.PROP.	V.ABS.	V.REL.	F.PROP.
LVP	6713	1	1	6274	1	1	5089	1	1	6017	1	1
NP	302	0,045	0,043	283	0,045	0,044	275	0,054	0,054	286	0,05	0,046
CC	3141	0,468	0,592	2875	0,458	0,515	2270	0,446	0,452	2723	0,45	0,517
G1	3,6	0,001	0,001	3,2	0,001	0,001	2,6	0,001	0,001	3,1	0,001	0,001
G2	112	0,017	0,022	99	0,016	0,019	81	0,016	0,016	97	0,02	0,019
G3	261	0,039	0,052	231	0,037	0,045	188	0,037	0,038	226	0,04	0,045
CD1	-47	-0,07	-0,009	-42	-0,007	-0,008	-34	-0,07	-0,07	-41	-0,01	-0,008
CD2	-70	-0,01	-0,014	-63	-0,01	-0,012	-51	-0,01	-0,01	-61	-0,01	-0,012

TABLA 2 .- RELACIONES ENTRE LOS DIFERENTES PESOS ECONOMICOS Y FACTORES DE PROPORCIONALIDAD PARA CADA UNO DE LOS CARACTERES

	LVP	NP	CC	G1	G2	G3	CD2	CD3
FP1/REL1(V1)	1.0	0.96	1.27	1.86	1.32	1.34	1.29	1.34
FP2/REL2(V2)	1.0	0.98	1.12	1.96	1.20	1.22	1.20	1.20
FP3/REL3(V3)	1.0	1.00	1.01	1.96	1.01	1.03	1.05	1.00
REL1/REL2	1.0	1.00	1.02	1.05	1.06	1.06	1.05	1.04
REL1/REL3	1.0	0.83	1.05	1.05	1.05	1.05	1.05	1.04
REL2/REL3	1.0	0.83	1.03	1.00	0.99	1.00	1.00	1.00
V1/V2	1.0	0.98	1.13	0.95	1.10	1.09	1.08	1.12
V1/V3	1.0	0.96	1.25	0.95	1.31	1.30	1.23	1.35
V2/V3	1.0	0.98	1.11	1.00	1.20	1.19	1.14	1.20
XREL/REL1	1.0	1.06	0.97	0.97	0.97	0.97	0.97	0.97
XREL/REL2	1.0	1.05	0.99	1.02	1.02	1.02	1.02	1.01
XREL/REL3	1.0	0.88	1.01	1.01	1.01	1.02	1.02	1.01
XFP/FP1	1.0	1.07	0.87	1.00	0.86	0.87	0.89	0.86
XFP/FP2	1.0	1.05	1.00	1.00	1.00	1.00	1.00	1.00
XFP/FP3	1.0	0.85	1.14	1.00	1.19	1.18	1.14	1.20
FP1/FP2	1.0	0.98	1.15	1.00	1.16	1.16	1.13	1.17
FP1/FP3	1.0	0.80	1.31	1.00	1.38	1.37	1.29	1.40

TABLA 2.- (CONTINUACION)

	LVP	NP	CC	G1	G2	G3	CD2	CD3
FP2/FP3	1.0	0.81	1.14	1.00	1.19	1.18	1.14	1.20
XFP/XREL	1.0	0.97	1.14	1.93	1.18	1.20	1.17	1.18
XFP/V1	1.0	1.01	0.90	1.03	0.89	0.90	0.91	0.88
XFP/V2	1.0	0.99	1.02	0.98	0.98	0.98	0.98	0.99
XFP/V3	1.0	0.97	1.13	0.99	1.17	1.16	1.12	1.19

TABLA 3 .- RESULTADOS TECNICOS (I)

INDICES	AÑO	
TECNICOS	1988	1989
PRODUCTIVIDAD CERDA:		
Lechones destetados cerda presente/año	18.34	18.00
Lechones destetados cerda productiva/año	19.95	20.90
RITMO DE REPRODUCCION:		
N° partos cerda presente/año	2.16	1.94
Intervalo entre partos	169.40	153.00
Intervalo destete-cubrición	-	6.00
Intervalo destete-cubrición Fértil	28.63	12.00
% de Fertilidad	-	83.40
% de Repeticiones	-	14.50
Duración de la gestación	-	116.00
Edad de destete	26.48	27.00
Edad media de las cerdas en meses	28.50	31.00
Edad media al primer parto en meses	18.16	13.70
% de abortos	0.67	0.66
PRODUCTIVIDAD:		
Lechones destetados por camada	8.72	8.77
Lechones nacidos totales por camada	10.23	10.58
Lechones nacidos vivos por camada	9.71	9.96
Lechones nacidos muertos por camada	0.52	0.60
% lechones nacidos muertos/nacidos totales	5.05	4.97
% lech.muertos lactancia/nacidos totales	14.70	17.02
% lech.muertos lactancia/nacidos vivos	9.72	12.05
Tasa de renovación	15.72	24.70
% camadas destetadas sobre total camadas	99.57	98.30
Camadas destetedas/cerda renovada	-	-
Número de cerdas dadas de baja	-	146.00
Edad media de cerda dada de baja	-	41.50
Camadas destetadas por cerda dada de baja	-	5.50
Intervalo último movimiento-Baja	-	69.70
Número medio de cerdas	547.10	590.90

TABLA 4.- RESULTADOS TECNICOS (II)

INDICES	AÑO	
TECNICOS	1990	1991
PRODUCTIVIDAD CERDA:		
Lechones destetados cerda presente/año	15.10	10.20
Lechones destetados cerda productiva/año	18.80	19.30
RITMO DE REPRODUCCION:		
N° partos cerda presente/año	1.82	1.52
Intervalo entre partos	160.00	161.00
Intervalo destete-cubrición	7.00	5.00
Intervalo destete-cubrición Fértil	18.00	24.00
% de Fertilidad	74.30	75.70
% de Repeticiones	21.40	24.00
Duración de la gestación	115.00	111.00
Edad de destete	27.00	26.00
Edad media de las cerdas en meses	30.00	29.00
Edad media al primer parto en meses	12.80	14.00
% de abortos	3.60	1.86
PRODUCTIVIDAD:		
Lechones destetados por camada	8.28	8.48
Lechones nacidos totales por camada	10.13	9.95
Lechones nacidos vivos por camada	9.45	9.49
Lechones nacidos muertos por camada	0.68	0.47
% lechones nacidos muertos/nacidos totales	6.71	4.72
% lech.muertos lactancia/nacidos totales	18.33	14.81
% lech.muertos lactancia/nacidos vivos	12.33	10.63
Tasa de renovación	25.70	46.10
% camadas destetadas sobre total camadas	97.00	98.19
Camadas destetedas/cerda renovada	5.6	6.1
Número de cerdas dadas de baja	221.00	55.00
Edad media de cerda dada de baja	41.30	37.70
Camadas destetadas por cerda dada de baja	5.50	6.10
Intervalo último movimiento-Baja	85.50	119.70
Número medio de cerdas	615.90	700.30

Recherche et identification de gènes à effets quantitatifs chez les animaux domestiques

J. M. ELSEN
INRA-SAGA, BP 27, 31326 Castanet-Tolosan, France

Introduction

La valeur opérationnelle, pour l'amélioration des populations, du modèle infinitésimal dans lequel un très grand nombre de gènes à faible effet explique la ressemblance entre les apparentés ne saurait être mise en doute. Un progrès génétique substantiel a été obtenu chez les plantes et les animaux domestiques par l'application des règles d'évaluation des valeurs génétiques et de sélection qui en découlent. Il est aujourd'hui difficile de prévoir si l'arsenal biotechnologique "moderne" supplantera un jour la machinerie traditionnelle en efficacité technique, et plus encore en rentabilité.

Certaines applications sont cependant dès aujourd'hui prometteuses, notamment l'introgression entre races ou le suivi intra race de gènes à effet quantitatif majeur ouvrant la porte à la construction de génotypes. D'un point de vue plus fondamental l'identification de gènes contrôlant les caractères à améliorer est une clé inestimable pour la compréhension des mécanismes régulant le fonctionnement du vivant, avec des perspectives d'applications dépassant largement la sélection des populations.

Nous présentons ici quelques pistes pour l'identification de tels gènes.

VII Reunión Nacional sobre Mejora Genética Animal, Lleida, 21-22 avril 1994.

Mise en évidence de gènes quantitatifs à partir des performances

Les généticiens appliqués ont essentiellement travaillé sur l'estimation de paramètres descriptifs de l'hérédité polygénique : héritabilité, corrélation génétique, effet d'hétérosis... Une abondante littérature est disponible sur les méthodes statistiques et protocoles expérimentaux nécessaires ainsi que sur les valeurs de ces paramètres pour de nombreux caractères dans de très nombreuses populations.

Dans le domaine de la génétique humaine au contraire, l'effort a surtout porté sur la mise en évidence de gènes à effet majeur sur la sensibilité aux maladies. Les fondations de ces recherches se trouvent dans Elston et Stewart (1971), Morton et McLean (1974) et Lalouel *et al.*(1983) pour l'analyse de ségrégation, Cannings *et al.*(1978) et Hasstedt (1982) pour l'analyse des pedigree.

Il est amusant de noter que ces dernières années, les généticiens appliqués ont exploré les phénomènes d'hérédité monogénique ou mixte cependant que la génétique humaine découvrait le concept d'hérédité multigénique de la sensibilité aux maladies. L'intérêt nouveau des améliorateurs de populations animales pour ces gènes à effet majeur s'explique en grande partie par la découverte successive des gènes "culards" chez le porc et les bovins et du gène Booroola chez les ovins. De nouveaux exemples ont été trouvés plus récemment dont un gène de rendement technologique pour le porc (Le Roy *et al.*, 1990), un autre gène d'hyperovulation sur le chromosome X de la brebis (Davis *et al.*, 1991) et un gène de vitesse de traite chez la chèvre (Ricordeau *et al.*, 1990).

Plusieurs critères très simples ont été proposés pour une exploration rapide de données quantitatives disponibles pour la mise en évidence de ces locus majeurs. Leur comparaison numérique (Le Roy et Elsen, 1989) montre que le test de Bartlett de comparaison des variances intra famille ainsi que le test de régression de la variance sur la moyenne familiale de Fain (1978) sont des critères puissants mais peu robustes. Une généralisation aux variables discrètes en a été proposée par Elsen et Le Roy (1990). D'autres critères sont imaginables. L'accent devrait être mis sur des critères non paramétriques ou pour le moins sur des critères plus robustes aux déviations par rapport aux hypothèses sur les distributions.

L'observation attentive de jeux de données existants, si possible avec l'aide des critères simples mentionnés ci dessus, peut suggérer la ségrégation d'un gène majeur. Des méthodes d'analyse plus sophistiquées sont alors disponibles pour tester cette hypothèse. Les quelques gènes majeurs solidement prouvés ont dans un premier

temps été repérés par la simple observation de valeurs extrêmes et du comportement mendélien du caractère dans certaines familles. A notre connaissance, seule l'hypothèse de l'existence d'un gène contrôlant l'adiposité du porc a été suggérée par une exploration systématique d'un ensemble de données disponibles (Le Roy *et al.*, 1990), cette hypothèse faisant actuellement l'objet d'une vérification expérimentale.

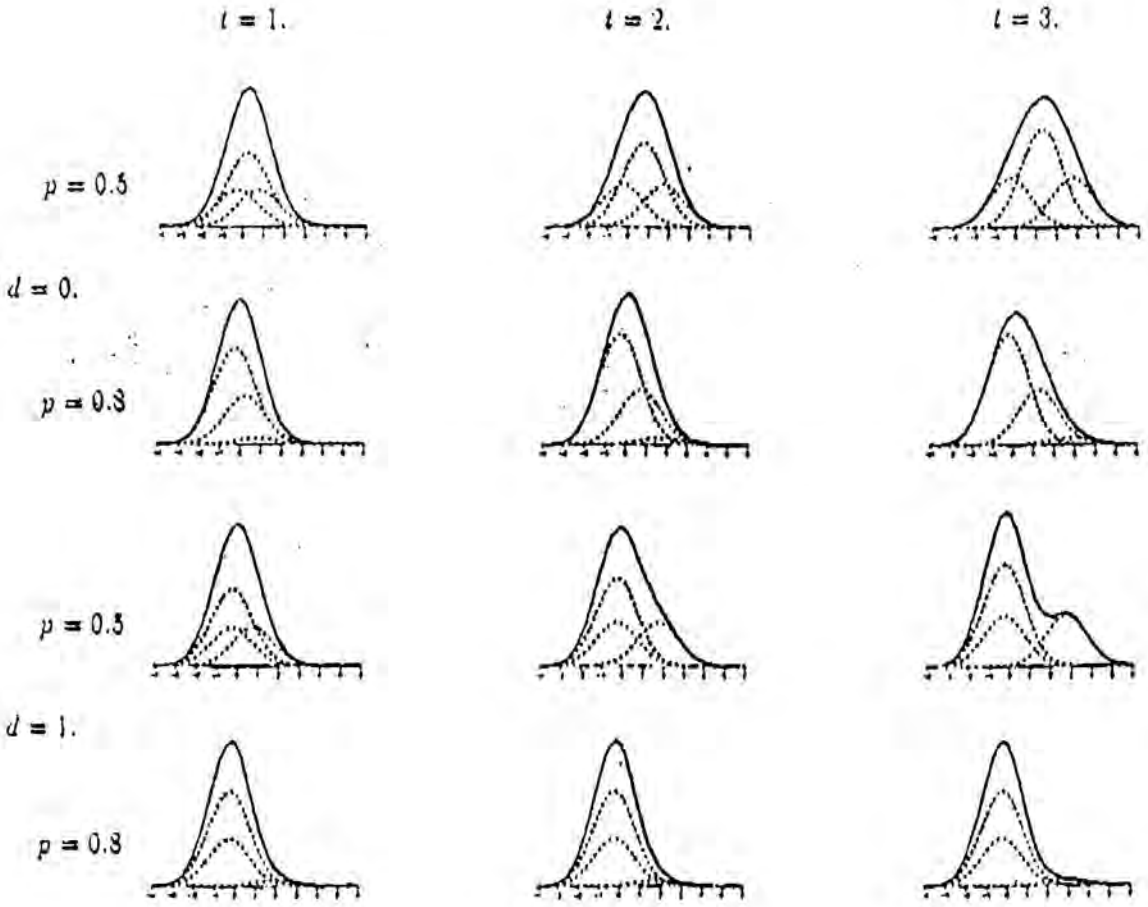
Les tests statistiques de l'existence de gènes majeurs font pour la plupart appel à la théorie des tests du rapport de vraisemblance, une hypothèse particulière H_0 (par exemple celle d'une hérédité polygénique, dans laquelle le contrôle génétique du caractère est dû à un grand nombre de gènes à petit effet, pouvant être décrite par une distribution normale des effets génétiques) étant testée contre une hypothèse générale l'emboîtant H_1 (par exemple une hérédité mixte dans laquelle un gène à effet important s'ajoute à l'hérédité polygénique). En principe, sous H_0, le rapport de vraisemblance entre ces deux hypothèses est asymptotiquement distribué selon une loi de χ^2, ce qui donne un seuil de rejet de l'hypothèse H_0.

Les méthodes proposées peuvent être schématiquement classées en trois catégories selon la quantité d'information généalogique considérée.

Les premières sont basées sur la seule distribution du caractère dans la population, sans considération de sa structure familiale. Par rapport à une normalité du caractère le plus souvent observée, la présence d'un gène majeur sera indiquée par un applatissement, une leptokurtie, voire une multimodalité si l'effet du gène est suffisamment important.

Sur la figure 1, la distribution globale du caractère est représentée, ainsi que les sous jacentes correspondant aux trois génotypes au locus majeur, pour diverses valeurs de l'écart t entre homozygotes (en unité d'écart type), de la dominance d (d variant entre 0, additivité et 1 dominance complète) et de la fréquence q de l'allèle défavorable. La déformation de la distribution globale due à la présence d'un gène n'est sensible que pour des effets très importants. Par ailleurs, cette approche globale du phénomène est très peu robuste puisque n'importe quel effet parasite inconnu avec deux ou trois classes peut générer les mêmes déformations qu'un gène majeur.

Figure 1
Distribution phénotypique du caractère au niveau de la population (traits continus) et intra génotype (traits pointillés) pour diverses hypothèses sur l'écart t entre homozygotes , la dominance d et la fréquence q du gène défavorable.



L'analyse de ségrégation forme la deuxième catégorie de méthodes. Ici, deux générations interviennent, soit simplement par la prise en compte des structures familiales (frères ou demi frères), soit par l'analyse simultanée des performances des deux générations, compte tenu des relations familiales. Par essence, c'est l'hétérogénéité des distributions entre familles, due à la variabilité des génotypes parentaux au locus majeur, qui est utilisée ici comme indicateur de la présence d'un gène majeur : au premier ordre, il s'agit de différences de moyennes pouvant être regroupées dans un nombre de classes correspondant au nombre de génotypes au locus majeur; au deuxième ordre, de variance supérieure chez les descendants de parents hétérozygotes; à l'ordre supérieur, du mélange de 2 ou 3 sous distributions chez ces mêmes descendants.

L'analyse de ségrégation a permis de "prouver" la ségrégation de nombreux gènes de sensibilité aux maladies et a été utilisée pour l'analyse de quelques corps de données chez les animaux domestiques. Elle est puissante, comme le rappelle le tableau 1 issu de Le Roy *et al.*(1989), mais peu robuste à la non normalité des distributions. Il est recommandé, quand la distribution globale du caractère est non normale, d'inclure dans l'analyse une transformation normalisante telle que la transformation de Box Cox (Go *et al.*, 1978).

Tableau 1
Puissance de l'analyse de ségrégation au niveau 5 p.cent (en p.cent) pour différentes structures de famille et cas de gène majeur (d'après Le Roy *et al.*, 1989)

Nombres de pères	Nombres de descendants	Cas 1	Cas 2	Cas 3
5	5	9	3	9
	10	17	11	7
	20	24	9	19
10	5	17	7	5
	10	27	8	19
	20	54	15	34
20	5	26	9	15
	10	51	21	33
	20	90	25	48

cas1 : $t = 2, d = 1, q = 0,5$
cas2 : $t = 2, d = 0, q = 0,5$
cas3 : $t = 2, d = 1, q = 0,9$

L'analyse de pedigree a pour objectif la modélisation des performances d'une population compte tenu de l'ensemble des relations de parenté entre les individus la constituant. Avec l'augmentation du nombre de générations dans l'analyse, une meilleure robustesse peut être espérée : pour que l'hypothèse monogénique (resp. mixte) soit retenue au détriment de l'hypothèse sporadique (resp. polygénique), il faut que la distribution conjointe du caractère dans l'ensemble du pedigree respecte les proportions mendéliennes. Des solutions directes, le "peeling" (Cannings *et al.*, 1978) ou itératives (Van Arendonk *et al.*, 1989) ont été proposées pour traiter le cas d'une hérédité monogénique. Les premières garantissent l'obtention de la vraisemblance exacte mais sont très exigeantes en temps de calcul. La convergence des secondes, beaucoup plus rapides, vers la solution exacte, n'est pas prouvée.

Aujourd'hui, il n'existe pas de méthodes exactes pour le traitement de l'hérédité mixte. Hasstedt (1982) a proposé une approximation efficace qui est très largement utilisée en génétique humaine. L'ordre de lecture du pedigree est fondamental dans le temps de résolution de l'algorithme. Le choix du meilleur ordre est obscur, et dans le cas des fichiers animaux, de très grandes tailles, la démarche n'est pas utilisable. Le développement récent des méthodes de "Gibbs sampling" résoudra peut être une partie de ces difficultés. Il faudrait par ailleurs évaluer précisément l'intérêt de multiplier le nombre de générations considérées dans l'analyse. Si pour les populations humaines caractérisées par de très petites familles nucléaires, il est sans aucun doute intéressant de suivre la ségrégation des gènes suspectés sur le plus grand nombre de générations possible, cet intérêt est probablement limité dans le cas des animaux pour lesquels un même père peut avoir un très grand nombre de descendants.

En conclusion, l'analyse des données quantitatives a largement contribué à mettre en évidence des gènes majeurs. Elle présente cependant trois difficultés :

- manque de robustesse par rapport aux déviations aux hypothèses relatives à la distribution du caractère, les solutions possibles étant la normalisation des données, l'augmentation du nombre de générations analysées, et peut être à terme l'utilisation de tests non-paramétriques.

- lourdeur des calculs nécessaires à l'évaluation des vraisemblances. Des méthodes approchées peuvent être proposées (par exemple Le Roy *et al.*, 1989 ou Knott *et al.*, 1991), mais le traitement exact des grands pedigree est totalement irréaliste.

- conséquence des deux premières difficultés, l'évaluation d'hypothèses oligogéni-

ques, dans lesquelles le caractère est supposé sous la dépendance de plusieurs (quelques) gènes, n'est pas possible avec ces méthodes. Ceci réduit par conséquent la puissance de détection des gènes pris indépendamment.

Apport des marqueurs pour la détection des gènes majeurs

Des cartes génétiques commencent à voir le jour pour les animaux domestiques (Bishop *et al.*, 1994; Barendse *et al.*, 1994; Rohrer *et al.*, 1994). Ces cartes vont progressivement se renforcer, se détailler et devenir disponibles pour les différentes espèces. Elles seront une source d'information pour la mise en évidence de gènes à effet quantitatif.

L'intérêt des marqueurs découle de la possibilité d'identifier dans la descendance d'un parent commun, les individus ayant reçu l'un ou l'autre des deux segments chromosomiques entourant le locus marqueur chez le parent. Si un locus à effet quantitatif est situé sur ce segment, et si le parent est double hétérozygote aux locus marqueur et majeur, une différence phénotypique systématique sera observée entre les descendants ayant reçu l'un ou l'autre des allèles marqueurs. L'existence d'un gène majeur sera alors testée par l'effet du marqueur intra famille sur le caractère quantitatif, ce qui peut être réalisé dans le cadre classique de l'analyse de variance. La famille peut être un ensemble de demi frères de pères (Soller et Genizi, 1978; Weller *et al.*, 1990), de plein frères (Soller et Genizi, 1978), ou l'ensemble des animaux F2 ou Backcross si les souches parentales sont homozygotes (Soller *et al.*, 1976).

Le phénomène peut être présenté plus largement. Dans la démarche précédente (utilisation des seules données quantitatives), la transmission des gènes entre générations est suivie statistiquement : un descendant reçoit aléatoirement en tout locus l'un ou l'autre des allèles possédés par son père (et par sa mère). Si ce locus a un effet majeur, la distribution du caractère pour ce descendant, conditionnellement aux génotypes des parents, sera un mélange de 4 distributions en proportions égales, certains éléments du mélange pouvant être identiques si l'un au moins des parents est homozygote.

L'observation d'un marqueur du gène en question a fondamentalement pour effet de limiter le nombre d'éléments du mélange. Le génotype conjoint aux locus marqueur et majeur de chaque parent est tout d'abord déduit de l'observation de

son ascendance ou de sa descendance. L'information vient de l'ascendance dans le cas de parents F1 nés d'un croisement entre lignées homozygotes. Elle vient de la descendance dans le cas de familles de frères ou demi frères appartenant à des populations "outbred". Dans une analyse de pedigree les deux sources d'informations se mélangent.

Le tableau 2 donne à titre d'exemple le nombre de composantes de la distribution du caractère chez un descendant issu de parents hétérozygotes au locus majeur selon les informations apportées par le locus marqueur. Une liaison parfaite est supposée entre les deux locus.

Tableau 2
Nombre de composantes de la distribution d'un caractère contrôlé par un gène majeur chez un individu (cas où les deux parents sont supposés hétérozygotes au locus majeur).

Génotype marqueur			Nombre de
Père	Mère	Descendant	composantes
non mesuré			4
M1M1	M1M1	M1M1	4
M1M1	M1M2	M1M1	2
M1M1	M1M2	M1M2	2
M1M1	M2M2	M1M2	4
M1M2	M1M2	M1M1	1
M1M2	M1M2	M2M2	1
M1M2	M1M2	M1M2	2

Des méthodes d'analyse de ségrégation prenant en compte l'information apportée par des marqueurs ont été étudiées par McLean *et al.*(1984), Risch (1984), Knott et Haley (1992), Le Roy et Elsen (1994). Contrairement aux critères de Soller et Genizi (1978) et de Weller *et al.*(1990) qui n'utilisent que les différences de moyennes entre groupes, ces méthodes exploitent les différences de formes de distributions expliquées ci dessus. Elles sont donc plus puissantes, comme l'indique le tableau 3 tiré de Le Roy et Elsen (1994).

Table 3
Puissance du protocole de Weller *et al.*(d'aprés Le Roy et Elsen, 1994)
5 grand pères, 100 fils/grand père, 100 filles/fils

Moyennes † AA	AB	BB	Taux de recombinaison	WKS	SA	SAM
-0.1	0	0.1	0	75	63	95
			0.1	61	64	85
			0.2	37	64	77
			0.5	3	72	74
0	0	0.2	0	75	66	95
			0.1	61	73	91
			0.2	33	67	87
			0.5	4	66	63

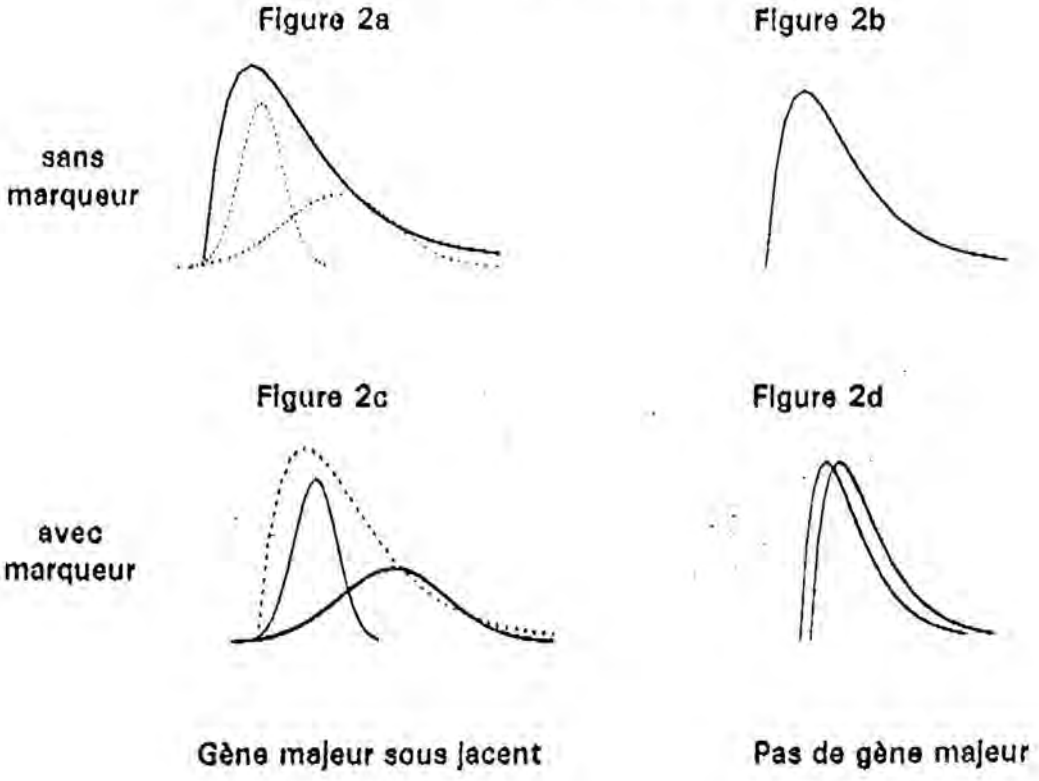
† intra génotype, en écart type phénotypique
WKS = Statistique de test de Weller *et al.*(1990)
SA = Analyse de ségregation sans marqueur
SAM = Analyse de ségrégation utilisant l'information d'un marqueur

Il est probable cependant que ce surcroît de puissance soit obtenu au détriment de la robustesse, le test statistique du rapport de vraisemblance utilisé supposant connues les distributions intra génotype.

Par rapport aux méthodes basées sur les seules données phénotypiques, l'analyse simultanée de ces données et de la transmission d'allèles marqueurs entre générations a un triple intérêt :

- plus de robustesse (notamment pour les méthodes d'analyse de variance, le cas des méthodes du maximum de vraisemblance restant à chiffrer). La figure 2 illustre le phénomène : une distribution leptokurtique peut avoir pour origine la nature même du caractère (épaisseur de lard chez le porc par exemple) (figure 2a) ou l'existence de deux ou trois distributions sous jacentes correspondant à des génotypes majeurs (figure 2b). L'absence de robustesse de l'analyse de ségrégation sans marqueur vient de cette confusion. Quand un marqueur est disponible, les distributions sous jacentes sont repérées par les génotypes individuels en ce locus (figures 2c et d).

Figure 2
Apport des marqueurs pour l'analyse des causes de leptokurtie des distributions intra famille.



- la possibilité de traiter l'hérédité oligogénique. Ce cas ne pose aucune difficulté dans le cadre des modèles linéaires d'analyse des effets marqueurs intra famille, sauf dans le cas de gènes à effets quantitatifs liés, dont l'effet sur le caractère sera surestimé ou sous estimé selon la répartition des haplotypes chez les parents (par exemple, deux locus très liés avec des allèles A1, A2 et B1, B2 (A1 et B1 augmentant la valeur phénotypique par rapport à A2 et B2) ne seront pas detectés dans la descendance de parents A1B2/A2B1).

  Le cas de l'analyse de ségrégation est loin d'être résolu. Une première approche pourrait consister dans un premier temps à appliquer successivement le même modèle d'hérédité mixte (un gène majeur + des polygènes) aux différents

points marqués du génome puis à reprendre l'analyse en multilocus pour les seuls marqueurs s'avérant positifs. La difficulté algébrique de ces modèles et la quantité de calcul nécessaire à leur traitement devra être mis en balance avec le gain de puissance qu'ils donneraient en comparaison avec des modèles linéaires.

- comme cela sera décrit plus loin, l'étape suivant la mise en évidence d'un gène à effet quantitatif est sa localisation sur le génome. L'utilisation des marqueurs pour la detection des QTLs (Quantitative Trait Locus) fourni simultanément des premiers éléments pour leur localisation.

De nombreuses difficultés théoriques restent par ailleurs à résoudre pour l'utilisation des cartes génétiques. Elles sont essentiellement liées au caractère multidimensionnel du problème : présence simultanée de nombreux marqueurs indépendants et non indépendants, avec des quantités d'informations variables d'un marqueur à l'autre (Haley *et al.*, 1994) ; forte probabilité de la ségrégation de plusieurs (voir un grand nombre) de locus quantitatifs liés ou non liés et à effets variés; interactions épistatiques entre locus quantitatifs; effets pléiotropiques de ces gènes sur plusieurs caractères. Les biotechnologies ont donc ouvert un champ de recherche très vaste pour les mathématiciens de la génétique quantitative. Il faut à la fois définir des tests statistiques appropriés, en estimer les propriétés et proposer des algorithmes efficaces.

Populations expérimentales

En génétique humaine la recherche de familles informatives est une difficulté importante justifiant le recours à l'analyse de pedigree complexes qui tire au mieux partie de l'information disponible. Pour les généticiens des plantes ou des animaux la situation est bien meilleure puisqu'il est possible de créer de toutes pièces les populations d'étude. Bien souvent d'ailleurs ces populations existent avant même la mise en place d'un projet de détection de gènes quantitatifs : populations croisées Backcross ou F2 utilisant les complémentarités entre races et effets d'hétérosis, grandes familles de demi frères de pères, en particulier quand l'insémination artificielle est disponible.

Tous les protocoles proposés, que des marqueurs soient ou non disponibles, reposent sur l'observation de la ségrégation du caractère quantitatif (ou de sa coségrégation avec un locus marqueur) dans la descendance de reproducteurs : si un gène ma-

jeur existe et qu'au moins un des parents est hétérozygote A1A2 pour ce gène, alors la distribution du caractère sera un mélange en proportions mendéliennes. Dans une F2, par exemple, les éléments du mélange seront 1/4 (A1A1), 1/2 (A1A2) et 1/4 (A2A2) ou 3/4 (A1) et 1/4 (A2A2)selon que le gène est additif ou dominant.

Pour accroître la puissance des protocoles, plusieurs solutions ont été proposées.

La première est d'augmenter la probabilité que les parents soient hétérozygotes au locus quantitatifs. Pour les plantes et les espèces de laboratoires, il est possible de disposer de lignées homozygotes. Tous les descendants F1 seront hétérozygotes pour les locus fixés en des états opposés dans les lignées parentales. Le projet européen PiGMaP qui utilise des croisements entre Large White et Meishan essaie de faire profit de ce phénomène, les races parentales étant supposées porteuses d'allèles différant pour un grand nombre de caractères dont la prolificité, la croissance et l'adiposité. Au contraire, les protocoles proposés par Soller et Genizi (1978) et par Weller *et al.*(1990) sont applicables intra race. Les parents peuvent donc avoir divers génotypes aux locus quantitatifs et seule une partie d'entre eux seront hétérozygotes. Le tableau 4 donne une idée de la puissance d'un protocole "père" selon le pourcentage de pères effectivement hétérozygotes étudiés.

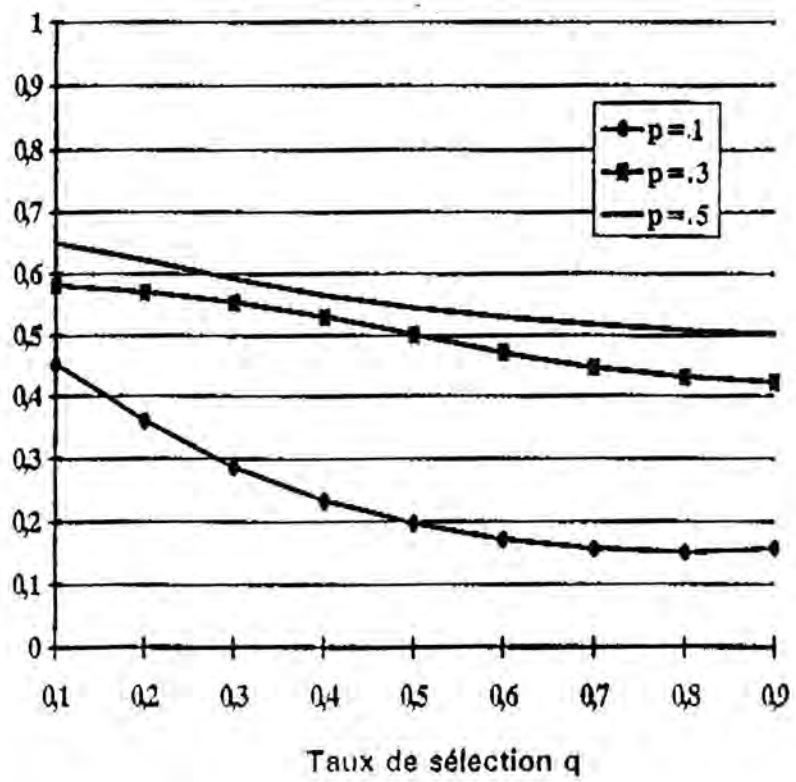
Table 4
Puissance du protocole de Soller et Genizi (1978) selon la proportion de pères hétérozygotes au locus majeur.
(30 pères, 40 descendants/père)

Nombre de pères AB	Puissance au seuil 5 pcent	Puissance au seuil 1 pcent
Gène dominant de 1 écart type d'effet		
5	37	21
10	75	53
15	100	95
20	100	96
25	100	98
30	100	100
Gène dominant de 0.5 écart type d'effet		
5	13	3
10	16	6
15	30	13
20	38	18
25	48	27
30	61	36

Cette proportion est un élément fondamental pour l'efficacité du dispositif. Dans le cas des protocoles visant simultanément à trouver des QTLs pour plusieurs caractères, le degré d'hétérozygotie de la population des pères ne peut pas être contrôlée. Par contre, si l'objectif est l'analyse génétique d'un caractère particulier, les reproducteurs peuvent être choisis *a priori*.

Deux solutions sont envisageables : choisir, pour l'analyse, des reproducteurs dont une première série de descendants s'est avérée très polymorphe pour le caractère, ou accoupler des grands parents de phénotypes extrêmes opposés. La figure 3 illustre l'efficacité du procédé (il y est supposé que le gène est en équilibre de Hardy Weinberg avec une fréquence q pour l'allèle A1 avant la selection et que le taux p de sélection est le même pour les grands parents des deux sexes).

Figure 3
Augmentation du pourcentage d'hétérozygotes AB au locus majeur par sélection divergeante de leur parents
(fréquence initiale p, taux de sélection q)



Une seconde solution est d'augmenter la différence phénotypique entre les groupes de descendants définis par l'allèle reçu du père. C'est le principe du protocole de Weller *et al.*(1990) comparé au protocole de Soller et Genizi (1978) : la comparaison ne porte plus sur des différences individuelles mais sur des différences de moyennes de descendance de n filles, avec une variance intra groupe diminuée d'un facteur $\sqrt{n}$. Une autre solution pourrait être la comparaison de descendants de pères supposés hétérozygotes accouplés à deux groupes de mères choisies de façon opposée sur le caractère étudié, et donc plutôt A1A1 pour le premier groupe, plutôt A2A2 pour le deuxième groupe.

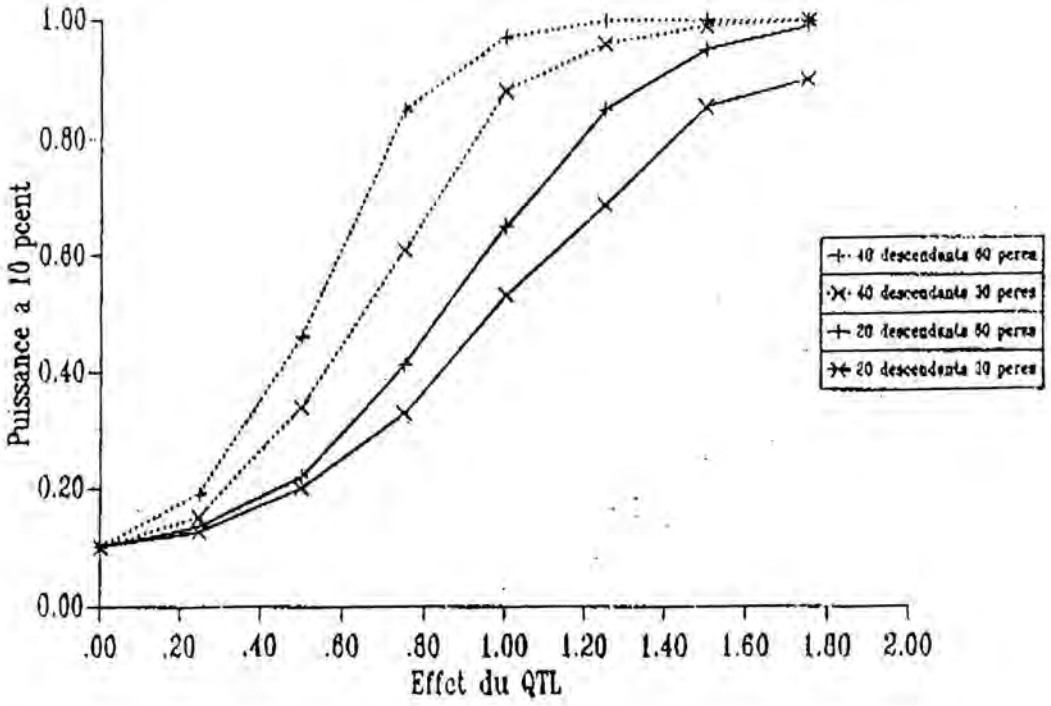
Le tableau 5 donne la valeur moyenne des groupes de descendants en supposant que la sélection des femelles a créé un écart polygénique δ entre le groupe des positives "+" et des négatives "−", et que les moyennes génotypiques sont respectivement de $-\Delta$, 0 et Δ pour les A1A1, A1A2 et A2A2.

Tableau 5
Moyennes intragroupes dans un protocole de détection de QTL avec sélection des conjointes

Hypothèse génétique	Pas de QTL		Père A1M1/A2M2		Père A1M2/A2M1	
Type de mère	+	−	+	−	+	−
M1 reçu	δ	$-\delta$	δ	$-\delta-\Delta$	$\delta+\Delta$	$-\delta$
M2 reçu	δ	$-\delta$	$\delta+\Delta$	$-\delta$	δ	$-\delta-\Delta$

Soit μ_{ij} la valeur moyenne de la cellule ij. Alors que le contraste proposé par Soller et Genizi (1978) entre descendants ayant reçu l'allèle M1 et M2 de leur père a une espérance de Δ, la statistique de test $t = ||\mu_{M1+} - \mu_{M2-}| - |\mu_{M1-} - \mu_{M2+}||$ a pour espérance 2Δ. Un test basé sur cette quantité devrait donc être sensiblement plus puissant que le test étudié par Soller et Genizi (1978). La figure 4 qui rapporte le nombre d'individus nécessaires à l'obtention d'une puissance de 90 pcent (au niveau 5 pcent) en fonction de l'effet du QTL dans le protocole "père" montre l'intérêt de cette selection des conjointes.

Figure 4
Puissance d'un QTL dominant localié à 0 cM du marqueur



Approche du gène

L'observation de la ségrégation d'un gène majeur dans une population n'est qu'une première étape à partir de laquelle il faut essayer de situer son locus puis d'identifier sa séquence. En outre, la mise en évidence de marqueurs présente souvent un grand intérêt appliqué puisqu'ils sont un outil performant pour la sélection des candidats reproducteurs sèlon leur génotype au locus majeur, notamment quand le phénotype correspondant au gène incriminé ne s'exprime que dans un seul sexe (gène d'ovulation des ovins), tardivement (débit de traite de la chèvre), après l'abattage (rendement technologique du porc) ou avec une pénétrance incomplète (pratique-

ment tous les gènes).

Quand le gène a été prouvé à partir des seules informations phénotypiques la première étape est donc de trouver un premier marqueur de son locus. Cette étape est déjà franchie dans le cadre des protocoles de détection de QTLs utilisant la carte génétique.

Recherche d'un marqueur génétique

Si les nouveaux polymorphismes de l'ADN (microsatellites, minisatellites, RAPDs) sont aujourd'hui la source principale de marqueurs, les systèmes plus classiques comme les RFLPs, voire les groupes sanguins ou le complexe majeur d'histocompatibilité ont toujours leur intéret. On peut rappeler à ce propos l'utilisation très large qui fut faite des locus PHI et PGD pour l'élimination du gène de la sensibilité à l'halothane chez le porc.

La création de "familles informatives" de bonne qualité est un préalable fondamental pour avoir de bonnes chances de trouver un marqueur. Par nature, la recherche d'un tel gène est une opération à caractère systématique qui donnera obligatoirement un résultat positif à partir du moment où ces familles ont la structure et l'effectif nécessaires. Le facteur chance est par contre plus important dans le nombre de candidats marqueurs qu'il faudra tester avant l'obtention d'un résultat favorable (cf par exemple la découverte après le test de 6 locus uniquement, d'un marqueur très lié au gène callipyge chez les ovins : Cockett *et al.*, 1994), même si un choix raisonné de ces candidats peut limiter la quantité de travail maximum.

La structure optimale, définie par les génotypes majeurs des parents dont on étudiera les méioses, dépend du type de gène majeur que l'on veut marquer (tableau 6). Dans tous les cas de figure, un des deux parents au moins doit être hétérozygote au locus majeur. Dans les "backcross" (le deuxième parent homozygote) seules les méioses du parent double hétérozygote (aux locus majeur et marqueur) sont utiles pour l'estimation du degré de liaison entre les gènes. Dans le cas de l'"intercross" (les deux parents hétérozygotes) les méioses observées chez les deux parents sont utiles, mais seuls les descendants homozygotes au locus majeur sont interprétables. La quantité totale d'information est donc a priori la même dans les deux solutions. Le plus souvent cependant les femelles ont peu de descendants par rapport aux mâles. Si la phase ne peut pas être déduite de l'observation de leur ascendance, la solution du backcross devra être préférée.

Le choix du type d'accouplement reposera également sur d'autres considérations telles que la pénétrance selon les génotypes et la viabilité des homozygotes.

Tableau 6
Choix des parents de la population pour la recherche d'un marqueur d'un gène majeur selon les caractéristiques de ce gène

Type de gène	Génotype du parent 1	Génotype du parent 2	Exemple
Autosomique avec aa récessif non viable	Aa	Aa	Intersexualité (porc)
Autosomique avec dominance	Aa	aa	Rendement technologique (porc)
Autosomique additif	AB	AA ou BB	Booroola (ovin)
Lié au sexe codominant	X_AY ou X_BY	X_AX_B	Gène Inverdale (ovin)
Lié au sexe dominant	X_aY	X_AX_a	Gène de nanisme (poule)

Selon l'espèce, les familles de référence pourront être des familles "nucléaires" (un couple parental et ses descendants), des familles de demi frères de père ou des mélanges hiérarchisés de frères et demi frères (P pères étant accouplés à M mères ayant chacune D descendants). Des structures croisées, dans lesquelles mâles et femelles ont plusieurs conjoints sont également concevables mais plus difficiles à analyser.

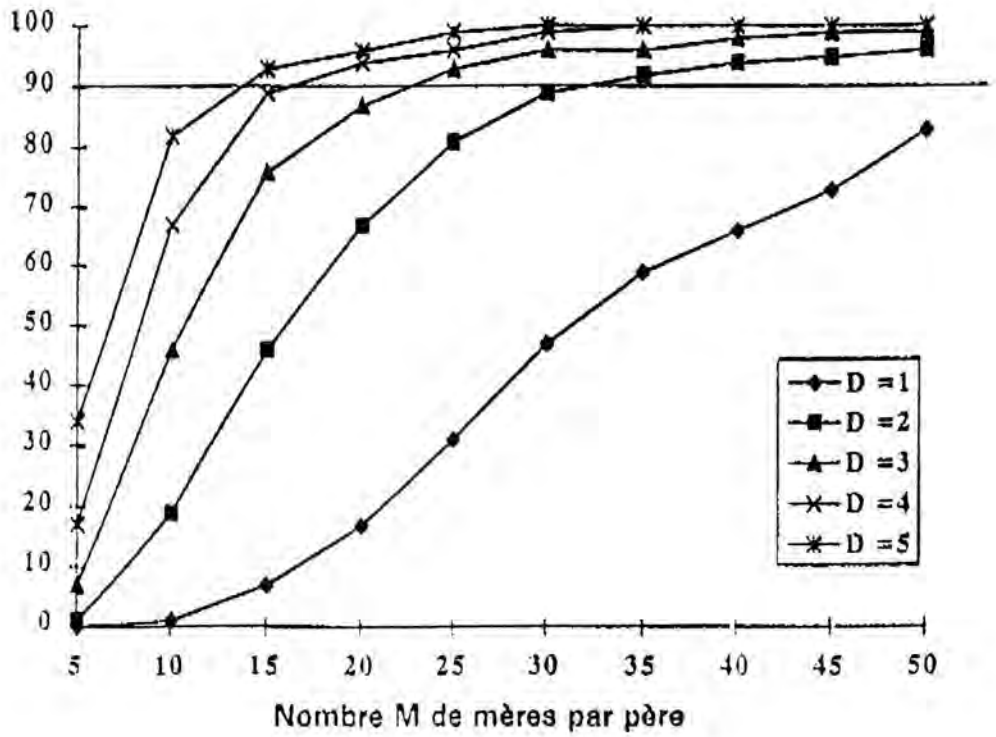
Le choix du nombre total de descendants mesurés et de la structure familiale (définie par les paramètres P, M et D) peut être raisonné à partir de la distribution du Lod Score, critère statistique habituellement retenu pour le test de la liaison entre deux locus (le lod score $z(\theta)$ est le logarithme décimal du rapport des vraisemblances de l'échantillon sous les hypothèses d'absence de liaison et de liaison au taux de recombinaison θ). Suivant les recommandations de Morton (1955), l'hypothèse d'une liaison entre deux gènes peut être conservée si le lod score est supérieur à 3, rejetée s'il est inférieur à -2.

Une première démarche, suivant Ott (1991) est de comparer la valeur de l'espérance du lod score maximum (l'ELOD, $E(z(\hat{\theta}))$) pour différentes populations. Les propriétés d'additivité de l'ELOD facilitent l'utilisation de ce critère. Il a été notamment appliqué par Van der Beek et Van Arendonk (1993) à l'optimisation d'une population informative chez les animaux. Ces auteurs ont trouvé que pour un taux de recombinaison de 20 pour cent, 27 descendants donnent un ELOD supérieur à 3 quand 5 allèles sont en ségrégation et 55 quand il n'y a que 2 allèles.

Le critère de l'ELOD n'a pas de signification statistique précise. Une autre démarche, beaucoup plus satisfaisante, serait d'estimer la puissance du test du lod

score ($\pi(s) = prob(z(\theta) > s)$) et de retenir les structures expérimentales garantissant que cette puissance $\pi(s)$ soit supérieure à une probabilité $1 - \beta$ donnée. Cette démarche avait été proposée par Demenais *et al.*(1988) dans un cadre humain, les auteurs montrant par simulation que 10 à 40 familles nucléaires avec 6 enfants assurent une puissance supérieure à $1 - \beta = 0.80$ pour un seuil de lod score $s = 2$ quand le taux de recombinaison varie de 0 à 20 pour cent. Nous avons entrepris ce type d'analyse afin d'optimiser des populations expérimentales pour la recherche de différents gènes majeurs chez le porc et la chèvre. A titre d'exemple, la figure 5 représente, pour un marqueur à 20 pour cent de recombinaison du gène majeur, la valeur de $\pi(5)$ pour $P = 6$ pères, et un nombre varié de mères ($M = 5$ à 50) et de descendants par mère ($D = 1$ à 5). On notera que l'effectif total de descendants à mesurer est de l'ordre de 60 par père pour les valeurs retenues de s et β. Un travail est en cours pour généraliser ces résultats.

Figure 5
Probabilité d'obtenir un lod score supérieur à 5 pour diverses populations de 6 familles de pères



Le seuil s était fixé à 3 par Morton (1955) dans le cas de l'étude d'un couple de gènes. Dans la recherche d'un marqueur d'un gène quantitatif, une série de marqueurs est successivement testée jusqu'à l'obtention d'une liaison. Le seuil de rejet doit donc être augmenté en conséquence, en considérant aussi la non indépendance entre gènes localisés sur le même chromosome. Ott (1991) a proposé comme seuil pour l'acceptation d'une liaison génétique la quantité $3 + \log(g - 100)$ où g est le nombre de marqueurs testés, si $g > 100$.

Comment réduire les coûts d'analyse ?

Plusieurs propositions ont été faites pour limiter la quantité de génotypes marqueurs à mesurer, en supposant que le travail de laboratoire est d'un coût très supérieur aux coûts d'élevage des animaux. Les deux premières propositions considèrent le phénomène de pénétrance incomplète qui prévaut pour les gènes majeur contrôlant les caractères quantitaifs. Les deux suivantes sont basées sur l'utilisation d'un mélange d'ADN.

Le "selective genotyping" de Lander et Botstein (1989) consiste à sélectionner parmi l'ensemble des descendants d'un backcross, les individus extrêmes, vers les basses et hautes valeurs. Le raisonnement est tout à fait transposable à une famille de frères ou demi frères de père. Une grande partie de l'information sur la liaison entre le QTL et le marqueur est en effet apportée par ces individus extrêmes qui ne sont pas ambigus pour leur génotype majeur. D'après les auteurs, une sélection des individus situés à plus de 1 écart type de la moyenne, avec une augmentation de 25 pour cent d'individus élevés, permet de réduire le travail de laboratoire de 60 pour cent.

Le "sequential sampling" de Motro et Soller (1993) est basé, comme le lod score de Morton (1955), sur le principe des tests séquentiels de Wald (1947). Le critère est ici une comparaison des valeurs moyennes des individus selon l'allèle marqueur reçu du parent double hétérozygote. Le nombre d'individus dont le génotype en un marqueur est mesuré est augmenté jusqu'à un seuil d'acceptation ou de rejet définitif de la liaison génétique. D'après les auteurs, un tel dispositif permettrait de réduire considérablement (jusqu'à 50 pour cent) le travail de laboratoire.

Dans la méthode dite de "bulk segregant analysis" (Michelmore *et al.*, 1991), un test RAPD est utilisé pour évaluer en parallèle deux mélanges d'ADN qui diffèrent l'un de l'autre par le génotype des individus le composant au locus majeur dont on cherche un marqueur. La technique demande qu'un nombre suffisant d'individus soient mélangés dans chaque groupe afin qu'aucun biais d'échantillonage ne fasse

apparaître une différence fortuite entre les groupes. La technique a été utilisée avec succès chez les plantes (Michelmore *et al.*, 1991); nous la testons sur des familles de Boorola en collaboration avec l'Université de Davis, Californie.

Un article récent de Lisitsyn *et al.*(1994) propose une nouvelle approche, la "Genetically Directed Representational Difference Analysis" (GDRDA) basée sur une technique de soustraction d'ADN (RDA, Lisitsyn *et al.*, 1993). Deux groupes d'individus différant statistiquement par un seul locus sont constitués, le "tester" et le "driver". Il peut s'agir de deux individus congéniques si de telles lignées homozygotes sont disponibles, ou d'un père hétérozygote AB en un locus majeur (tester) et un mélange de ses filles homozygotes AA en ce locus (driver) produites par l'accouplement avec des conjointes homozygotes AA. Les auteurs proposent un procédé utilisant la PCR qui permet d'isoler le fragment chromosomique autour du gène majeur. Outre sa rapidité d'exécution, cette technique ne nécessite pas la connaissance préalable d'une carte génétique de l'espèce et pourrait donc être universellement utilisée.

Conclusion

Le développement récent des techniques de biologie moléculaire a modifié profondément notre approche de la génétique quantitative. Le repérage de gènes contrôlant les caractères est facilité, et leur localisation sur le génome possible. Ce sont les premiers pas vers leur identification qui pourra ouvrir de nouveaux horizons dans la compréhension de mécanismes et l'amélioration des espèces.

Les généticiens quantitatifs doivent jouer un rôle actif dans cette nouvelle approche. L'efficacité de la sélection traditionnelle n'est pas remise en cause mais il faut intégrer dans son organisation les nouvelles possibiltés offertes par l'identification des QTLs, comme l'ont été l'insémination artificielle, le transfert des embryons et le contrôle des paternités par les groupes sanguins. Par ailleurs, toutes ces techniques de laboratoires utilisent de l'ADN extrait d'animaux vivants dans des populations dont la struture doit être optimisée, tâche relevant pleinement du savoir faire statistique des généticiens quantitatifs. Enfin un champ d'investigation très large nous est ouvert comprenant les méthodes statistiques d'analyse du génome, les techniques d'évaluation des reproducteurs incluant les informations issues de la carte génétiques et la mise au point de plans de sélection, d'introgression, de contruction de génotypes... adaptés aux contraintes et objectifs de l'élevage.

References

Barendse W., Armitage S.M., Kossarek L.M., Shalom A., Kirkpatrick B.W., Ryan A.M., Clayton D., Li L., Neibergs H.L., Zhang N., Grosse W.M., Weiss J., Creighton P., McCarthy F., Ron M., Teale A.J., Fries R., McGraw R.A., Moore S.S., Georges M., Soller M., Womack JE., Hetzel D.J.S. (1994) A genetic linkage map of the bovine genome. *Nature Genetics*, 6 : 227-.

van Arendonk J.A.M., Smith C., Kennedy B.W. (1989) Method to estimate genotype probabilities at individual loci in farm livestock. *Theor. Appl. Genet.*, 78 : 735-740.

van der Beek S., van Arendonk J.A.M. (1994) Criteria to optimize designs for detection and estimation of linkage between marker loci from segregating populations contining several families. *Theor. Appl. Genet.*, 86 : 269-280.

Bishop M.D., Kappes S.M., Keele J.W., Stone R.T., Sunden S.L.F., Kawkins G.A., Toldo S.S., Fries R., Grosz M.D., Yoo J., Beatttie C.W. (1994) A genetic linkage map in cattle. *Genetics*, 136 : 619-639.

Cannings C., Thompson E.A., Skolnick H.H. (1978) Probability functions on complex pedigrees. *Adv. Appl. Prob.*, 10 : 26-61.

Cockett N.E., Jackson S.P., Shay T., Nielsen D., Moore S.S., Steele M.R., Barendse W., Green R.D., Georges M. (1994) Chromosomal localisation of the callipyge gene in sheep (*Ovis aries*) using bovine DNA markers. *Proc. Natl. Acad. Sci., USA*, 91.

Davis G.H., McEwan J.C., Fenessy P.F., Dodds K.G., Farquhar P.A. (1991) Evidence for the presence of a major gene influencing ovulation rate on the X chromosome of sheep. *Biol. Reprod.*, 44 : 620-624.

Demenais F., Lathrop G.M., Lalouel J.M. (1988) Detection of linkage between a quantitative trait and a marker locus by the lod score method : sample size and sampling consideration. *Ann. Hum. Genet.*, 52 : 237-246.

Elsen J.M., Le Roy P. (1990) Detection of major genes and determination of genotypes. Application to discrete variables. *In : Hill W.G., Thompson R.,*

Wolliams J.A. (eds.), 4th world congress on genetics applied to livestock production, vol XV, pp 37-49, Edinburgh.

Elsen J.M., Vu Tien J., Bouix J., Ricordeau G. (1985) Linear programming model for incorporating the Booroola gene in an other breed. *In : Land R.B., Robinson D.W. (eds.), Genetics for reproduction in Sheep*, pp 175-181, Butterworths, London, U.K.

Elston R.C., Stewart J. (1971) A general model for the genetic analysis of pedigree data. *Hum. Hered.*, **21** : 523-542

Fain P.R. (1978) Characteristics of simple sibship variance tests for the detection of major loci and application to height, weight and spatial performance. *Ann. Hum. Genet.*, **42** : 109-120.

Go R.C.P., Elston R.C., Kaplan E.B. (1978) Efficiency and robustness of pedigree segregation analysis. *Am. J. Hum. Genet.*, **30** : 28-37.

Haley C.S., Knott S.A., Elsen J.M. (1994) Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics*, in press.

Hasstedt S.J. (1982) A mixed-model likelihood approximation on large pedigrees. *Comput. Biomed. Res.*, **15** : 295-307.

Knott S.A., Haley C.S. (1992) Maximum likelihood mapping of quantitative trait loci using full-sib families. *Genetics*, **132** : 1211-1222.

Knott S.A., Haley C.S., Thompson R. (1991) Methods of segregation analysis for animal breeding data : a comparison of power. *Heredity*, **68** : 299-311.

Lalouel J.M., Rao D.C., Morton N.E., Elston R.C. (1983) A unified model for complex segregation analysis. *Am. J. Hum. Genet.*, **35** : 816-826.

Lander E.S., Botstein D. (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121** : 185-199.

Le Roy P., Elsen J.M. (1994) Numerical comparison between powers of maximum likelihood and analysis of variance methods for QTL detection in progeny test designs. The case of monogenic inheritance. *submitted to Theor. Appl. Genet.*

Le Roy P., Elsen J.M. (1992) Simple test statistics for major gene detection : a numerical comparison. *Theor. Appl. Genet.*, **83** : 635-644.

Le Roy P., Elsen J.M. (1992b) Numerical comparison between analysis of variance and maximum likelihood methods for QTL detection in progeny test designs. *43rd Annual Meeting of the EAAP, Madrid (Spain) : 14-18 september 1992*, Commission on animal genetics, Session II : advances in molecular genetics, vol I, 68-69 (abstr.).

Le Roy P., Elsen J.M., Naveau J. (1990) Etude de la variabilité génétique de l'adiposité dans la lignée Laconic. *Journées de la Recherche porcine en France*, **22** : 11-16.

Le Roy P., Elsen J.M., Knott S.A. (1989) Comparison of four statistical methods for detection of a major gene in a progeny test design. *Génét. Sél. Evol.*, **21** : 341-357.

Le Roy P., Naveau J., Elsen J.M., Sellier P. (1990) Evidence for a new major gene influencing meat quality in pigs. *Genet. Res.,Camb.*, **55** : 33-40.

Lisitsyn N.A., Lisitsyn N.M., Wigler M.H. (1993) Cloning the difference between two complex genome. *Science*, **259** : 946-951.

Lisitsyn N.A., Segre J.A., Kusumi K., Lisitsyn N.M., Nadeau J.H., Frankel W.N., Wigler M.H., Lander E.S. (1994) Direct isolation of polymorphic markers linked to a trait by genetically directed representational difference analysis. *Nature Genetics*, **6** : 57-63.

MacLean C.J., Morton N.E., Yee S. (1984) Combined analysis of genetic segregation and linkage under an oligogenic model. *Computers and Biomedical Research*, **17** : 471-480.

Michelmore R.W., Paran I., Kesseli R.V. (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis : a rapid method to detect markers in specific regions by using segregating populations. *Proc. Natl. Acad. Sci., USA*, **88** : 9828-9832.

Morton N.E., (1955) Sequential tests for the detection of linkage. *Am. J. Hum. Genet.*, **7** : 277-318.

Morton N.E, McLean C.J. (1974) Analysis of family resemblance. III Complex segregation of quantitative traits. *Am. J. Hum. Genet.*, 26 : 489-503.

Motro U., Soller M. (1993) Sequential sampling in determining linkage between marker loci and quantitative trait loci. *Theor. Appl.. Genet.*, 85 : 658-664.

Niemann-Sorensen A., Robertson A. (1961) The association between blood groups and several production characteristics in three Danish cattle breeds. *Acta Agric. Scand.*, 11 : 163-196.

Ott J. (1991) *Analysis of human genetic linkage.* The John Hopkins University press, Baltimore.

Piper L.R., Bindon B.M. (1982) The *Booroola Merino* and the performance of medium *non – peppin* crosses at Armidale. *In : Piper L.R., Bindon B.M., Nethery R.D. (ed.), The Booroola Merino*, pp 9-20, CSIRO, Melbourne.

Ricordeau G. (1984) Major genes in sheep and goats. *In : 2th world congress on genetics applied to livestock production*, vol VI, pp 451-461, Madrid.

Ricordeau G., Bouillon J., Le Roy P., Elsen J.M. (1990) Déterminisme génétique du débit de lait au cours de la traite des chèvres. *INRA Prod. Anim.*, 3, (2) : 121-126.

Risch N. (1984) Segregation analysis incorporating linkage markers. I. Single-locus models with an application to type I diabetes. *Am. J. Hum. Genet.*, 36 : 363-386.

Rohrer G.A., Alexander L.J., Keele J.W., Smith T.P., Beatttie C.W. (1994) A microsatellite linkage map of the procine genome. *Genetics.* 136 : 231-245.

Soller M. (1978) The use of loci associated with quantitative traits in dairy cattle improvement. *Anim. Prod.* 27 : 133-139.

Soller M., Brody T., Genizi A. (1976) On the power of experiments to detect marker linked quantitative effects in crosses between inbred lines. *Theor. Appl. Genet.*, 47 : 35-39.

Soller M., Genizi A. (1978) The efficiency of experimental designs for the detection of linkage between a marker locus and a locus affecting a quantitative trait in segregating populations. *Biometrics*, **34** : 47-55.

Spearow J.L., Nutson P., Mailliard W., Fua L., McGee S., Barkley M. (1991) Mapping QTL controlling hormone induced ovulation rate in mice. *J. Anim. Sci.*, :

Teyssier J., Bosc P., Molénat G., Vincent M., Thimonier J., Elsen J.M., Boomarov (1992) First french results on productivity of prolific sheep including heterozygous Booroola ewes.*In : 43rd Annual Meeting of EAAP*, Madrid, Spain, 14-18 semptember 1992, vol I, pp 128-129, (abstr.).

Wald A. (1947) *Sequential analysis*. Wiley, New York.

Weller J.L., Kashi Y., Soller M. (1990) Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Science*, **73** : 2525-2537.

Recientes progresos en el rastreo de genes en especies domésticas: implicaciones en mejora animal

M.A. Toro y L. Silió
Area de Mejora Genética Animal
CIT-INIA, Madrid

Introducción

En su discurso de apertura del primero de los Congresos Mundiales de Genética Aplicada a la Producción Ganadera, celebrado en Madrid en 1974, asimilaba Miguel Odriozola el ataque experimental a los genes que gobiernan las funciones al rastreo de un animal por un cazador y diferenciaba en el mismo tres posibles estrategias, que corresponderían a los enfoques estadístico, fisiológico y molecular, aventurando para esta última (la búsqueda de la pieza en su madriguera de DNA), entonces incipiente, el logro de importantes capturas en el plazo de una década. Han pasado dos y, aunque con retraso, comienzan a cobrarse algunas piezas como sin duda tendremos ocasión de comprobar en el 5º de esta serie de Congresos Mundiales que tendrá lugar el próximo Agosto.

El propósito de esta comunicación es el de revisar el estado actual de este rastreo de genes, es decir, del desarrollo de mapas genéticos en especies ganaderas y de la identificación de loci con efecto en caracteres productivos así como discutir algunos de los problemas que plantea el uso de esta información molecular en programas de selección, introgresión génica o conservación.

Mapas genéticos de especies domésticas

Los instrumentos básicos en el rastreo de genes con efecto sobre caracteres cuantitativos (QTL) son los mapas de ligamiento, que definen las relaciones entre genes marcadores en términos de distancias génicas o frecuencias de recombinación. Los mapas físicos son descripciones de las relaciones físicas de los componentes de un genomio animal (cromosomas, bandas cromosómicas, genes, secuencias de bases) que pueden facilitar el desarrollo de los mapas de ligamiento. La metodología seguida en el levantamiento de mapas genéticos de especies de animales domésticos se inspira en parte en la orientación del Proyecto Genoma Humano y sus progresos son atribuibles a las siguientes causas:

- Disponibilidad actual de gran número de marcadores DNA (RFLPs, DNA hipervariable: mini y microsatélites), que permiten iniciar la construcción de mapas genéticos, limitada por las escasas decenas de marcadores tradicionales: genes codificantes de isoenzimas y grupos sanguíneos.

- Desarrollo de técnicas de mapeo físico: hibridación *in situ*, análisis de líneas híbridas de células somáticas y análisis de flujo clasificado de cromosomas, que permiten establecer grupos sinténicos de genes situados en un mismo cromosoma y localizar finalmente éste.

□ Empleo como familias de referencia de cruces (parentales ⇒ F· ⇒ F_2) entre poblaciones genéticamente divergentes (*Bos indicus x Bos taurus*; Large White x Jabalí o Meishan) que permiten una alta heterocigosidad y mayor información para la construcción de mapas (Soller y Beckmann, 1988).

□ Utilización de este material biológico por diversos laboratorios en proyectos coordinados a escala nacional (proyecto USA de mapa genético porcino) o internacional (PigMap europeo) que posibilitan formas más eficientes de división del trabajo.

Entre los mamíferos domésticos, los mayores logros en la construcción de mapas genéticos se han alcanzado en vacuno y en porcino.

Barendse et al. (1994) acaban de publicar un mapa genético del genoma bovino obtenido a partir del análisis de 295 inidividuos nacidos en un conjunto de 14 familias de referencia procedentes de cruces Bhrama x Angus (Texas), N'Dama x Boran (Kenia) y *Bos indicus x Bos taurus* (Australia). Aunque en vacuno se dispone de más de 300 loci asignados a cromosomas o grupos sinténicos, para este trabajo se han analizado 202 polimorfismos de DNA (144 microsatélites) de los que 171 se encuentran ligados en 36 grupos de ligamiento con una distancia genética media entre marcadores de 15 cM. Al menos uno de estos grupos ha sido asignado a 24 de los 29 pares de autosomas y al par de cromosomas sexuales. En la Tabla 1 se resumen el número de marcadores asignados a estos cromosomas, la heterocigosis media y longitud mapeada por cromosoma.

Tabla 1. Resumen de mapa de ligamiento del genoma vacuno (Barendse et al., 1994)

Cromosoma	N° de marcadores	Het. media	Longitud mapeada,cM	Cromosoma	N° de marcadores	Het. media	Longitud mapeada,cM
1	12	0.63	170	15	6	0.60	65
2	9	0.70	104	16	5	0.67	59
3	8	0.62	121	17	7	0.65	107
4	6	0.62	80	18	4	0.71	57
5	8	0.57	100	19	4	0.65	39
6	6	0.53	28	20	6	0.60	44
7	3	0.66	16	21	8	0.78	93
8	7	0.63	129	23	9	0.85	93
9	5	0.64	83	24	4	0.76	36
10	9	0.80	115	26	2	0.86	27
11	7	0.73	131	28	5	0.59	65
13	4	0.73	43	X	3	0.77	45
14	6	0.64	74				

Además hay 3 grupos de ligamiento asignados a grupos sinténicos (U12, U25 y U27) pero sin localización cromosómica y un último grupo (LG36) sin asignación física. La longitud total mapeada se estima en 2.513 cM, equivalente al 90% del genomio bovino (2.800 cM).

El cariotipo del cerdo tiene 18 pares de autosomas más el par X/Y de cromosomas sexuales. El proyecto más ambicioso de mapeo del genoma porcino es el PiGMap iniciado en 1991 por un grupo de 17 laboratorios de Bélgica, Dinamarca, Francia, Alemania, Italia, Holanda, Noruega, Suecia y Gran Bretaña. Sus objetivos son desarrollar un mapa genético que cubra un 90% del genomio con una distancia media entre marcadores de 20 cM, producir un mapa físico con al menos dos loci por brazo cromosómico e iniciar los experimentos necesarios para mapear QTL (Haley y Archibald, 1992). Las poblaciones utilizadas para la obtención de cruces F_2 de referencia en PiGMaP son de tres tipos: razas europeas comerciales (Large White y Pietrain), la raza china hiperprolífica Meishan y el Jabalí europeo. En tres laboratorios se está trabajando con Large White x Meishan y en otros dos con Jabalí x Large White y Jabalí x Pietrain.

En el proyecto americano de mapa genético porcino intervienen 12 centros experimentales y universidades norteamericanas con cruces con razas chinas como familias de referencia, objetivos similares de mapeo genético y físico y un particular énfasis en el desarrollo de métodos para preservar el material genético en líneas celulares, problema aún no resuelto en cerdos.

Un primer avance de los posibles resultados de estos proyectos en marcha lo constituye el mapa genético primario establecido recientemente a partir del análisis de 200 individuos F_2 procedentes de los cruzamientos Large White x Jabalí realizados en Suecia (Andersson et al., 1994). Se han utilizado un total de 117 marcadores (36 RFLPs, 59 microsatélites y 22 marcadores convencionales) que permiten mapear 1500 cM, que pueden representar un 75% del mapa de ligamiento porcino. De los 16 grupos de ligamiento obtenidos, 15 han sido asignados a cromosomas específicos, no disponiéndose en este estudio de marcadores de los cromosomas 10, 17, 18 y X. Un resumen de los datos de este mapa de ligamiento se presenta en la Tabla 2.

Tabla 2. Resumen de mapa de ligamiento del genoma porcino (Andersson et al., 1994)

Cromosoma	Nº de marcadores	Het. media	Longitud mapeada, cM	Cromosoma	Nº de marcadores	Het. media	Longitud mapeada, cM
1	4	0.60	43.3	9	11	0.48	75.8
2	5	0.78	44.3	11	4	0.31	9.5
3	9	0.52	86.3	12	5	0.67	72.8
4	10	0.85	102.0	13	6	0.66	102.6
5	6	0.75	81.4	14	8	0.47	67.0
6	9	0.57	79.2	15	4	0.67	35.9
7	10	0.67	99.3	16	6	0.60	65.0
8	8	0.55	47.8				

Rastreo de genes cuantitativos (QTL)

La utilización de cruces entre poblaciones genéticamente distantes favorece asimismo la detección de QTL. La historia de la mayor parte de las razas animales es la de un largo proceso de selección combinada con consanguinidad que lleva a los genes de efecto

importante a frecuencias alélicas extremas o a la fijación. En los cruces entre poblaciones seleccionadas en dirección divergente (Large White y Meishan, por ejemplo), o entre una población seleccionada y una salvaje (Large White y Jabalí) cabe esperar la mayor heterocigosidad para los QTL de mayor efecto sobre los caracteres seleccionados.

En el experimento sueco anteriormente citado (Anderson et al., 1994), existen marcadas diferencias entre las poblaciones parentales en caracteres de crecimiento y composición corporal, fruto de la intensa selección realizada en el cerdo doméstico, por lo que se ha realizado un análisis de QTL para estos caracteres a partir de registros obtenidos en los 200 animales F_2 y utilizando los 105 marcadores asignados a cromosomas (Tabla 2).

En el análisis, la información molecular es utilizada para calcular la probabilidad de que cada individuo F_2 haya heredado, de cada población, 2, 1 o ningún alelo para un QTL en una determinada posición del genomio. Si los parentales fueran líneas consanguíneas, los marcadores serían completamente informativos, por lo que estas probabilidades podrían calcularse a partir de sólo 2 marcadores flanqueando cada posición (Lander y Botstein, 1989). Pero en los casos, más frecuentes, las poblaciones parentales segregan para algunos marcadores, éstos pueden ser no informativos en individuos concretos, siendo preciso utilizar todos los marcadores del grupo de ligamiento para calcular las probabilidades de los diferentes genotipos posibles, asumiendo que en las poblaciones parentales se encuentran fijados alelos QTL alternativos (Haley et al., 1994).

En la Tabla 3 se presentan los resultados obtenidos con la aplicación de esta metodología. Los efectos dominante y heterocigoto de un QTL en una posición dada se han definido, respectivamente, como las desviaciones de los animales heterocigotos y homocigotos para el alelo procedente de Jabalí respecto a la media de ambos heterocigotos. Se estiman por regresión de los registros sobre los coeficientes de ambos efectos calculados a partir de las probabilidades genotípicas en cada posición. El procedimiento se repite a intervalos de 1 cM. Los valores presentados en la Tabla corresponden a las posiciones en que se ha encontrado un valor mayor, a lo largo del cromosoma, del estadístico F que expresa la significación de la regresión.

Tabla 3. Mapeo de QTL con efectos significativos sobre crecimiento y composicion corporal en individuos F_2 procedentes del cruce Large White x Jabalí (Anderson et al., 1994)

Carácter	*F*	Posicion Mapa	Efecto Aditivo	Efecto Dominante	Porcentaje Varianza F_2
		Cromosoma 4			
Crecimiento hasta 70 Kg (g/d)	11.8**	58	-23.5 ±4.9	0.7 ±7.3	11.9
Longitud intestino delgado (m)	11.1**	27	-0.87 ±0.18	0.03±0.27	11.3
Espesor grasa dorsal (mm)	18.0***	3	2.30 ±0.42	1.47±0.65	17.6
Porcentaje grasa abdominal (%)	19.4***	7	0.38 ±0.06	0.15±0.10	18.7
		Cromosoma 13			
Crecimiento hasta 30 Kg (g/d)	7.6*	53	-13.5 ±3.6	6.0 ±5.2	7.5

Estos resultados, aun limitados por el reducido número de individuos F_2 que sólo permite la detección de QTL de efecto grande, indican la existencia de QTLs en el cromosoma 4 con importantes efectos aditivos sobre crecimiento, longitud del intestino delgado y espesor de grasa dorsal. Asimismo se encuentra un QTL en el cromosoma 13 que afecta el crecimiento inicial del lechón. Todos ellos han debido contribuir a la respuesta a la selección para crecimiento magro efectuada en las razas europeas de cerdos.

Un enfoque alternativo al cruce de líneas divergentes, que posibilita el rastreo de QTL en una población de élite como el vacuno lechero Holstein-Friesian, se está desarrollando en los Estados Unidos por la empresa Genmark Inc. Un amplio avance de los resultados presentados en Birmingham con ocasión del 7º Congreso Internacional de Genética ha sido revisado por Rowlett et al. (1993). Para identificar QTL con efecto sobre la producción de leche se aprovechan las ventajas de un diseño abuelos/nietas, a partir de la estructura clásica del esquema de prueba de progenie, que genera registros y permite disponer de evaluaciones del mérito genético para los caracteres productivos analizados: kg de leche, grasa y proteina y porcentajes de grasa y proteina.

El trabajo supone la construcción de un mapa genético primario a partir de 1532 novillos probados, distribuidos en 14 familias de hermanos de padre con 33 a 208 novillos por familia. Para ello se tipificaron los 14 toros fundadores para 180 microsatélites, de los que en 159 se encontró al menos un toro heterocigoto. Posteriormente se tipificaron las familias informativas para cada marcador particular, estableciéndose grupos de ligamiento confirmados en paneles de celulas somáticas híbridas. Al menos un marcador fue identificado para cada uno de los 29 autosomas, con una media de 5 por cromosoma, mapeando el 60% del genoma.

En la identificación de QTL se emplearon métodos de máxima verosimilitud, analizándose separadamente cada familia y grupo de ligamiento. Los resultados indican la presencia de QTL con efecto sobre la produccion de leche en los cromosomas 1, 2, 6, 10 y 20, aunque la interpretación de los efectos no es simple: alguno de los QTL pudiera afectar al metabolismo de la lactosa y otro al desarrollo del tejido mamario. Cabe esperar resultados más satisfactorios cuando se disponga de mapas de mejor resolución.

Selección asistida por marcadores

Utilizando el desequilibrio de ligamiento poblacional

Hay varias causas que pueden originar desequilibrio de ligamiento entre marcadores y QTL: a) el cruce entre razas o líneas consanguíneas; b) mutaciones recientes; c) deriva genética y d) selección. De todas ellas la primera es la más importante y ha sido considerada en detalle por Lande y Thompson (1990). Se parte de una población en desequilibrio y suponemos que disponemos de muchos marcadores dialélicos codominantes. En una primera fase estimamos por regresión múltiple el valor mejorante aparente para cada marcador. La suma de ellos (o mejor aún la del conjunto más significativo) constituye lo que se denomina valor molecular (*molecular score*). Este valor molecular puede combinarse con el valor fenotípico del individuo utilizando la teoría de los índices de selección como se indica a continuación.

La mejor predicción del valor mejorante de un individuo utilizando tanto el fenotipo (z) como el valor molecular (m) vendrá dada por

$$\hat{u} = \text{cov}\,(u,\ z\ m) \begin{bmatrix} \sigma_z^2 & \sigma_m^2 \\ \sigma_m^2 & \sigma_m^2 \end{bmatrix}^{-1} \begin{bmatrix} z \\ m \end{bmatrix} = zh^2\,(1-p)/(1-h^2p) + m\,(1-h^2)/(1-h^2p)$$

siendo $p = \sigma_{m2}/\sigma_u^2$ la proporción de la varianza aditiva asociada a los marcadores.

La respuesta esperada a la selección será

$$R = i\sigma_{\hat{u}} = ih^2\sigma_z \sqrt{p/h^2 + (1-p)^2/(1-ph^2)}$$

fórmula que nos permite calcular su eficiencia relativa relativa respecto a la selección fenotípica, que es mayor para caracteres de heredabilidad baja y si una fracción considerable de la varianza genética está asociada a los marcadores como indica la siguiente tabla.

Tabla 4. Eficiencia de la selección por indice combinando datos fenotípicos y moleculares respecto a la selección fenotípica

	Proporción de varianza aditiva asociada a marcadores, p		
h^2	0.10	0.30	0.50
0.10	1.35	1.87	2.29
0.25	1.11	1.31	1.51
0.50	1.03	1.08	1.15

Estos resultados, excesivamente optimistas, han sido matizados posteriormente mediante estudios de simulación de Zhang y Smith (1993) y Gimelfarb y Lande (1994) a partir de una población cruce de dos líneas consanguíneas:

- la eficacia es mayor al aumentar el censo poblacional
- la eficiencia no siempre es mayor al aumentar el número de marcadores que contribuyen al índice de forma que la inclusión de un número excesivo de ellos puede disminuir la respuesta
- si la evaluacion genética se ha llevado a cabo con BLUP la ventaja de incluir información molecular es menos evidente ya que su efecto es, en cierta medida, análogo al de incrementar la heredabilidad
- puesto que el desequilibrio de ligamiento va disipándose con el transcurso de las generaciones es importante reevaluar la contribución de los marcadores cada generación.

En la Tabla 5 se presenta un ejemplo del tipo de resultados encontrados por Zhang y Smith (1993).

Tabla 5. Progreso genético utilizando desequilibrio de ligamiento ($h^2 = 0.25$; $p \approx 0.30$)

Generación	Solo marcadores	BLUP	Combinada
1	1.03	1.10	1.42
2	1.62	1.90	2.18
3	2.08	2.41	2.98
4	2.38	2.81	3.60
5	2.58	3.11	3.90
6	2.75	3.42	4.10
7	2.88	3.69	4.27
8	2.96	3.95	4.34
9	3.03	4.11	4.42

Utilizando el desequilibrio de ligamiento dentro de familias

Incluso si en una población existe equilibrio de ligamiento entre los marcadores y los loci cuantitativos hay desequilibrio de ligamiento dentro de las familias. Meuwissen y van Arendonk (1992) han estudiado la ventaja de utilizar marcadores para predecir con más precisión las desviaciones intrafamiliares en programas de selección de vacuno de leche y concluyeron que en un esquema de selección de prueba de progenie los marcadores apenas aportan ventajas ya que las desviaciones intrafamiliares están estimadas con precisión a partir de los datos individuales y de la progenie. Sin embargo en núcleos MOET en los que la selección se basa en la información registrada en ancestros y contemporáneos la repuesta a la selección puede incrementarse 10-20 %.

Futuro de la selección asistida con marcadores

Casi todos los especialistas consideran que la información molecular no va a sustituir a los programas de selección convencionales pero sí puede complementarlos. Esto es más evidente si consideramos que los actuales métodos de evaluación basados en la metodología del modelo mixto presentan un marco adecuado en el que incluir la información molecular tanto en poblaciones en equilibrio (Fernando y Grossman, 1990) como en desequilibrio de ligamiento via grupos genéticos (Goddard, 1992).

Algunos autores sin embargo, especialmente Smith y Smith (1993) consideran que el futuro depende no tanto del desarrollo de modelos estadísticos complejos y de métodos numéricos costosos computacionalmente que aprovechen la informacion proporcionada por marcadores mas o menos débilmente ligados a genes de interés económico sino del desarrollo de las técnicas de biología molecular que permitan, a partir de unos marcadores inciales mas o menos provisionales, localizar marcadores estrechamente ligados (1 a 2 cM) de forma que la selección se convierta efectivamente en selección directa para los QTL de interés económico. La utilización de la información molecular dependerá a medio plazo del abaratamiento de sus costes.

Introgresión asistida por marcadores

Se denomina introgresión al proceso de introducir en el genoma de una estirpe A, un gen favorable que exista en otra estirpe B. Y clásicamente se lleva a cabo mediante retrocruzamientos sucesivos con la estirpe A pero reteniendo únicamente aquello animales portadores del gen de interés. Obviamente la utilización de marcadores puede permitir:

- detectar los animales portadores del gen de interés, lo que resulta especialmente interesante si se expresa en un sólo sexo, o con penetrancia incompleta o tardíamente en la vida del animal o su detección requiere la destrucción del animal; la eficacia de se vera incrementada si disponemos de dos marcadores flanqueando al gen en cuestión

- recuperar lo más rápidamente el genoma A al elegir entre los portadores del gen de interés aquellos más similares al tipo A.

Las estrategias óptimas de utilización de marcadores en programas de introgresión han sido investigadas por varios autores y de forma destacada por Hospital et al. (1992). En la Tabla 6 se presenta un ejemplo del tipo de resultados que encuentran en una situación en la que consideran 20 cromosomas de 100 cM de longitud y se selecciona (proporción seleccionada= 10%) de acuerdo a dos marcadores situados a 20 cM de los loci teloméricos. Uno de los cromosomas es portador en el centro del gen de interés.

Tabla 6. Recuperación del genoma receptor en un proceso de introgresión

Generación	Introgresión sin marcadores		Introgresión con marcadores	
	Cromosoma portador	Cromosomas no portadores	Cromosoma portador	Cromosomas no portadores
RC1	0.607	0.750	0.636	0.817
RC2	0.684	0.875	0.747	0.943
RC3	0.741	0.938	0.855	0.983
RC4	0.784	0.968	0.891	0.993
RC5	0.817	0.984	0.903	0.996
RC6	0.842	0.992	0.913	0.998

Cromosoma portador es aquel donde se localiza el gen de interés para el que obviamente la recuperacion del genoma es más lenta mientras que cromosomas no portadores son los restantes. Como se observa en la tabla la introgresión asistida por marcadores puede suponer una ventaja de aproximadamente dos generaciones menos de retrocruzamientos. Además los autores indican que dos o tres marcadores por crosmosoma son suficientes e incrementar su número no implica un aumento de eficiencia.

Conservación asistida por marcadores

En un programa de conservacion de recursos genéticos, el objetivo es el mantenimiento de la variabilidad genética. Por ello se ha sugerido, en este tipo de programas, practicar selección utilizando como criterio la proporción de heterocigosidad en un conjunto de genes marcadores. Chevalet (1992) ha realizado algunos cáculos sobre esta propuesta y señala que el método es eficaz en poblaciones pequeñas ($N < 100$) siempre que se disponga de al menos tres marcadores por cromosoma y la proporción seleccionada sea baja (80-90%). A título de ejemplo con este sistema una población de $N=25$ mantiene en la generación 15 el mismo índice de heterocigosidad que se hubiera mantenido con un programa de elección aleatoria de los reproductores hasta la generación 10.

Perspectivas

Junto a los logros alcanzados hasta el momento son aún varios los problemas que plantea el uso de información molecular en mejora animal. Para abordar algunos de ellos, inicia sus trabajos una red de 13 grupos de investigación de 9 países europeos, que se ha fijado para los próximos 4 años las tareas siguientes:

- Métodos de mapeo de marcadores
- Mapeo de QTL en cruces entre poblaciones no consanguíneas
- Mapeo de QTL en poblaciones de hermanos y medios hermanos
- Mapeo de QTL con genealogías complejas
- Mapeo de QTL para caracteres no normales
- Analisis multivariante
- Genotipado selectivo y RAPDs
- Estudios de especies modelo
- Introgresión asistida por marcadores
- Selección asistida por marcadores en poblaciones panmícticas

Referencias

Anderson L. et al., 1994. **Science 263**: 1771-1774.
Barendse W. et al., 1994. **Nature Genetics 6**: 227-235.
Chevalet C., 1992. En *Eléments de génétique quantitative et application aux populations animales*, pp. 295-297, INRA Eds.
Fernando R.L. y Grossman M., 1990. **Genet. Sel. Evol. 21**: 467-477.
Gimelfarb A. y Lande R., 1994. **Genet. Res. 63**: 39-47.
Goddard M.E., 1992. **Theor. Appl. Genet. 83**: 878-886.
Haley C.S. y Archibald A.L., 1992. En *Genome Analysis* Volume 4: *Strategies for Physical Mapping*, pp. 99-129, CSHLP.
Haley C.S. et al., 1994. **Genetics 136**: 1195-1207.
Hospital et al., 1992. **Genetics 132**: 1199-1210.
Lande R. y Thompson R., 1990. **Genetics 124**: 743-756.
Meuwissen T.H.E. y van Arendonk J.A.M., 1992. **J. Dairy Science 75**: 1651-1659.
Rowlett K. et al., 1993. **AgBiotech News & Info. 5 (10)**: 337-360.
Smith C. y Smith D.B., 1993. **Anim. Breed. Abst. 61**: 197-204.
Zhang W. y Smith C., 1992. **Theor. Appl. Genet. 83**: 813-820.

LA DESCRIPCION GENETICA DE LOS CARACTERES CUANTITATIVOS

Carlos López-Fanjul de Argüelles

Departamento de Genética, Universidad Complutense, 28040 Madrid

La cualidad esencial de la herencia de los caracteres cuantitativos es su flexibilidad, esto es, que la expresión de los loci pertinentes pueda variar en relación con las características, asimismo variables, del medio en que se expresen. Sin embargo esta capacidad imposibilita, por su propia naturaleza, la clasificación genotípica y, por tanto, impide igualmente el análisis pormenorizado de la variación genética poblacional de un atributo en términos del número de loci segregantes, sus relaciones de ligamiento, el de alelos por locus, sus frecuencias y efectos sobre el carácter en cuestión y sobre otros, especialmente eficacia biológica, así como el tipo de acción génica en cada caso. Esta limitación obliga a la adopción de un postulado genético general, el mismo para cualquier carácter, cuya relación con la realidad concreta de cada uno de ellos muy rara vez se ha podido establecer. Este modelo, propuesto por Fisher en 1918, considera que la variación genética de un rasgo cuantitativo en una población se debe a la segregación de un número indefinidamente grande de loci neutros con efectos

infinitesimales, de ahí el que se le conozca por el nombre de "modelo infinitesimal". Su validez debe ser investigada experimentalmente y, en lo que sigue, expondremos los resultados obtenidos hasta la fecha en Drosophila melanogaster, el organismo piloto que ha proporcionado la mayor parte de los datos disponibles.

I. Experimentación clásica

A lo largo de los últimos cuarenta años se ha tratado de averiguar el grado de coincidencia entre lo esperado bajo el modelo infinitesimal y lo observado con respecto a diversos caracteres cuantitativos en poblaciones naturales. Para ello se han sometido éstas a dos tipos de manipulación: selección artificial o consanguinidad forzada, en experimentos inspirados en su mayoría por consideraciones teóricas desarrolladas por Alan Robertson, en especial las derivadas de la teoría de límites a la selección.

i) Número de loci

El procedimiento de Wright-Castle (y otros asimilables a él) para calcular el número de loci determinantes de las diferencias genéticas entre dos líneas para un carácter cuantitativo, suele redundar en estimas sesgadas e imprecisas, incluso en las condiciones más favorables. No obstante el modelo infinitesimal predice que la variabilidad de la respuesta aumentará gradualmente a medida que el proceso selectivo avanza. Por el contrario cuando el número de loci causantes de esa respuesta es reducido y, por tanto, sus efectos son considerables, la probabilidad de fijación del alelo favorable en cada locus se aproximará a la unidad y, en consecuencia, la variabilidad de la respuesta disminuirá

temporalmente hasta desaparecer.

ii) Magnitud de efectos alélicos y frecuencias iniciales

De la especificación del modelo infinitesimal se desprende que, en poblaciones panmícticas, la distribución conjunta de valores genotípicos en padres e hijos es normal bivariante. En consecuencia la correspondiente regresión es lineal y su valor no estará afectado por la práctica de selección, que sólo cambiará las frecuencias alélicas por infinitésimos. Por ello la respuesta a corto plazo es prácticamente independiente de la situación genética subyacente y, en la práctica, lo mismo ocurre con la respuesta a largo plazo siempre que sea simétrica y se trate de caracteres neutros. Sin embargo la comparación de los límites alcanzados en poblaciones que han sido sometidas o no a una reducción previa de su censo efectivo, puede aportar datos de interés acerca de las frecuencias iniciales en el caso de que una porción importante de la variación original se deba a loci con efectos grandes mientras que, si éstos son pequeños, la respuesta sólo es función de la magnitud del estrangulamiento del censo y no depende de las frecuencias génicas. Por otra parte el efecto de un periodo previo de selección en un sentido sobre la respuesta obtenida posteriormente en el opuesto, depende básicamente de las frecuencias de partida, si la duración del periodo previo es corta o de la magnitud de los efectos alélicos, si es larga.

iii) Acción génica

Los límites alcanzados mediante diferentes procedimientos de selección dependen del tipo de acción génica y, por ello, su comparación puede resultar ilustrativa en este respecto. Por ejemplo, la selección recíproca recurrente sólo superará a la selección en estirpe cerrada cuando exista acción génica

sobredominante, al menos en algunos de los loci implicados. Por otra parte los esquemas cíclicos de subdivisión y selección interlíneas conllevan límites inferiores a los alcanzados por selección en una única línea, salvo que existan alelos recesivos con frecuencia inicial baja o epistasias importantes. Sin embargo al descender al campo empírico la replicación de estos experimentos es obligadamente escasa o nula, con la consiguiente pérdida de precisión estadística que obscurece la interpretación de los resultados. En este orden de cosas es mucho más esclarecedor el estudio de los cambios experimentados por las distribuciones de medias y varianzas aditivas de un carácter en un conjunto de líneas derivadas del mismo origen y sometidas todas ellas a la misma deriva. Se dispone de predicciones para los casos en que los efectos de los loci son independientes o cuando la epistasia es aditiva x aditiva.

iv) Pleyotropía y neutralidad

El grado de pleyotropía del conjunto de loci considerados puede investigarse a partir de estimas de correlaciones genéticas puntuales o realizadas. Sin embargo el efecto pleyotrópico de mayor interés es el que se refiere a eficacia biológica. Si el carácter es neutro (periférico) desplazamientos moderados de la media en ambos sentidos (obtenidos mediante selección artificial divergente) no acarrearán disminución de la eficacia biológica promedio y, por tanto, la respuesta se mantendrá una vez desaparecida su fuerza causante, mientras que en cualquier otro caso la relajación de la selección resultará en un retorno de la media hacia su valor primitivo, aunque de esta observación no pueda deducirse el tipo de relación causal entre carácter y eficacia.

v) Resultados

Los dos caracteres que han sido estudiados con mayor detalle son el número de cerdas abdominales y esternopleurales en D. melanogaster. Aplicando las distintas técnicas mencionadas (y otras) se ha podido averiguar que la mayor parte de la variación genética de estos atributos en poblaciones naturales se debe a la segregación de unos pocos loci semineutros, del orden de una decena, con efectos aditivos grandes y frecuencias iniciales intermedias. El resto de los caracteres comunmente utilizados en experimentos de selección artificial no ha sido examinado con la misma intensidad pero, en la medida en que ha sido así, parecen compartir las propiedades de los anteriores. Con todo la homogeneidad de este conjunto no es necesariamente extrapolable a otros, por cuanto se trata de un grupo de rasgos elegido por los investigadores por poseer a priori las propiedades aludidas, en particular heredabilidad media o alta y depresión consanguínea escasa o nula. Por otra parte la variabilidad genética poblacional de los componentes principales de la eficacia, en particular fecundidad, viabilidad y éxito en el apareamiento, está en buena parte generada por la segregación de alelos recesivos desfavorables y, por tanto, con frecuencias iniciales bajas, aunque los escasos datos disponibles coinciden en señalar la presencia adicional de unos pocos loci con efectos aditivos considerables que pudieran mantenerse segregando con frecuencias intermedias, como consecuencia de pleyotropía antagónica o de la acción de otros mecanismos con consecuencias semejantes.

II. Experimentación neoclásica

Los estudios teóricos de William G. Hill y colaboradores sobre

el comportamiento poblacional de mutaciones con efecto sobre caracteres cuantitativos, han propiciado la obtención de datos experimentales referentes a dichas mutaciones. El procedimiento utilizado sólo difiere del clásico en que el material de partida es totalmente homocigótico, en lugar tratarse de una muestra de una población natural genéticamente variable, pero los métodos utilizados para su análisis siguen siendo los mismos: selección natural o consanguinidad forzada. Cualquier cambio permanente de la media de un carácter en una línea debe, pues, atribuirse a la aparición de una mutación que se fija por deriva o alcanza su máxima frecuencia posible por selección. Mientras que la experimentación clásica sólo proporcionaba información promediada sobre el conjunto de los genes implicados en la respuesta a la selección o la depresión consanguínea observadas, los neoclásicos permiten ciertos progresos en la descripción genética de atributos cuantitativos, tales como la caracterización de la distribución de efectos de las mutaciones sobre el carácter o caracteres de interés, el cálculo del aporte de variación por generación debido a mutación y, además, establecer el efecto y acción génica de mutaciones concretas sobre cualquier carácter. Los dos métodos experimentales aludidos son complementarios. Así los datos procedentes de líneas seleccionadas proporcionan una descripción de la distribución de efectos de las mutaciones tal y como éstas aparecen, mientras que los correspondientes a líneas consanguíneas se aproximan más a la situación resultante tras la acción de la selección natural.

i) Heredabilidades de mutación

Recibe este nombre el incremento de variación debido a mutación por generación expresado como fracción de la varianza ambiental

del carácter (σ^2). Las estimas estadísticamente precisas son escasas y corresponden casi exclusivamente a D. melanogaster, con todo disponemos de datos referentes tanto a atributos morfológicos como a componentes principales de la eficacia. Todas ellas son muy semejantes y oscilan entre $0,5 \times 10^{-3}$ y $3,5 \times 10^{-3}$. Debido fundamentalmente a que muchas de estas mutaciones no son neutras, las cifras anteriores son subestimaciones que, de acuerdo con los cálculos de Caballero y Keightley, deben multiplicarse por un factor del orden de tres. En cualquier caso ponen claramente de manifiesto la importante influencia sobre la respuesta a la selección de las mutaciones ocurridas a lo largo del proceso.

ii) Distribución de efectos de las mutaciones

En general su forma se aleja de la esperada bajo el modelo infinitesimal, presentando asimetrías de distinto signo dependiendo del carácter estudiado y, además, leptocurtosis. En estas circunstancias las mutaciones con efectos muy grandes o muy pequeños son más frecuentes de lo esperado en el caso de normalidad, para un valor prefijado de la varianza aportada por mutación. En consecuencia la respuesta a la selección artificial depende del cuadrado medio de los efectos de las mutaciones cuyo signo es el mismo que el del sentido de selección seguido y la variación de la respuesta se hace proporcional a la curtosis de la distribución de efectos.

iii) Análisis de mutaciones individuales

Este estudio está limitado por el poder de resolución experimental que impone el efecto mínimo detectable de las mutaciones ($0,1-0,2\ \sigma$). Aparecen aquí claras diferencias entre los dos grupos de caracteres que venimos considerando. Los efectos de las mutaciones sobre caracteres morfológicos pueden ser de

distinto signo y de acuerdo con la magnitud de su efecto, tanto sobre el propio atributo como sobre eficacia, pueden dividirrse en tres clases: letales con efecto en heterocigosis sobre el carácter de magnitud variable que, en experimentos de selección, pueden ascender a un 40% del total; mutaciones de efecto muy grande ($>\sigma$), casi sin excepción recesivas y perjudiciales; y el resto, de efecto menor ($<\sigma/2$), en buena parte semineutras y con acción génica variable aunque predomine la aditividad. De aquí el que una importante fracción de la variación genética de estos rasgos en poblaciones naturales pueda deberse a la segregación de unos pocos loci con efectos aditivos considerables. En contraposición las mutaciones con efecto sobre componentes principales de eficacia suelen ser perjudiciales, con la excepción de las que influyen sobre fecundidad, preponderando la acción génica recesiva.

iv) Pleyotropía

Los datos referentes a mutaciones individuales revelan una pleyotropía muy amplia, de tal manera que sólo en muy escasas ocasiones tienen efecto sobre uno sólo de los caracteres estudiados. Sin embargo las observaciones están limitadas a un grupo de mutaciones caracterizado por tener efectos muy considerables, al menos sobre uno de los atributos en custión y, por tanto, su extrapolación es arriesgada. Por otra parte el tipo de acción génica de una determinada mutación presenta una fuerte tendencia a conservarse con respecto a los distintos caracteres sobre los que ésta influye, de acuerdo con la predicción del modelo Kacser-Burns.

III. Conclusión

Los datos correspondientes a mutaciones con efecto sobre caracteres cuantitativos confirman los obtenidos anteriormente analizando su variabilidad genética en poblaciones naturales y proporcionan una descripción más precisa de ésta. Reiteradamente apuntan hacia situaciones muy distintas de las supuestas en el modelo infinitesimal. Si los datos procedentes de D. melanogaster pueden ampliarse a las especies domésticas, y hasta ahora nunca ha dejado de ser así, las predicciones de la respuesta a la selección natural y su variabilidad estarían ambas sesgadas, resultando en sobrevaloración la primera y en infravaloración la segunda. Los resultados proporcionan, además, una comprobación empírica de la importancia de la contribución de nuevas mutaciones a la respuesta a la selección artifical a distintos plazos.

STATISTICAL MODELS FOR THE DETECTION OF GENES CONTROLLING QUANTITATIVE TRAIT LOCI EXPRESSION (*)

Carbonell, E.A., Asíns M. J.

Instituto Valenciano de Investigaciones Agrarias

Apartado Oficial

46113 Moncada, Valencia

1. Introduction

2. Statistical Methodologies

2.1. A marker at a time

2.1.1. Linear models: Comparison of means and ANOVA

2.1.2. Maximum likelihood estimation

2.2. Pairs of markers

2.2.1. Linear and non-linear models

2.2.2. Maximum likelihood estimation

2.3. All markers simultaneously

3. Power studies

4. Perspectives and further remarks

(*) Este trabajo es una versión reducida del que se publicará en: "In vitro haploid production in higher plants". Jain SM and Sopory SK (Eds.). Kluwer Academic Publisher, Dordrecht, The Nederlands.

1. Introduction

Many important traits in plant breeding exhibit continuos variation (yield, maturity, biotic and abiotic stress tolerance, etc.) and therefore are named metric, polygenic, or quantitative traits (QTs). The genetic principles underlying their inheritance are basically the same than those affecting Mendelian or qualitative traits, but since the segregation of the genes concerned could not be followed individually, new methods of study had to be developed and new concepts introduced. Due to the continuos variation, genetic studies of QTs relied upon statistics to describe the nature and characteristics of this variation. These statistical models are based on the assumptions that "many" loci are involved in the expression of the trait and that the effect of the individual genes is "very small" compared to the environmental effects. Consequently, a branch of the science of Genetics based on these models for quantitative traits grew up, the Biometrical or Quantitative Genetics. However, Sax (1923) was the first to introduce the concept of using qualitative genes to locate genes of lesser effects controlling QTs. Thoday (1961) used single gene morphological markers to conduct detailed studies of QTs in *Drosophila melanogaster*. More systematic attempts to resolve QTs into their individual genetic components were initially limited by the scarceness of polymorphic qualitative markers that could cover large parts of the genome. These limitations were partly overcome by the use of isozymes (Tanksley et al., 1982; Edwards et al., 1987), later by restriction fragment length polymorphisms (RFLPs) which have provided a major source of markers with many desirable attributes (Paterson et al., 1988, 1991) and lately by RAPDs (random amplified polymorphic DNA) and "microsatellites". Thus, genetic markers (isoenzymatic loci, RFLPs, RAPDs, etc.) linked to loci affecting a QT of interest can be used to follow their individual segregation. The availability of detailed linkage maps of molecular markers makes possible to dissect QTs into discrete factors, called QTLs (Quantitative Trait Loci) by Gelderman (1975); then, when a high enough number of markers are scored in a family segregating for a given polygenic trait, accurate estimates of genic effects and QTL locations can be obtained. Molecular markers have been used to study

many QTs in several species. In all of them, genes controlling the character of interest have been detected what probes that this approach is feasible. Thus, great efforts are recently being focused on constructing saturated linkage maps based on molecular markers to locate genes and QTLs affecting important traits by means of specific statistical methods. Four types of experimental designs are generally used in these studies: F_2's, backcrosses, recombinant inbred lines and doubled haploid lines. The first two designs are the most frequently used for gene mapping mainly due to the less time involved in annual crops, but the last two allow unlimited replications. Some annual crops have been extensively studied (tomato, corn, etc.) others, like perennial crops or forest trees show inherent complexities like long juvenile periods making their studies to need another approach.

Therefore, in this paper we will discuss the different statistical models used for QTL analysis devoting special attention to the power of the different methods to detect the presence of a putative QTL. Comments about the unresolved problems and the future of QTL analysis will be also presented

2. Statistical methodologies

QTL analysis have been approached using different strategies depending on the number of markers involved in the analysis. Early studies considered the relationship between a single marker and a QTL; later, models considered a pair of markers flanking the QTL and studied the association between the QTL and the interval defined by the flanking markers. More recently, statistical methodologies focus their attention to the whole linkage group considering all markers of the group as being associated with the QTL.

Although any type of classification of statistical approaches is always incomplete and biased, we will try to describe the current statistical models being organized under two types of view-points; first, according the number of markers being studied, and then by the basic statistical methodology being employed.

2.1. A marker at a time.

2.1.1. Linear models: Comparison of means and ANOVA

Traditional methods to estimate the association between a marker locus and a quantitative trait are based on a very simple and appealing approach: Individuals are classified according to the genotype of the marker; if the means for the quantitative trait for those groups are not statistically different it is an indication that the classification into groups was somehow arbitrarious as far as the quantitative trait is concerned; therefore, the marker and the QTL are independent (i. e., no QTL is linked with the marker). Conversely, if means are statistically different, it is an indirect proof of the association between the grouping structure (the genetic marker) and the quantitative trait. So, a group of statistical methodologies are based on the comparison of means or ANOVA approach. Formally, the theory behind this intuitive approach is as follows:

Let us consider a population derived from a cross between two inbred lines. Denoting by M the marker locus and Q the locus involved in the expression of a quantitative trait. Consider two codominant alleles for the marker locus and that both loci are linked, r being their recombination fraction (2r would be the probability of crossing over during meiosis) independent of the sex of the parents from which the gametes were produced. Defining *a* and *d* after Falconer (1989) as the genotypic value of the homozygote and heterozygote genotypes of the Q locus, the phenotypic value of an individual i could be expressed for a given QTL by the following general linear model (Carbonell et al., 1992):

$$Y_i = \mu + aX_i + d(1 - X_i^2) + \varepsilon_i \qquad i = 1,\ldots,n \qquad (1)$$

where X is a dummy variable taking values of +1 (for the dominant homozygote), 0 (for the heterozygote) or -1 (for the recessive homozygote) in such a way that the expected genotypic values are:

$$\begin{aligned} \mu_{Q_1Q_1} &= \mu + a \quad (X = +1) \quad \text{for} \quad Q_1Q_1 \\ \mu_{Q_1Q_2} &= \mu + d \quad (X = 0) \quad \text{for} \quad Q_1Q_2 \\ \mu_{Q_2Q_2} &= \mu - a \quad (X = -1) \quad \text{for} \quad Q_2Q_2 \end{aligned} \qquad (2)$$

and ε_i is random variable that is assumed to be normally distributed with mean 0 and variance σ^2. This general model can be particularized for any mating design.

F_2 and backcross populations have been extensively studied by many authors (Soller and Beckmann, 1983; Kahler and Wehrhahn, 1986; Edwards, et al., 1987; Weller, 1987). Most of them used the ANOVA approach and presented the frequencies of the QTL genotypes for each genotype of the marker (Table 1), and the expected values of the means of the quantitative trait for each marker class (Table 2).

Table 1. Frequencies of QTL genotypes within each genotypic class of the marker for an F_2 population.

Marker genotype	Q_1Q_1 $[\mu + a]$	Q_1Q_2 $[\mu + d]$	Q_2Q_2 $[\mu - a]$
M_1M_1	$(1-r)^2$	$2r(1-r)$	r^2
M_1M_2	$r(1-r)$	$(1-2r+2r^2)$	$r(1-r)$
M_2M_2	r^2	$2r(1-r)$	$(1-r)^2$

Table 2. Means and variances of the QTL within each genotypic marker class for F_2 populations.

Marker genotype	Means	Variances [a]
M_1M_1	$\mu_{M_1M_1} = \mu + a(1-2r) + 2dr(1-r)$	$2a^2r(1-r) + 2d^2ru - 4adr(1-3r+2r^2)$
M_1M_2	$\mu_{M_1M_2} = \mu + d(1-2r+2r^2)$	$2a^2r(1-r) + 2d^2ru$
M_2M_2	$\mu_{M_2M_2} = \mu - a(1-2r) + 2dr(1-r)$	$2a^2r(1-r) + 2d^2ru + 4adr(1-3r+2r^2)$

(a) $u = 1-3r+4r^2-2r^3$

If both loci segregate independently (r = 0.5), the means of the quantitative trait for each genotypic marker class are all equal; therefore, a way to test if r = 0.5 is by testing the null hypothesis of equal class means.

However, none of the above studies went a step further calculating the second order moments. Asíns and Carbonell (1988) analytically showed that, even if all of the genotypes of the QTL have equal variance, variances of the QTL within each marker class in an F_2 population are different due to dominance and/or linkage between the marker locus and the QTL (see table 2) and hence, standard ANOVA procedures should be avoided in this case. It is recommended that the variance-covariance matrix should be used in the ANOVA analysis or, as a simpler alternative, calculate a t-test comparing the means of both homozygotes using non-pooled variances. Variances will be equal in the case of no association (which is precisely what we want to test) or under no dominance of the trait. Therefore, with this type of experimental data, if there is no previous knowledge of the type of gene action, the first step would be to test the homogeneity of variances. If it is significant, it could be an indication of linkage and/or dominant gene action. If this test is not significant but means are, it could be an indirect proof that there is not dominance and that the QTL is associated with the marker. If data are pooled from both types of backcrosses, the same problem of unequal variances arises (Asíns and Carbonell, 1988). This fact was experimentally noted by Edwards et al. (1987), Stuber et al. (1987) and Weller et al. (1988). Zhuchenko et al. (1979) also found that partial linkage between a QTL and a genetic marker should result in a difference in skewness between the F_2 groups homozygous for the opposite alleles of the marker. The homozygote with the higher mean should have the lower skewness value.

For the F_2 population, the contrasts of interest to estimate the genetic effects are:

$$\mu_{M_1M_1} - \mu_{M_2M_2} = 2a(1-2r)$$

$$2\mu_{M_1M_2} - (\mu_{M_1M_1} + \mu_{M_2M_2}) = 2d(1-2r)^2 \qquad (3)$$

Since the quantity in parenthesis is always equal or lower than one, the bias in the estimate of the dominant effect is larger than that of the additive effect.

Several other mating designs using replicated progenies like S_1, recombinant inbred lines, and crosses of the above populations with another inbred line have been studied by Cowen (1988) and Dudley (1992). Others like daughter and granddaughter design were introduced by Weller et al. (1990).

Unless the QTL shows complete dominance (a = d), there is no way to distinguish between a QTL tightly linked having a small genetic effect from another with a large effect but loosely linked. Hence, this type of approach could be considered as preliminary or exploratory in order to at least be able to detect the presence of a QTL. In order to have unbiased estimates of both the effects and the recombination fraction, maximum likelihood techniques should be used.

2.2.2. Maximum likelihood estimation

To have a separate estimates of the recombination fraction and the biometrical parameters of the QTL, alternative methods should be used. Maximum likelihood is a powerful tool for certain problems of estimation; however, it requires that the type of distribution be known. Many studies apply the central limit theorem and assume an underlying normal distribution of the trait unless there is reason to assume otherwise. Weller (1986) used a combination of moments method and maximum likelihood estimation to obtain an estimate of the seven parameters involved in an F_2 population (the recombination fraction plus the three means and three variances). The moments method allowed to reduce form seven to three the number of parameters to be estimated. The three parameters are obtained by trial and error.

For a doubled haploid population, the method is as follows (Luo and Kearsey, 1991): Let Y_i and Y_j denote trait values for two random individuals belonging to the maker genotypic classes M_1M_1 and M_2M_2, respectively. The likelihood of the entire population can be written as:

$$L = \prod_{i=1}^{n_1} f(Y_i) \prod_{j=1}^{n_2} f(Y_j) \qquad (4)$$

where L is the likelihood function; $f(Y_i)$ is the density function of Y_i; and n_1 and n_2 are the total number of individuals in each genotypic marker class. In turn, the density functions are:

$$f(Y_i) = (1-r)f(Q_1Q_1) + rf(Q_2Q_2)$$
$$f(Y_j) = rf(Q_1Q_1) + (1-r)f(Q_2Q_2) \qquad (5)$$

That is, each density function is expressed as the sum of the two disjoint events: since the linkage between marker and QTL is incomplete, the individuals within each marker class are a mixture of the two genotypes at the QTL, i. e. Q_1Q_1 and Q_2Q_2. Hence, for a given genotypic marker class, the sampled individual is either Q_1Q_1 or Q_2Q_2. In turn, the probability of each event is a set of conditional probabilities: the product of the probability of having that particular genotype at the Q locus given the genotype of the marker (either M_1M_1 or M_2M_2), taken from table 1 times the marginal density of Y_i or Y_j, conditional on the fact that the individual has that genotype at the QTL (f(QQ)). In order to obtain the form of the latter density, the underlying distribution of the genotypes at the QTL in the entire population (not within marker class) is assumed to be normal with means $\mu_{Q_1Q_1}$ and $\mu_{Q_2Q_2}$ and variances $\sigma^2_{Q_1Q_1}$ and $\sigma^2_{Q_2Q_2}$. The variances include not only environmental variation but also genetic variation at other loci (QTL) affecting the quantitative trait.

For, say, $f(Q_1Q_1)$ the conditional density will have the form:

$$f(Q_1Q_1) = (2\pi\sigma_{Q_1Q_1})^{-\frac{1}{2}} \exp\left(-\frac{(Y_i - \mu_{Q_1Q_1})^2}{2\sigma^2_{Q_1Q_1}}\right) \qquad (6)$$

In order to obtain the estimates of the recombination fraction and the means and variances, partial differentials of the log of the likelihood function L for all five parameters are set to zero and solving the resultant five equations. If an analytical

solution is not readily evident, iterative solutions are available, but they tend to be cumbersome and time consuming and not necessarily provide good solutions.

2.2 Pair of markers

The basic idea about using a pair of markers flanking the putative QTL was put forward by Lander and Botsein (1986). Once a map of markers is available, we test whether a QTL lies in an interval of known size between two adjacent markers.

If the putative QTL is located within the interval defined by the two marker locus A and B with known recombination fraction between them equal to p, we denote by r and r' the recombination fractions to the left and right markers. If the distance between the two markers is very small, one may assume that the frequency of double cross-over is negligible; then $p = r + r'$. If not, $p = r + r' - 2rr'\delta$, where δ is the coefficient of coincidence. Under no interference, $\delta = 1$, hence using the Haldane function $p = r + r' - 2rr'$

2.2.1 Linear and non-linear models

Similarly to the single marker case (see table 2), one could obtain expressions for the means of the quantitative trait for each genotypic class of the pair of markers. Here, the number of different genotypes for the marker loci is four for doubled haploid and backcross population and nine for F_2. The expected frequencies for each genotype of the QTL given the genotypic array of the marker loci are given in Table 3 for doubled haploids.

Knapp et al. (1990) described several statistical models and estimators of the recombination frequencies between marker loci and the QTL and the means of the QTL for the case of no double crossing-over. Under this situation, expressions are very simple because in two cases, the genotype of the QTL is directly estimated from the marker loci.

The problem with this approach is that if the interval between the two flanking markers is not small, the assumption of no double crossing over is not realistic. In this case, all marker classes are comprised of mixtures of the two genotypes of the QTL and the nonlinear model is not as simple.

Table 4. Marginal marker genotypic frequencies and expected and conditional frequencies of the two types of QTL genotypes given the genotypes of the flanking markers for a doubled haploid population.

		Conditional frequencies			
		Under Haldane's function		Under no double cross-over	
Marker genotypes	Marginal	Q_1Q_1 $\mu_{Q_1Q_1}=[\mu+a]$	Q_2Q_2 $\mu_{Q_2Q_2}=[\mu-a]$	Q_1Q_1	Q_2Q_2
A_1B_1/A_1B_1	(1 - p)/2	(1 - s)	s	1	0
A_1B_2/A_1B_2	p/2	(1 - t)	t	r'/p	r/p
A_2B_1/A_2B_1	p/2	t	(1 - t)	r/p	r'/p
A_2B_2/A_2B_2	(1 - p)/2	s	(1 - s)	0	1

p is the recombination fraction between the markers, s = rr'//1 - p), t = r(1 - r')/p, (assuming Haldane's function); and r and r' are the distance of the putative QTL to the left (A) and right (B) marker, respectively

The model is this case is

$$Y_i = \left[(1-s)\mu_{Q_1Q_1} + s\mu_{Q_2Q_2}\right]X_{1i} + \left[(1-t)\mu_{Q_1Q_1} + t\mu_{Q_2Q_2}\right]X_{2i} + \left[t\mu_{Q_1Q_1} + (1-t)\mu_{Q_2Q_2}\right]X_{3i} + \left[s\mu_{Q_1Q_1} + (1-s)\mu_{Q_2Q_2}\right]X_{4i} + \varepsilon_i \quad (7)$$

where Y_i is the phenotypic value of the quantitative trait for the ith individual and the X's are coding variables taking values of 1 or 0 depending on the genotypes of the four marker loci (i. e., X_{1i} will take value 1 if the genotype of the flanking markers is $A_1A_1B_1B_1$ and 0, otherwise).

Remembering that s and t are functions of r, r' and p, and that r' could be expressed in terms of r and p, the above model has only three unknowns (one recombination fraction and two means); so, only three partial derivatives should be computed.

Similar approach was used by Haley and Knott (1992) but in order to overcome the non-linearity problem, the estimates of the genetic effects for a given r are obtained by standard linear regression methods. Then, letting r to vary, the position which gives the best fitting model is the most likely position of the QTL.

The above methods are based in the calculation of the expectation of the mean genotypic effect of a QTL conditional on the marker genotype but in none of them the variance within each marker class is calculated and checked for homogeneity prior to continue with the linear model estimation. It has been shown by Balansard and Carbonell (unpublished data) that in fact variances are different for F_2's. Then, the estimators of the parameters (recombination fraction and means) using standard least squares not including the variance structure, although unbiased, are not consistent and have larger asymptotic variances. That will imply that, due to large mean square errors, some QTL with relatively small effects may be undetected by these methods.

Martinez and Curnow (1992) used a different approach. The linear model was based not in the expected means of the trait but in the probability that a given allele is present in an individual from a particular marker group. For a backcross population, if Y_{ij} is the observation for the quantitative trait on the jth plant (j= 1, 2, ...n) in the ith marker class (i = 1,.., 4), the expectation of Y_{ij} is given by:

$$E(Y_{ij}) = -a + (a+d)G_i(Q_1) = \beta_0 + \beta_1 G_i(Q_1) \qquad (8)$$

where $G_i(Q_1)$ is the conditional probability that an allele Q_1 is present in an individual from the ith marker class genotype. That conditional probability can be easily obtained by standard mendelian genetics procedures as a function of p and r and r'. $G_i(Q_1)$ is a function of the unknown parameter r. Hence, the parameters of the linear model can be obtained for a fixed r. Therefore, for a putative QTL at a given position $r = r_0$ we obtain a fitted value for the phenotype:

$$\hat{Y}_{ij} = \hat{\beta}_0(r_0) + \hat{\beta}_1(r_0)G_i(Q_1, r_0) \qquad (9)$$

A measure of the fit of the model is given by the residual sum of squares of the regression for a given value of r, r_0. If we graph this sum of squares against the corresponding chromosome positions letting r to vary, we obtain a measure of the evidence of the presence of a QTL in the interval defined by the markers involved if a clear minimum is found. If the presence of such a QTL is accepted by the significance of the regression model, we can select as an estimate of its position that value of r that minimizes the sum of squares, say r_m. If there is no segregation distortion in the cross, then r_m is a consistent estimator of r, and $E\left[\hat{\beta}_0(r_0)\right] = -a$ and $E\left[\hat{\beta}_1(r_0)\right] = (a+d)$.

The authors presented several numerical values for different genetical situations. If only one QTL was in the linkage group, the method gave consistent estimates of both position and size of effects. However, when two linked QTL are located in two non contiguous intervals, the method is not longer consistent for the positions nor the effects and a "ghost" QTL could be identified in a region neighboring the intervals where the QTL were actually present. If the effects of both linked QTLs are of similar magnitude but of opposite sign, then the ghost effect will not be present but the size of the effects will be underestimated and their positions moved closer to the outside markers, or even remain undetected. Hence, when two QTLs are linked, the sum of squares function may present a minimum somewhere in between the two QTLs. This ghost effect can be wrongly interpreted as a true QTL.

2.2.2. Maximum likelihood.

The simplest case to obtain maximum likelihood estimates of the recombination frequencies between a QTL and a pair of flanking markers was studied by Jensen (1989) involving only two markers. He applied this method to data from chromosome-doubled monoploid lines of barley originating from the F_1 generation of a cross between two well-adapted lines of barley. The lines segregated for two marker loci, DDT resistance and short rachilla hairs located on the same chromosome; the quantitative trait was single-kernel weight.

If genetic markers have been scored throughout the genome Lander and Botsein (1989) developed a method based in the interval mapping in order to find a QTL at any given location in the linkage group. The approach integrates information on all employed markers into one likelihood map. The integration of all markers makes this method very appealing, giving detailed insight into the part of the genome being investigated. The method is based in the linear model of equation (1). For an additive trait, *a* represents the estimated phenotypic effect of a single allele substitution at a putative QTL. Lander and Botsein (1989) used the LOD score to test the presence of a QTL: the estimates are compared to the constrained estimates under the null hypothesis of H_0: a = 0 corresponding to the hypothesis that no QTL is linked, the LOD score is then defined as:

$$LOD = \log_{10}(L(\hat{\mu},\hat{a},\hat{\sigma}^2)/L(\tilde{\mu},0,\tilde{\sigma}^2)) \qquad (10)$$

where $\hat{\mu}, \hat{a}, \hat{\sigma}^2$ are maximum likelihood estimates of the parameters and $\tilde{\mu}, \tilde{\sigma}^2$ are constrained maximum likelihood estimates obtained under H_0. However, distributional properties of this test statistic are not completely known. An alternative way of testing the presence of the QTL (i. e., a = d = 0) is given by the significance of the linear model in equation (1) using the Fisher-Snedecor F statistic (Carbonell et al., 1992).

If the QTL is located exactly at one of the two markers defining the interval, the problem is easily solved by standard regression procedures because in this case, the genotype of the QTL is known to be that of the marker; so, the values of X_i in equation (1) are known. However, if the QTL is located somewhere between the flanking markers (which is the normal situation), these values cannot be directly obtained but only their probability distributions. In this case, the likelihood of the parameters $\theta = (\mu, a, d, \sigma^2)$ based on the observations $Y = (Y_1, Y_2, \ldots\ldots Y_n)$ is given for an F_2 population by:

$$L(\theta|Y) = \prod_{i}^{n}\left[G_i(-1)z(y_i|-1;\theta) + G_i(0)z(y_i|0;\theta) + G_i(+1)z(y_i|+1;\theta)\right] \qquad (11)$$

where $G_i(X)$ represents the probability distributions of X over the values of -1, 0, and +1 (corresponding to QTL genotypes Q_2Q_2, Q_1Q_2 and Q_1Q_1) given the genotypes of the flanking markers for individual i. As before, this probability is a function of p, r and r'; however, given that p is known and that r' can be expressed in terms of p and r assuming the Haldane function ($p = r + r' - 2rr'$), this probability has only r as unknown. Therefore, the recombination fraction between the QTL and the left marker of the interval, r, is included in the likelihood through $G_i(X)$. Hence, for each value of r, a different likelihood will be obtained. The probabilities for all combinations of genotypes of the markers will be calculated for each individual i according to the specific genotype of its markers.

The function

$$z(y_i|x_i;\theta) = (2\pi\sigma^2)^{-1/2}\exp\left(-\frac{(y_i - \mu - ax_i - d(1-x_i^2))^2}{2\sigma^2}\right) \qquad (12)$$

denotes the conditional density of Y_i given $X_i = x_i$ assuming normal distribution of the trait.

In order to obtain the maximum likelihood estimates of θ, the function $L(\theta|Y)$ may be maximized through choices of θ. This may be accomplished by any maximizing method, for instance using the EM algorithm (Dempster et al., 1977) for a fixed r.

The E and M steps should be repeated for convergence of the estimates of θ. Once the parameters are obtained, the most likely position may be estimated by either, the maximum of the F statistic or the LOD score, through all possible values of r varying along the intervals. The LOD or the F values could be plotted against r; then, the most likely position of the QTL is given by the value of r, say r_M, which gives the maximum of the curve if that maximum is higher than a given significant threshold. The

maximum found for the LOD is not necessarily equal to that of the F. For an F_2 showing dominance, Carbonell et al. (1992) have given a detailed method to obtain the most likely position of the putative QTL and its genetic effects. A FORTRAN 77 computer program that implements the procedure has been written for F_2 populations (Carbonell and Gerig, 1991) and also for backcross and doubled haploids (unpublished).

Carbonell et al. (1993) studied the effect of two linked QTLs on the estimation of the recombination fraction and the genetic effects using their model with simulated data from doubled haploids. Results from 100 simulations using different heritabilities consistently gave bimodal distributions for the higher heritability of the trait (h^2 = 0.50). For the heritability of 0.08, few cases showed a highly significant unimodal distribution in a location intermediate to both QTL like the "ghost" QTL observed by Martinez and Curnow (1992), or a multimodal or plateaued curve indicating that QTLs could be found in a long range of values of the recombination fraction with no clear location of the maximum of the likelihood function. More frequently a bimodal curve like in the high heritability case was found indicating the presence of two QTLs; however, the predicted locations were biased in the sense that the leftmost QTL was shifted to the right and the leftmost one to the left as Martinez and Curnow (1992) have also found.

Therefore, implicit in the above methods is the fact that a single QTL is investigated at a time and that the QTL segregate independently from the remaining loci contributing to the quantitative trait. The full model is actually a summation of the a_j and d_j effects over the total number of QTLs (j = 1, 2,....q) so that when a particular QTL is investigated, it is assumed that its effect is not influenced by the rest of the QTL. Therefore, it is not intended to resolve clusters of QTLs but only those that are not too close to make tenable the assumption of independent action on the measured quantitative trait (Carbonell et al., 1992). Lander and Botsein (1989) suggest that when a LOD score graph shows evidence of two QTLs by the presence of two distinct peaks, each of the QTLs should, in turn, be fixed at their estimated positions. The difference

between the LOD scores calculated for the fixed QTL and a possible second QTL and the LOD scores for the fixed QTL only is then plotted against distance along the linkage group. A high peak indicates the presence of the second QTL. However, this procedure is not totally correct because the estimated positions and size of the effects of the two QTLs are not independent. This procedure could be a good approximation when the two QTLs are sufficiently well separated as in the case presented by Lander and Botsein (1989) where both QTLs were 80 cM apart but it does not work well when the two QTLs are much closer located (Knapp, 1991; Martinez and Curnow, 1992; Carbonell et al., 1993) or the presence of epistatic effects is suspected. Hence, when multiple linked QTLs affect the trait, these methods could lead to unpredictable results if linkage between QTLs is tight. This fact may have important implications if applied to a breeding program for marker-assisted selection

2.3 All markers simultaneously.

Multiple regression techniques (Draper and Smith, 1981) have been suggested by Cowen (1989) to identify multiple QTLs linked to several markers. This procedure seems to be the most reasonable analytic procedure when compared to the linear regression by "leaps and bounds" presented by Romero-Severson et al. (1989).

The basic procedure consists in assigning a coded value to each marker genotype and use stepwise regression procedures to identify which markers are associated with the variation of the quantitative trait. The problems with this approach are those of the stepwise selection of variables: several marked segments are usually available in the same chromosome for the analysis of the QTL effects; these segments likely do not segregate independently. Since stepwise regression procedures incorporate the variables into the model in a sequential manner, one by one, the estimates of the QTL effects which are already included in the model may be biased by other linked and unlinked QTL which may or may not be yet in the model. Some QTL already in the model may prevent the entrance of a new QTL in the model; therefore, some QTL may cover the presence of another QTL that will remain undetected by the statistical methodology. Moreover, the method may produce very different results depending on

the path of selection followed by the computer programs; i. e., if two variables have very similar F values to be in or out of the equation and are highly correlated, the final equation may depend on which of the two variables is chosen to enter in the equation.

Moreno-González (1992) used the ideas of Knapp et al. (1990) to include multiple segments and several putative QTL in a general model for different populations assuming the presence of digenic epistatic interaction and no double crossing-over. For the backcross population the model is:

$$Y_i = \mu_0 + \sum_j^s (a_j X_{1ij} + d_j X_{2ij} + \rho_j a_j X_{3ij} + \rho_j d_j X_{4ij}) + \text{ epistatic interactions} \qquad (13)$$

where a_j and d_j are the additive and dominance values (Falconer, 1989) of the putative QTL located in the jth segment, ρ_j is the probability that a recombinant gamete carries the Q_2 allele in the jth segment, and X's are dummy variables. The above non-linear model can be fitted by iterative least squares routines. The approach takes into consideration many terms in the equation which makes the computation to estimate all of them almost impractical and meaningless. One should make reasonable assumptions and decide which parameters should be retained in the model. The effect of double crossing-over when the markers are far apart also needs to be investigated.

Taking into account the information from three markers simultaneously will result in a search for one or more QTLs over a bivariate surface (two intervals at a time). Although this search could be done by the maximum likelihood approach, Martinez and Curnow (1992) used an extension of his regression method for pair of markers to the case of multiple markers and several linked QTLs. When the assumption that two QTLs are segregating, one in each of the neighbor regions, is true then their method is consistent and the positions of both QTLs are well estimated. Even when there is only one QTL, the method gives a surface that is very flat as the positions of the other hypothesized but non-existent QTL varies. In the situation of an empty flanked region between the two QTLs, the bivariate regression method obtains again estimates for two QTLs, the real one for which estimations are biased both in size and position, and a

"ghost" QTL with size of effect different from zero and a position between the two markers flanking the empty region. Clues to decide if the QTL is a real one or a ghost can be obtained using several intervals in turn. The decisive test is to use information provided by studying four markers simultaneously in a multivariate space search. Unfortunately, this search can be performed only if the sample size of the segregating population is large enough to have an appropriate representation of each one of the 16 different marker classes; then, the number of individuals to be screened becomes almost intractable.

Considering the genome as a whole, Rodolph and Lefort (1993) proposed a method based on the linear model that can be used in a very general genetic context. The purpose is to detect regions displaying QTL effects rather than to detect isolated QTLs and estimate their effects. Their linear model is actually an extension of the basic model in equation (1) summing over the set of m markers, The major advantage of the methods based on all markers simultaneously is the possibility of making global tests, recovering the whole or a part of a chromosome. In doing so, the existence of a set of linked QTLs can be detected. Conversely, the main characteristic of the interval mapping methods is their sequential nature. This is the reason why they are not well suited for the detection of closely linked QTL but they are more powerful in the case of isolated QTLs. Hence, a compromise should be reach using both approaches.

3. Power studies

The accuracy of any mapping procedure depends not only on the ability of the statistical method to detect the location of the QTLs involved in the expression of the trait, and to estimate their genetic effects. Other factors have a very important influence on this accuracy: the heritability of the trait, the contribution of each QTL to the total genotypic variance of the trait, the number of QTLs, their interactions, their distribution over the genome, their distance to the markers, the statistical distribution of the random non-genetic factors, the type of segregating population studied, its size, the genome size, the number of marker loci employed, etc. Given the large number of

factors, authors restrict themselves to partial studies investigating some ot these factors at a time.

For the single marker approach, results from the comparison between the maximum likelihood approach and the comparison of means seems to be influenced by the poor convergence during maximum likelihood estimation if the effects of the QTL are small and for recombination fractions near to 0.5. That was the case reported by Simpson (1992) when correcting his previous findings (Simpson, 1989). When a better computing algorithm was used the power of both tests was shown to be similar.

For the interval mapping approach Carbonell et al. (1993) used 100 replicated sets of 250 simulated individuals to investigate the power of the interval mapping to detect and estimate the genetic effects of several QTLs. The simulated individuals had eight linkage groups with six markers each separated 20 cM. Six QTLs were involved in the expression of the quantitative trait, with different gene action and the assumption of dominance in some QTLs or a completely additive model. From the results they concluded that F_2 and doubled haploid populations allow experiments to be conducted with smaller sample sizes because they show much higher power than backcrosses. Moreover, they give more accurate estimates of the location of the QTL and with less variance. This result could be expected from the fact that the model uses more spread-out values (+1 and -1) than backcrosses (0 and -1) and as Luo and Kearsey (1991) pointed out, from the fact that the absolute difference between the means of both homozygous genotypes is larger than that between the heterozygous an the recessive homozygous genotypes. If dominance is present backcrosses not only give biased estimation of the effects, because additive and dominant effects are completely confounded in this mating design, but also some QTL could not be detected if the genotypic values of the homozygote and the heterozygote at those loci are the same. That could be the case for complete dominance when the F_1 population is backcrossed with a "high" producing parental line and also when, independently of the direction of the backcross, the recurrent parental line had the "high" allele at some QTL. For doubled haploids, the power of detecting a given QTL is clearly related to its relative

contribution to the heritability of the character; i. e., proportional of to the square of its additive value a_j^2. The higher its contribution, the higher the probability of being detected. Even QTLs having small effects were identified by the model; the power of the test was about 90% for heritabilities of the QTL as low as 5%. To obtain similar power for backcrosses, the heritability attributable to the individual QTL should be around 14%. For a given type of gene action, the comparison with results given by Carbonell et al. (1992) for F_2's is difficult because two different criteria were used for this population but it seems that doubled haploids have similar power than F_2. In fact, similar values of 5% for the heritabilities of the QTL were obtained by van Ooijen (1992) for this type of population. However, if dominance is present, doubled haploids will only detect the additive component of a particular QTL. This fact could be of extreme importance for QTLs showing overdominant effects or when the quantitative trait needs to be studied to design the most efficient breeding strategy in order to exploit the non-additive variation hidden in some QTLs. The major technical advantage for doubled diploids, independently of any effect of replication on the required number of offspring, lies in the fact that the lines can be reproduced independently and continually evaluated with respect to additional quantitative traits and markers with all the information cumulative (Burr et al., 1988). If the effect of replication is taken into account, replicated progenies can bring about a major reduction in the number of individuals that need to be scored. Savings are the greatest when heritability of the trait is low, but are much reduced when heritability is moderate to high under the assumption of codominance at all QTL (Soller and Beckmann, 1990).

For a given nominal type I error, tests using F values are more powerful while keeping the realized type I error to values similar to nominal ones. It seems that the nominal significance level used applying the Bonferroni correction, was too low for LOD and a more conservative test should be used in this case (Carbonell et al., 1993).

Regarding the genetical approach it is interesting to compare the classical "single marker" methodology with the interval mapping. The methodologies based in all

markers simultaneously are much more specifically suitable for the case of multiple linked QTLs were the entire genome is investigated.

Lander and Botsein (1989) concluded that even in the worst case where the QTL is located in the middle of the interval, interval mapping decreases the number of offspring by a factor of (1 - p) as compared with the single marker approach; so for maps having markers at distances of 10, 20, 30 and 40 cM, the savings are about 9%, 16%, 23% and 28%, respectively (assuming the Haldane mapping function). However, Knott and Haley (1992) compared both maximum likelihood procedures for simulated F_2 data and found that the use of flanking markers only provides substantially more power for QTL detection than using single markers when the markers are widely spaced and the QTL effects large but interval mapping both correctly position the QTL and estimates its effect on the average. The estimates obtained using single markers are not close to the simulated values with both the distance from the nearest marker and the additive effect overestimated. Similar results were obtained by Rebai et al. (1994) when analytically compared the power of the maximum likelihood approach using interval mapping with the ANOVA approach for single markers for backcross populations. They showed that the ANOVA test is less powerful for all non null values of the distance between the flanking markers; however, for small intervals (distances from 0 to 30 cM) the difference in power between the two tests do not exceed 5%. For large intervals and quite large QTL effects the advantage of interval mapping is 10% to 30%. For simplicity in the analytical derivations, their approach assumed absence of double crossing over; therefore, the results for large intervals would be approximate and conclusion should be taken with caution. For small effects of the QTL (less than 1% of the phenotypic variance) powers of the tests are close to the significance level and could not be compared.

Single marker methodologies based on ANOVA and comparison of means are much simple to implement using standard statistical computer packages but they can be only used to the detection of several unlinked QTLs. They can be used as an exploratory phase of the study where we are simply interested in the detection of association. If an

better knowledge of the distributional properties of the test statistics being used, models for non-normally distributed traits, models for dominant markers like RAPDs, or for populations of allogamous plants, etc.

There has been a fast development of techniques in the area of genetic markers assisted by specific statistical methods that provide powerful means for QTL identification. However, in addition to QTL location and gene effects estimation, important aspects of QTL analysis are been studied such as the predictive value of QTLs within a pedigree and across evolutionary distance (Paterson et al., 1991); the effects of different experimental designs, heritabilities of the QT, and types of gene action on QTL detection (Carbonell et al., 1993); the efficiency of generations for estimating marker-associated QTLs (Moreno-González, 1993), etc. Therefore, there is a considerable body of knowledge that should enable us to deeply investigate other well known and important phenomena related to QTs such as: pleiotropy, gene interactions, heterosis, transgressive segregations, and genotype by environment interaction. Very recently, several authors have focused studies towards these subjects (de Vicente and Tanksley, 1993; Bretó et al., 1994; Asíns et al., 1994). Undoubtedly, designs that allow unlimited replications are most adequate to get a deeper insight on these phenomena using the methodology developed for QTL analysis.

References

Asíns, M. J., P. Mestre, J. E. García, F. Dicenta & E. A. Carbonell, 1994. Genotype x environment interaction in QTL analysis of an intervarietal almond cross by means of genetic markers. Theor. Appl. Genet. (in press)

Asíns, M. J. & E. A. Carbonell, 1988. Detection of linkage between restriction fragment length polymorphism markers and quantitative traits. Theor. Appl. Genet. 76: 623-626.

Bretó, M. P., M. J. Asíns & E. A. Carbonell, 1994. Salt tolerance in *Lycopersicon* species. III. Detection of quantitative trait loci by means of molecular markers. Theor. Appl. Genet. (In press).

estimate of the localization of the QTL and an unbiased estimation of effect is required, interval mapping should be used although previous mapping of the markers is needed. If precise estimation of the recombination fraction between the flanking markers is not available, Knott and Haley (1992) found that the use of an incorrect value for the recombination frequency between the flanking markers, does not affect the ability to map the QTL as long as the order of the markers along the chromosome is correct. On the other hand, if a QTL located in the distal part of the linkage group, the local maximum of the LOD score is likely to arise in the interval closer to the QTL or even at one of the marker giving a mislocalization of the QTL and an under-estimation of its effect (Martinez and Curnow, 1992). Therefore, as Rodolphe and Lefort (1993) conclude, the comparison between results obtained from different statistical methods can give an interesting insight especially on the number of QTLs; if the purpose is a precise dissection of the genetic determinism of the trait under study, this cannot be achieved with only one experiment; the search for QTLs should be an iterative process. The multimarker model enables the experimenter to focus on the interesting segments detected. New experiments must be made, with more markers on these segments and more individuals in order to get more recombinant genotypes between these closely linked markers. These new experiments can be analyzed using different models and methods adapted to different situations and purposes. Then, simultaneous use of several methods should lead to better efficiency in QTL technology.

4. Perspectives and further remarks

In the preceding sections it has been shown that there still are some unresolved matters that deserve further development. They are mainly related to the presence of closely linked and interacting QTLs, the development of efficient ways to detect QTL showing small effects, unveiling QTLs masked by other QTLs having very large effects, the presence of epistatic effects, studies not only directed towards methods to increase the power of the designs but also to decrease the probability of false positives,

Burr, B., F. A. Burr, K. H. Thompson, M. C. Albertson & C. W. Stuber, 1988. Gene mapping with recombinant inbreds in maize. Genetics 118: 519-526.

Carbonell, E. A., M. J. Asíns, M. Baselga, E. Balansard & T. M. Gerig, 1993. Power studies in the estimation of genetic parameters and the localization of quantitative trait loci for backcross and doubled haploid population. Theor. Appl. Genet. 86: 411-416.

Carbonell, E. A. & T. M. Gerig, 1991. A program to detect linkage between genetic markers and non-additive quantitative trait loci. J. Hered. 82: 435.

Carbonell, E. A., T. M. Gerig, E. Balansard & M. J. Asíns, 1992. Interval mapping in the analysis of nonadditive quantitative trait loci. Biometrics 48: 305-315.

Cowen, N. M., 1988. The use of replicated progenies in marker based mapping of QTL's. Theor. Appl. Genet. 75: 857-862.

Cowen, N. M., 1989. Multiple regression analysis of RFLP data sets in mapping QTLs. In: Helentjaris, T.& B. Burr (Eds.), Development and application of molecular markers to problems in plant genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor.

Dempster, A. P., N. M. Laird & D. B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Royal Stat. Soc., Series B 39: 1-38.

Draper, N. R. & H. Smith, 1981. Applied Regression Analysis. Wiley, New York.

Dudley, J. W., 1992. Theory for identification of marker locus-QTL associations in population by line crosses. Theor. Appl. Genet. 85: 101-104.

Edwards, M. D., C. W. Stuber & J. F. Wendel, 1987. Molecular marker-facilitated investigations of quantitative traits loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116: 113-125.

Falconer, D. S., 1989. Introduction to Quantitative Genetics. Longmans, London.

Gelderman, H., 1975. Investigation on inheritance of quantitative characters in animals by gene markers. 1. Methods. Theor. Appl. Genet. 46: 319-330.

Haley, C. S. & S. A. Knott, 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69: 315-324

Paterson, A. H., S. Damon, J. D. Hewitt, D. Zamir, H. D. Rabinowich, S. E. Lincoln, E. S. Lander & S. D. Tanksley, 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. Genetics 127: 181-197.

Paterson, A. H., E. S. Lander, J. D. Hewitt, S. Peterson, S. E. Lincoln & S. D. Tanksley, 1988. Resolution of quantitative traits into mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335: 721-726.

Rebai, A., B. Goffinet & B. Mangin, 1994. Comparing power of different methods for QTL detection. Biometrics (in press).

Rodolphe, F. & M. Lefort, 1993. A multi-marker model for detecting chromosomal segments displaying QTL activity. Genetics 134: 1277-1288.

Romero-Severson, J., J. Lotzer, C. Brown & M. Murray, 1989. Use of RFLPs for analysis of quantitative trait loci in maize. In: Helentjaris, T.& B. Burr (Eds.), Development and applications of molecular markers to problems in plant genetics, pp. 97-102 Cold Spring Harbor Laboratory, Cold Spring Harbor.

Sax, K., 1923. The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8: 552-560.

Simpson, S. P., 1989. Detection of linkage between quantitative trait loci and restriction fragment length polymorphisms using inbred lines. Theor. Appl. Genet. 77: 815-819.

Simpson, S. P., 1992. Correction: Detection of linkage between quantitative trait loci and restriction fragment length polymorphism using inbred lines. Theor. Appl. Genet. 85: 110-111.

Soller, M. & J. S. Beckman, 1983. Genetic polymorphism in varietal identification and genetic improvement. Theor. Appl. Genet. 67: 25-33.

Soller, M. & J. S. Beckmann, 1990. Marker-based mapping of quantitative trait loci using replicated progenies. Theor. Appl. Genet. 80: 205-208.

Jensen, J., 1989. Estimation of recombination parameters between a quantitative trait locus (QTL) and two marker gene loci. Theor. Appl. Genet. 78: 613-618.

Kahler, A. L.& C. F. Wehrhahn, 1986. Association between quantitative traits and enzyme loci in the F2 population of a maize hybrid. Theor. Appl. Genet. 121: 185-199.

Knapp, S. J., 1991. Using molecular markers to map multiple quantitative trait loci: models for backcross, recombinant inbred and doubled haploid progeny. Theor. Appl. Genet. 81: 333-338.

Knapp, S. J., W. C. Bridges & D. Birkes, 1990. Mapping quantitative trait loci using molecular marker linkage maps. Theor. Appl. Genet. 79: 583-592.

Knott, S. A. & C. S. Haley, 1992. Aspects of maximum likelihood methods for the mapping of quantitative trait loci in line crosses. Genet. Res. Camb. 60: 139-151.

Lander, E. S. & D. Botsein, 1986. Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. Proc. Natl. Acad. Sci. USA 83: 7353-7357.

Lander, E. S. & D. Botsein, 1989. Mapping mendelian factors underlying quantitative traits using RFLPs linkage maps. Genetics 121: 185-199.

Luo, Z. W. & M. J. Kearsey, 1991. Maximum likelihood estimation of linkage between a marker gene and a quantitative trait locus. II. Application to backcross and doubled haploid populations. Heredity 66: 117-124.

Martinez, O. & R. N. Curnow, 1992. Estimating the location and the sizes of the effects of quantitative trait loci using flanking markers. Theor. Appl. Genet. 85: 480-485.

Moreno-González, J., 1992. Estimates of marker-associated QTL effects in Monte Carlo backcross generations using multiple regression. Theor. Appl. Genet. 85: 423-434.

Moreno-González, J., 1993. Efficiency of generations for estimating marker-associated QTL effects by multiple regression. Genetics 135: 223-231

van Ooijen, J. W., 1992. Accuracy of mapping quantitative trait loci in autogamous species. Theor. Appl. Genet. 84: 803-811.

Stuber, C. W., M. D. Edwards & J. F. Wendel, 1987. Molecular marker-facilitated investigations of quantitative trait loci. II. Factors influencing yield and its component traits. Crop Science 27: 639-648.

Tanksley, S. D., H. Medina-Fhilo & C. M. Rick, 1982. Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49: 11-25.

Thoday, J. M., 1961. Location of polygenes. Nature 191: 368-370.

de Vicente, M. C.& S. D. Tanksley, 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134: 585-596.

Weller, J. I., 1986. Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. Biometrics 42: 627-640.

Weller, J. I., 1987. Mapping and analysis of quantitative trait loci in *Lycopersicon* (tomato) with the aid of genetic markers using approximate maximum likelihood methods. Heredity 59: 413-421.

Weller, J. I., Y. Kashi & M. Soller, 1990. Power of daughter and grandaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. J. Dairy Sci. 73: 2525-2537.

Weller, J. I., M. Soller & T. Brody, 1988. Linkage analysis of quantitative traits in an interspecific cross of tomato (*Lycopersicon esculentum* x *Lycopersicon pimpinellifolium*) by means of genetic markers. Genetics 118: 329-339.

Zhuchenko, A. A., A. B. Korol & V. K. Andyushchenko, 1979. Linkage between loci of quantitative characters and marker loci. Genetika 14: 771-778.

ANALISIS DE SENSIBILIDAD EN LA ESTIMACION MULTICARACTER DE COMPONENTES DE (CO)VARIANZA EN VACUNO LECHERO MEDIANTE MUESTREO DE GIBBS

L.Varona, C. Moreno, L.A. García Cortés, J. Altarriba

Unidad de Genética Cuantitativa y Mejora Animal

Facultad de Veterinaria. Universidad de Zaragoza.

C/M. Servet 177. 50013. ZARAGOZA

INTRODUCCION

El muestreo de Gibbs (Geman y Geman, 1984) es una técnica numérica de marginalización inscrita dentro de un conjunto de procedimientos basados en cadenas de Markov con distribuciones continuas de probabilidad (Metropolis et al.,1953; Hastings, 1970).

El muestreo de Gibbs destaca dentro de estos procedimientos por su simplicidad, debido a que las distribuciones implicadas en la construcción de la cadena de Markov son las condicionales de cada parámetro, o agrupación de parámetros, del modelo estadístico con respecto a todos los demás.

Estas distribuciones son fácilmente deducibles del conocimiento de la distribución conjunta de observaciones y parámetros. Esta distribución conjunta se construye mediante la multiplicación de la función de verosimilitud y las distribuciones a-priori de los parámetros. La definición correcta de las distribuciones a-priori para reflejar ignorancia es un proceso controvertido (Berger et al., 1992). El análisis con distintas distribuciones a-priori permite estudiar el peso de éstas en la marginal posterior obtenida (Gianola, 1994. comunicación personal). En este trabajo se intenta evaluar las repercusiones de la utilización de dos tipos de a-priori, por un lado se utiliza las recomendadas por Box y Tiao (1972), para el caso de un modelo multivariante fijo y por otro se utilizan distribuciones a-priori constantes.

De este modo, el muestreo de Gibbs se inicia con un grupo arbitrario de valores $\left(U_1^{(0)}, U_2^{(0)}, \ldots, U_i^{(0)}, \ldots, U_k^{(0)}\right)$; a partir de esta semilla se obtienen observaciones aleatorias de las siguientes distribuciones en un orden natural:

$$U_1^{(1)} \sim \left[U_1 | U_2^{(0)}, U_3^{(0)} \ldots . U_k^{(0)}\right] \quad U_2^{(1)} \sim \left[U_2 | U_1^{(1)}, U_3^{(0)} \ldots . U_k^{(0)}\right] \quad . \quad U_k^{(1)} \sim \left[U_k | U_1^{(1)}, U_3^{(1)} \ldots . U_{k-1}^{(1)}\right]$$

Después de un número suficientemente grande de iteraciones (i),se alcanza un valor muestreado de la distribución marginal de cada variable. Repitiendo este proceso m veces obtenemos m puntos que han sido muestreados de las distribuciones desconocidas. Promediando las distribuciones condicionales desde las cuales se han muestreado las observaciones, utilizando el argumento de Rao-Blackwell, se obtienen aproximaciones de las distribuciones marginales (Gelfand y Smith, 1990). Las estimaciones de momentos de interés se realizan a partir de estas distribuciones.

Frecuentemente, en lugar de utilizar diferentes series de muestreos de longitud i, se genera una única serie de gran longitud de la cual se utilizan elementos separados cada i elementos, ignorando un número inicial. Este procedimiento presenta la ventaja de necesitar un único periodo de aproximación (Raferty y Lewis, 1992), obteniéndose de este modo una considerable ventaja computacional.

Una importante característica es que, por su condición de proceso ergódico, la esperanza de las funciones de interés pueden calcularse mediante el promedio de todos los elementos de una única cadena, una vez se ha alcanzado la convergencia a un proceso estacionario (Chan, 1993). De este modo, si el parámetro de interés es exclusivamente la esperanza de un

parámetro o combinación de parámetros (p.j. heredabilidad), el procedimiento se muestra muy eficiente computacionalmente (Geweke, 1992).
La utilización del muestreo de Gibbs para aproximar los momentos de las distribuciones marginales presenta un serio limitante en cuanto a la convergencia del proceso a su distribución estacionaria. Existen distintos procedimientos para comprobar la convergencia (Geweke,1992; Smith y Roberts,1992; Gelman y Rubin, 1992, Raferty y Lewis, 1992).
En el campo de la mejora genética animal permite la utilización de inferencia bayesiana y la implementación práctica de modelos complejos (Sorensen et al. 1994a), por ejemplo, Sorensen et al. (1994b) aplica Muestreo de Gibbs en un modelo multiumbral analizado con el procedimiento de "Data Augmentation" (Tanner y Wong, 1987).

MATERIAL

Se dispone de los datos productivos de 26727 vacas, respecto a los caracteres cantidad de leche y porcentaje de grasa de 77001 lactaciones normalizadas a 305 días y de una genealogía de 31071 triadas individuo-padre-madre, con un 33% y un 47% de padre y madres faltantes respectivamente.
El modelo asumido para ambos caracteres es:

$$y_{ijkl} = EL_i + RAE_j + u_k + p_k + e_l$$

con 30 niveles de edad-lactación (EL) y 5471 niveles del efecto Rebaño-Año-Estacion (RAE); se han incluido 86 grupos de padre y madre desconocido.

METODOS

Se han construido las funciones de probabilidad conjunta de observaciones y parámetros utilizando dos tipos de distribuciones a-priori para los componentes de la varianza. En un primer caso se han utilizado a-priori no informativas para el modelo lineal multivariante fijo (Box y Tiao, 1972). La utilización de estas distribuciones a-priori puede conducir a distribuciones conjuntas impropias (Gianola, 1994. comunicación personal). En un segundo caso se han utilizado distribuciones a-priori planas para los componentes de varianza.
De este modo la distribución conjunta de parámetros y observaciones es:

$$p(\bar{y}_1, \bar{y}_2, \bar{\beta}_1, \bar{\beta}_2, \bar{u}_1, \bar{u}_2, \bar{p}_1, \bar{p}_2, \bar{g}_1, \bar{g}_2, G, H, R | Q, A, X, Z) \propto$$

$$p(\bar{y}_1, \bar{y}_2 | \bar{\beta}_1, \bar{\beta}_2, \bar{u}_1, \bar{u}_2, \bar{p}_1, \bar{p}_2, R, X, Z) \times p(\bar{u}_1, \bar{u}_2 | G, A, \bar{g}_1, \bar{g}_2, Q)$$

$$\times p(\bar{p}_1, \bar{p}_2 | H) \times p(G) \times p(H) \times p(R) \times p(\bar{\beta}_1, \bar{\beta}_2) \times p(\bar{g}_1, \bar{g}_2)$$

siendo:

$$p(\bar{y}_1, \bar{y}_2 | \bar{\beta}_1, \bar{\beta}_2, \bar{u}_1, \bar{u}_2, \bar{p}_1, \bar{p}_2, R, X, Z) \propto N\begin{pmatrix} X\bar{\beta}_1 + Z\bar{u}_1 \\ X\bar{\beta}_2 + Z\bar{u}_2 \end{pmatrix}, I \otimes R$$

$$p(\bar{u}_1, \bar{u}_2 | G, A, \bar{g}_1, \bar{g}_2, Q) \propto N\begin{pmatrix} Q\bar{g}_1 \\ Q\bar{g}_2 \end{pmatrix}, A \otimes G$$

$$p(\bar{p}_1, \bar{p}_2 | H) \propto N\begin{pmatrix} 0 \\ 0 \end{pmatrix}, I \otimes H$$

$$p(\bar{\beta}_1, \bar{\beta}_2) \propto k$$

$$p(\bar{g}_1, \bar{g}_2) \propto k$$

y

$$p(G) \propto |G|^{-\frac{1}{2}(n+1)}$$

$$p(H) \propto |H|^{-\frac{1}{2}(n+1)}$$

$$p(R) \propto |R|^{-\frac{1}{2}(n+1)}$$

en el caso a) y

$p(G) \propto k$

$p(H) \propto k$

$p(R) \propto k$

en el caso b).

donde:

y_i	Vector de observaciones para el carácter i.
β_i	Vector de efectos sistemáticos para el carácter i.
u_i	Vector de efectos genéticos para el carácter i.
p_i	Vector de efectos permanentes para el carácter i.
g_i	Vector de efectos grupos de progenitores desconocidos para el carácter i.
e_i	Vector de residuos para el carácter i.
G	Matriz de varianzas y covarianzas aditivas.
H	Matriz de varianzas y covarianzas permanentes.
R	Matriz de varianzas y covarianzas residuales.
A	Matriz de parentesco.
X_i	Matriz de incidencia de efectos sistemáticos del carácter i.
Z_i	Matriz de incidencia de efectos genéticos y permanentes del carácter i.
Q	Matriz de incidencia de efectos de progenitor desconocido.
n	Número de observaciones.
q	Número de animales.
t	Número de animales con datos.
m	Número de caracteres.

Fijando parámetros se obtienen las distribuciones condicionales que se necesitan en el muestreo. Se utiliza un muestreo independiente de cada parámetro de localización (efectos sistemáticos, aditivos, grupos de padre desconocido y ambientales permanentes). Estos parámetros poseen una distribución condicional normal, cuya media es la esperanza condicionada a los demás parámetros y su varianza coincidecon la inversa de la diagonal de la matriz de coeficientes de modelo mixto (Gelfand et al., 1990; Wang y Gianola, 1993).

Por otra parte, se procede a agrupar los parámetros de dispersión según su causalidad, obteniéndose las siguientes distribuciones condicionales conjuntas en el caso a):

$$p(R^{-1} | y_1, y_2, u_1, u_2, p_1, p_2, X_1, X_2, Z_1, Z_2) \propto W([SE]^{-1}, v = n - m + 1)$$

$$p(G^{-1} | u_1, u_2, A, Q, g_1, g_2) \propto W([SA]^{-1}, v = q - m + 1)$$

$$p(H^{-1} | p_1, p_2) \propto W([SP]^{-1}, v = t - m + 1)$$

donde:

$$[SE] = \begin{bmatrix} \tilde{e}_1' \tilde{e}_1 & \tilde{e}_1' \tilde{e}_2 \\ \tilde{e}_2' \tilde{e}_1 & \tilde{e}_2' \tilde{e}_2 \end{bmatrix}$$

$$[SG] = \begin{bmatrix} (\tilde{u}_1 - Q\tilde{g}_1)' A^{-1} (\tilde{u}_1 - Q\tilde{g}_1) & (\tilde{u}_1 - Q\tilde{g}_1)' A^{-1} (\tilde{u}_2 - Q\tilde{g}_2) \\ (\tilde{u}_2 - Q\tilde{g}_2)' A^{-1} (\tilde{u}_1 - Q\tilde{g}_1) & (\tilde{u}_2 - Q\tilde{g}_2)' A^{-1} (\tilde{u}_2 - Q\tilde{g}_2) \end{bmatrix}$$

$$[SP] = \begin{bmatrix} \tilde{p}_1' \tilde{p}_1 & \tilde{p}_1' \tilde{p}_2 \\ \tilde{p}_2' \tilde{p}_1 & \tilde{p}_2' \tilde{p}_2 \end{bmatrix}$$

y en el caso b)

$$p(R^{-1} | y_1, y_2, u_1, u_2, p_1, p_2, X_1, X_2, Z_1, Z_2) \propto W([SE]^{-1}, v = n - 2m - 1)$$

$$p(G^{-1} | u_1, u_2, A, Q, g_1, g_2) \propto W([SA]^{-1}, v = q - 2m - 1)$$

$$p(H^{-1} | p_1, p_2) \propto W([SP]^{-1}, v = t - 2m - 1)$$

De este modo el muestreo necesario se limita a la utilización de generadores de números aleatorios normales y de Wishart.

El cálculo de las distribuciones marginales aproximadas para cada uno de los componentes de varianza, se realiza promediando las distribuciones

desde las que han sido generados según el argumento de Rao-Blackwell (Gelfand y Smith, 1990).
La distribución independiente de cada uno de los componentes de varianza es, de este modo, una chi-cuadrado invertida con v grados de libertad (Box y Tiao, 1972), con función de probabilidad:

$$p(\sigma_{ii}|\,y,\tilde{\theta}) = \left[\Gamma\left(\frac{v}{2}\right)\right]^{-1}\left(\frac{SE_{ii}}{2}\right)^{v/2}\sigma_{ii}^{-\frac{1}{2}(v+2)}\exp\left(-\frac{SE_{ii}}{2\sigma_{ii}}\right)$$

Los coeficientes de correlación siguen la siguiente distribución (Box y Tiao, 1972):

$$p(\rho|\,y,\tilde{\theta}) \propto \left(1-\rho^2\right)^{\frac{1}{2}(v-2)}\int_0^{\infty}\omega^{-1}\left(\omega+\frac{1}{\omega}-2\rho r\right)^{-(v+1)}d\omega$$

donde:

$$\omega = \left(\frac{\sigma_{11}SE_{22}}{\sigma_{22}SE_{22}}\right)^{1/2} \qquad r = \frac{SE_{12}}{\left(SE_{11}SE_{22}\right)^{1/2}} \qquad \rho = \frac{\sigma_{12}}{\left(\sigma_{11}\sigma_{22}\right)^{1/2}}$$

Cuando el número de grados de libertad es suficientemente grande puede aproximarse por (Box y Tiao, 1972):

$$p(\rho|\,y,\tilde{\theta}) \propto \frac{\left(1-\rho^2\right)^{\frac{1}{2}(v-2)}}{\left(1-\rho r\right)^{v+\frac{1}{2}}}$$

con constante de normalización:

$$\left(\int_{-1}^{1}\frac{\left(1-\rho^2\right)^{\frac{1}{2}(v-2)}}{\left(1-\rho r\right)^{v+\frac{1}{2}}}d\rho\right)^{-1}$$

fácilmente calculable mediante técnicas de integración numérica.
Por otra parte, las distribuciones de la heredabilidad y de la proporción de varianza ambiental permanente se calculan mediante una transformación de las distribuciones condicionales de la varianza aditiva y de la varianza ambiental permanente. Para ello se han fijado las demás varianzas ($\dot{\sigma}_a, \dot{\sigma}_p, \dot{\sigma}_e$) a las muestras de Gibbs actualizadas (Wang et al., 1993), obteniéndose las siguientes distribuciones de probabilidad para la heredabilidad (h^2):

$$p(h^2|\,y,\tilde{\theta}) \propto \frac{\Gamma(v/2)^{-1}}{2^{v/2}}\left(h^2\right)^{-v/2-1}\left(1-h^2\right)^{v/2-1}\left(\frac{su}{\dot{\sigma}_p^2+\dot{\sigma}_e^2}\right)^{v/2}\exp\left\{-\frac{1}{2}\left(\frac{su}{\dot{\sigma}_p^2+\dot{\sigma}_e^2}\right)\frac{h^2}{\left(1-h^2\right)}\right\}$$

donde $su=SG_{ii}/v$. y v=q-m+1 en el caso a) y v=q-2m-1 en el caso b)
Para la proporción de varianza ambiental permanente (c^2):

$$p(c^2|\,y,\tilde{\theta}) \propto \frac{\Gamma(v/2)^{-1}}{2^{v/2}}\left(c^2\right)^{-v/2-1}\left(1-c^2\right)^{v/2-1}\left(\frac{sp}{\dot{\sigma}_a^2+\dot{\sigma}_e^2}\right)^{v/2}\exp\left\{-\frac{1}{2}\left(\frac{sp}{\dot{\sigma}_a^2+\dot{\sigma}_e^2}\right)\frac{c^2}{\left(1-c^2\right)}\right\}$$

para el porcentaje de varianzas permanente donde $sp=SP_{ii}/v$ y v=t-m+1 en el caso a) y v=q-2m-1 en el caso b)
El muestreo de Gibbs exige la obtención de valores de distribuciones condicionales secuencialmente. Para la construcción de las distribuciones de cada nivel de los efectos fijos y aleatorios se necesitan todos los elementos de una determinada fila de la matriz de coeficientes de modelo mixto. El coste de almacenamiento de esta matriz de coeficientes es elevado en casos de tamaño grande. Para ello se ha optado por una estrategia de listas ligadas que permite almacenar y manejar en CPU matrices vacías de gran dimensión.
Los muestreos de las distribuciones se realizan mediante el algoritmo de Box y Muller (Press et al, 1986), en el caso de las distribuciones normales, y mediante en el de Odell y Feivesson (Jonhson, 1987), en el caso de las distribuciones de Wishart.
Construida la estimación de distribuciones marginales, los momentos se estiman por la técnica de integración numérica de Gauss-Legendre (Press et al., 1986)

RESULTADOS

El número de muestras Gibbs realizadas fue de 14042 después de desechar 330 observaciones en el caso a) y de 8200 observaciones en el caso b). Estos valores fueron obtenidos mediante el método de Raferty y Lewis (1992) y tal como se muestra en Varona et al. (1994) permiten calcular percentiles centrales de la distribución marginal posterior con una precisión de 0.001 con un 95% de nivel de confianza. Los resultados de convergencia no diferían substancialmente al modificar las distribuciones a-priori de los componentes de la varianza.
Los resultados para heredabilidades, proporciones de varianza ambiental permanente y correlaciones se muestran en la TABLA I; en la TABLA II se presentan los intervalos de confianza al 95% obtenido mediante ambos procedimientos.

TABLA I.- HEREDABILIDADES, PROPORCIONES DE VARIANZA AMBIENTAL PERMANENTE Y CORRELACIONES ESTIMADAS.

	MEDIA-A	DESV. TIPICA-A	MEDIA-B	DESV. TIPICA-B
h^2 Leche	0.212	0.011	0.208	0.011
c^2 Leche	0.268	0.010	0.271	0.009
h^2 % Grasa	0.411	0.015	0.406	0.016
c^2 % Grasa	0.105	0.013	0.108	0.014
Corr. Aditiva	-0.282	0.029	-0.283	0.027
Corr. Permanente	-0.167	0.045	-0.166	0.035
Corr. Residual	-0.070	0.005	-0.070	0.005

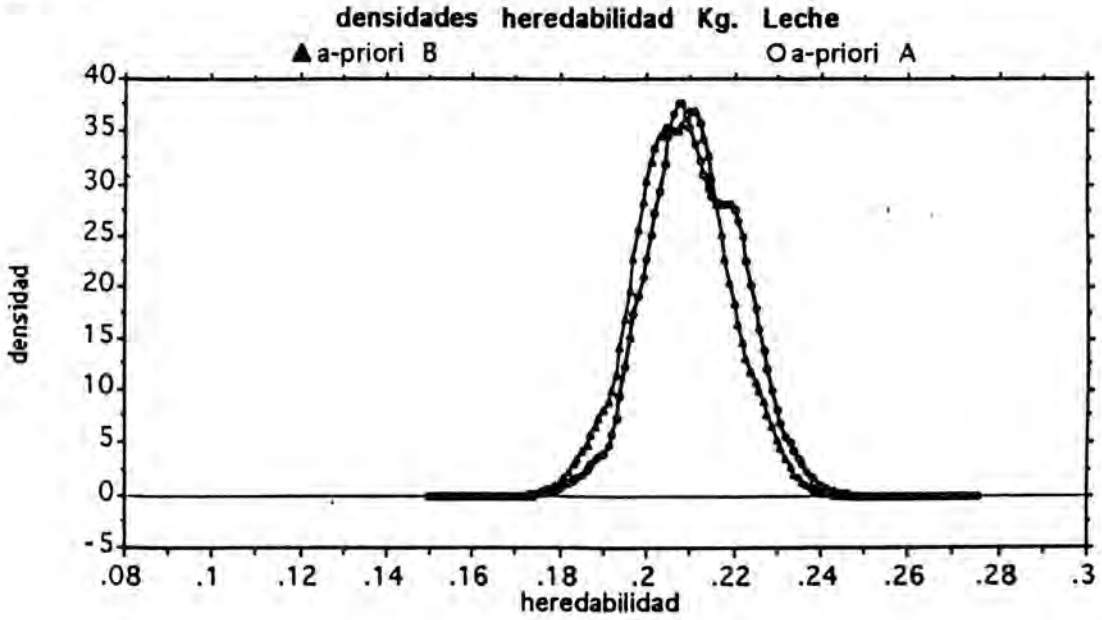
TABLA II.- INTERVALOS DE CONFIANZA PARA HEREDABILIDADES, PORCENTAJES DE VARIANZA PERMANENTE Y CORRELACIONES

	H.P.D. 95% - A	H.P.D. 95% - B
h^2 Leche	0.190/0.232	0.187/0.229
c^2 Leche	0.249/0.288	0.252/0.290
h^2 % Grasa	0.386/0.447	0.378/0.443
c^2 % Grasa	0.073/0.126	0.078/0.134
Corr. Aditiva	-0.230/-0.342	-0.227/-0.334
Corr. Permanente	-0.066/-0.242	-0.098/-0.237
Corr. Residual	-0.061/-0.079	-0.061/-0.079

Estos resultados están de acuerdo con los amplios márgenes descritos en la bibliografía. Destaca, en el carácter producción de leche, la magnitud de la varianza ambiental permanente respecto a la varianza genética. Las correlaciones calculadas son todas negativas, siendo la de mayor magnitud la correlación genética y la menor la residual.
Los resultados obtenidos con ambas distribuciones a-priori no difieren substancialmente, aunque se observa un ligero incremento de los porcentajes de varianza permanentes y una ligera disminución de las heredabilidades en el caso b). Estos resultados confirman la pretensión de ignorancia previa de las dos distribuciones utilizadas. Podemos concluir, por lo tanto, que

el peso de las distribución a-priori no informativas en problemas de tamaño grande queda minimizado por el gran peso de la verosimilitud en la distribución posterior. Hemos experimentado, por el contrario, una gran repercusión en el cambio de las distribuciones a-priori en problemas donde los datos proporcionen información mucho mas deficiente. En casos con muy poca información la utilización del a-priori del caso a) provoca estados absorbentes en varianzas nulas o en correlaciones uno.
Las distribuciones marginales calculadas quedaron muy poco definidas debido al pequeño número de puntos utilizados en su definición. En la gráfica I se muestran las distribuciones marginales calculadas para la heredabilidad del carácter cantidad de leche por los dos procedimientos.

Grafica I.



BIBLIOGRAFIA

BERGER J.O., BERNARDO, J.M. 1992. In: *Bayesian Statistics 4* (eds J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith),pp. 765. Oxford: Oxford University Press.

BOX G.E.P., TIAO G.C. 1972. Bayesian Inference in Statistical Analysis. Adison-Wesley, Reading.

CHAN K.S. 1993. J.A.S.A. 88:320.

GELFAND A.E., HILLS S.E., RACINE-POON A., SMITH A.F.M. 1990. J.A.S.A. 85:972.

GELFAND A.E., SMITH A.F.M. 1990. J.A.S.A. 85:398.

GEMAN S., GEMAN D. 1984. IEEE Transaction on pattern analysis and machine intelligence 6, 721.

GEWEKE J. 1992. In *Bayesian Statistics 4* (eds J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith), pp. 169. Oxford: Oxford University Press.

HASTINGS W.K. 1970. Biometrika 57:97.

JONHSON, M.E. 1987. Multivariate Statistical Simulation. John Wiley and sons. 230 pp.

METROPOLIS N., ROSENBLUTH A.W., ROSENBLUTH M.N., TELLER A.H., TELLER E. 1953. J. Chem. Phys.,21:1087.

PRESS W.H., FLANNERY B.P., TEUKOLSKY S.A., VETTERLING W.T. 1986. Numerical Recipes. The art of Scientific Computing. Cambridge University Press.

RAFERTY A. 1993. In: Discussion on Gibbs Sampler and Other MonteCarlo Methods. (pag. 83). J. R. Statist. Soc. B 55:53.

RAFERTY A., LEWIS S. 1992. In *Bayesian Statistics 4* (eds J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith),pp. 765. Oxford: Oxford University Press.

SMITH A.F.M., ROBERTS G.O. 1993. J.R. Statist. Soc. B 55:3-23.

SORENSEN D., ANDERSEN S., GIANOLA D., KORSGAARD I. 1994. Gen. Sel. Evol. (Enviado)

SORENSEN D., ANDERSEN S., JENSEN J.,WANG C.S., GIANOLA, D. 1994. In: Proc 5Th World Cong. on Genet. Appl. to Livest. Prod. (Enviado).

TANNER M.A., WONG, W.H. 1987. J.A.S.A. 82:528.

VARONA L., MORENO C., GARCIA CORTES L.A., ALTARRIBA J. 1994.. Revista Portuguesa de Zootecnia. 1:185-195.

WANG C.S., GIANOLA D. 1993. In: Proc. 44Th Annual Meeting European Association for Animal Production. Aarhus, Denmark. August, 1993.

WANG C.S., RUTLEDGE J.J., GIANOLA, D. 1993. Genet. Sel. Evol. 25:41.

Modelos matemáticos para la estimación de parámetros genéticos de la producción láctea y el Recuento de Células Somáticas (SCC) en ganado ovino.

Baro, J. A. J. A. Carriedo y F. San Primitivo.
Dept. de Producción Animal, Universidad de León.

Estudiamos los recuentos de células somáticas de las ovejas incluidas en el Control Lechero Oficial de ANCHE. Los análisis cualitativos de la mayor parte de la Comunidad Autónoma de Castilla y León se realizan en el CENSYRA situado en las inmediaciones de la Facultad de Veterinaria por lo que la logísitca de recogida de muestras se simplificaba mucho. En seis meses (Enero a Julio,1992) se analizaron 32000 muestras provenientes de unas 10000 ovejas de 52 rebaños en León, Palencia, Valladolid y Segovia. Siguiendo criterios tales como la edad media de las muestras a su llegada al laboratorio (menor de siete días), la modalidad de ordeño (se excluyeron las explotaciones de ordeño manual), la inclusión de una sola lactación por individuo y otros, se llegó a una muestra de 10171 controles de 3832 ovejas hijas de 217 sementales -de los cuales 33 pertenecían al hato de inseminación artificial- y 2011 hembras (de ellas 400 aportaban asimismo muestras).

Experiencias anteriores nos habían mostrado que se producen grandes variaciones en el recuento individual en días sucesivos, de acuerdo con lo observado por Reneau y por

Schultz en vacuno. Se decidió entonces que sería preferible un modelo para observaciones de "día de control", con varios registros por lactación, de forma que no se combinen estos valores en una medida única para toda la lactación y solventar asi la predecible baja repetibilidad de la variable SCC. Una ventaja adicional de los modelos lineales para obsevaciones del día de control es que permite incorporar efectos que consideren las circunstancias que afectaron a todos los individuos el día del test y reducir así la variabilidad no explicada entre controles consecutivos. Esto se hizo mediante la inclusión de un efecto llamado "rebaño-control". Meyer (1989) encontró que tal medida resulta en menores varianzas residuales que cuando se incluye el clásico efecto "estación-rebaño-año" y que ello se reflejaba asimismo en estimas de la heredabilidad ligeramente superiores, si bien a costa de manejar gran número de pequeños grupos para este efecto, que por otra parte, no se dan en el ganado ovino debido al mucho mayor tamaño medio de los rebaños. Es preciso asimismo incluir en el modelo efectos que den cuenta de la evolución de las variables, de su curva de lactación, tales como una regresión por días de lactación.

La contemplación de los estadísticos descriptivos de la muestra (Tabla 1) pone de manifiesto otra peculiaridad de la variable SCC que debe ser solventada antes de pretender estimar parámetros mediante métodos REML. El recuemto de células somáticas presenta una distribución fuertemente

sesgada hacia los valores más elevados, de forma que unos pocos recuentos altísimos elevan enormemente la media. Ali y Shook (1980) han mostrado que la transformación logarítmica del SCC hace que presente una distribución casi normal y mayor heredabilidad

Tabla 1. Estadísticos descriptivos.

	media	mediana	desv. típ.	coef. var.	sesgo
SCCo	2254	360	5370	238%	4.68
SCCv	692	392	1121	162%	
L_eSCCo	6.14	5.88	1.75	28%	.43
LDC	845	760	493	58%	1.78

SCCo = SCC en ovino
SCCv = SCC en vacuno
L_eSCC = logaritmo natural del SCC
LDC = producción láctea del día de control.

La repetición de observaciones en un mismo individuo se tuvo en cuenta en el modelo mediante la inclusión de un efecto aleatorio adicional para el que se asumió una estructura de covarianza proporcional a la matriz identidad ya que Ali y Schaeffer (1987) demostraron que no es mejor el ajuste que se obtiene cuando los elementos de la diagonal principal difieren según el estadio de la lactación de los controles, ni cuando se consideran correlaciones entre ellos.

Referencias:

Ali, T. E. y R. Schaeffer (1987). Accounting for covariances among test day milk yields in dairy cows. Can. J. Anim. Sci. 67:637-644.

Ali, T. E. y G. E. Shook (1980). An optimum transformation for somatic cell concentration in milk. J. Dairy Sci.63:487-490.

Ptak, E. y L. R. Schaeffer (1993). Use of test day yields for genetic evaluation of dairy sires and cows. Livestock production science, 34:23-34.

Meyer, K., H. U. Graser y K. Hammond (1989). Estimates of genetic parameters for first lactation test day production of Australian black and white cows. Livestock production science, 21:177-199.

Reneau, J. (1986). Effective use of dairy herd improvement somatic cell counts in mastitis control. J. Dairy Sci. 69:1708-1720.

Schultz, L. H. (1977). Somatic cells in milk - Physiological aspects and relationships to amount and composition of milk. Journal of food protection, 40:125-131.

MEJORA DEL MODELO DE EVALUACION GENETICA EN EL ESQUEMA DE SELECCION DE LA RAZA OVINA MANCHEGA.

Serrano, M.* , Pérez-Guzmán, M.D. ** , Montoro, V. ** y Jurado, J.J * .

* Area de Mejora Animal del CIT-INIA de Madrid.

** CERSYRA de Valdepeñas. Ciudad Real.

Introduccion:

En el estado actual de desarrollo del Esquema de Selección de la raza ovina Manchega, nos ha parecido necesario realizar un estudio para mejorar el modelo a aplicar en la valoración genética de los reproductores, principalmente en cuanto a los efectos fijos, y sus combinaciones. Uno de los objetivos principales de este trabajo ha sido la definición del efecto rebaño-año-estación móvil (Wiggans y col.,1988), con el fin de incluirlo en el modelo de evaluación, y evitar asi los sesgos que se producen cuando se tienen grupos de comparación con pocas observaciones (Ugarte y col., 1992).

Material y Métodos:

En este trabajo se han utilizado las lactaciones tipificadas a 120 dias de lactación del Esquema de selección de la raza ovina Manchega de la campaña de 1994. Los archivos constan de 26.668 lactaciones y 23.043 animales en genealogía para los caracteres leche normalizada (leche producida en 120 dias de lactación al 6% de grasa) (LN), leche total (leche producida en 120 dias de lactación) (LT), porcentaje de grasa (PG) y porcentaje de proteina (PP). En estos datos, los efectos fijos desconocidos en cualquiera de las observaciones, fueron incluídos en niveles con un código especial con el fin de no perder información.

La calidad del modelo estudiado se determinó mediante el cálculo del coeficiente de determinación, con el procedimiento GLM del paquete estadístico SAS (Sas, 1990) para los modelos de efectos fijos, y con un programa propio, BLUPAM-VERROR (Jurado y Hernandez, 1990) para los modelos mixtos. Este último calcula las predicciones de los efectos permanentes a partir de la predicción BLUP, y las utiliza junto con las estimas de los efectos fijos y predicciones genéticas para calcular el coeficiente de determinación y la varianza residual del modelo. Las heredabilidades y repetibilidades empleadas fueron 0.25 y 0.40 para LN y LT y 0.50 y 0.61 para PG y PP, respectivamente.

Para la construcción del efecto rebaño-año de parto-estación móvil se utilizó el programa AMV101 (Rozzi, 1992). Este programa construye a partir de los datos grupos de comparación con un mínimo de seis observaciones, con rebaño y año de parto común y cuya diferencia en mes de parto es como máximo de dos meses.

Los coeficientes de regresión de los caracteres leche normalizada y total y porcentajes de grasa y proteina sobre el intervalo parto-1er control, se calcularon mediante el paquete estadístico BMDP (Bmdp, 1990).

Resultados y discusión:

El análisis GLM1 (Tabla 1) corresponde al modelo de efectos fijos utilizado en la valoración genética realizada en el año 1993 (Serrano y col., 1993). Este modelo tiene el efecto estación de parto distribuído en cuatro niveles, y no incluye los factores edad al parto e intervalo parto-1er control. A partir del modelo GLM2, se incluyeron en los análisis estos dos últimos efectos.

Los coeficientes de regresión de los caracteres LN, LT, PG y PP sobre el número de días transcurridos entre el parto y el 1er control (INT) fueron: -0.2273 ± 0.0326, -0.5717 ± 0.0285, 0.0246 ± 0.0008 y 0.0071 ± 0.0003, respectivamente. Estos valores indican que cuanto más tarde se realiza el primer control menor es la producción controlada de LN y LT y algo superior la de PG y PP.

El factor INT fue incorporado al modelo de análisis como covariable en GLM2, GLM3, GLM4 y GLM5 y como efecto fijo con tres niveles (intervalo corto, medio y largo) en GLM6.

El año de nacimiento fue eliminado a partir del modelo GLM4, ya que se pretende incluirlo como grupo genético en futuras evaluaciones.

El efecto rebaño-año-estación del modelo GLM4 se construyó con el mes de parto distribuido en cuatro niveles, con el fin de comparar los resultados con los modelos GLM5 y GLM6 donde los efectos rebaño-año-estación móvil se constituyeron con el mes de parto compuesto por doce niveles.

Tabla 1: Coeficientes de determinación obtenidos mediante GLM (SAS), absorbiendo el efecto RAE en GLM4 y el efecto RAEM en GLM5 y GLM6.
MODELOS:

GLM1:	**y = reb + ap + est + np + tp + an + e.**	**(26.668 lactaciones)**
GLM2:	**y = reb + ap + est + np + ed + tp + b(int) + an + e.**	**(26.668 lactaciones)**
GLM3:	**y = reb + ap + est + nped + tp + b(int) + an + e.**	**(26.668 lactaciones)**
GLM4:	**y = rae + nped + tp + b(int) + e.**	**(26.668 lactaciones)**
GLM5:	**y = raem + nped + tp + b(int) + e.**	**(26.522 lactaciones)**
GLM6:	**y = raem + nped + tp + int + e.**	**(26.522 lactaciones)**

Caracter	GLM1	GLM2	GLM3	GLM4	GLM5	GLM6
LN	0.326947	0.335316	0.336322	0.413369	0.420446	0.420638
LT	0.428638	0.437973	0.438593	0.497315	0.502032	0.502207
PG	0.334120	0.356947	0.357372	0.524073	0.535504	0.534719
PP	0.251777	0.291285	0.291760	0.367079	0.376834	0.375196

reb = rebaño (54 niveles).
ap = año de parto (15 niveles).
est = mes de parto (4 niveles en GLM1 y 12 niveles en el resto).
np = número de parto (4 niveles).
ed = edad al parto (105 niveles).
tp = tipo de parto (4 niveles).
int = intervalo parto-1er control (covariable en GLM2, GLM3, GLM4 y GLM5 y 3 niveles en GLM6).
an = año de nacimiento (23 niveles).
nped = número de parto-edad al parto (102 niveles).
rae = rebaño-año-estación (503 niveles).
raem = rebaño-año-estación móvil (548 niveles).
e = residuo.
b = Coeficiente de regresión del carácter sobre la covariable.

En la tabla 1 se observa un progresivo incremento del CD desde el modelo GLM1 hasta los modelos GLM5 y GLM6. El mayor aumento se produce al incorporar el efecto RAE, por tener en cuenta las interacciones entre los factores que lo componen. Para PG y PP el coeficiente de determinación superior se obtuvo con GLM5; y para LN y LT con GLM6. En GLM1,GLM4 y GLM6 todos los efectos fijos estudiados fueron altamente significativos para todos los caracteres. En GLM2 el INT no fue significativopara el carácter LN, y el NP y TP no fueron significativos para el PP. En GLM3 el INT no fue significativo para LN y el TP no fue significativo para PP. Y por último en GLM5 el TP no fue significativo para el carácter PP.

En la tabla 2 se presentan los coeficientes de determinación y varianzas residuales, obtenidas con el programa BLUPAM-VERROR con las predicciones genéticas calculadas por el método de Gauss-Seidel tras 5000 iteraciones, para los modelos mixtos que incluyen los efectos fijos de GLM1, (MOD1), GLM4, (MOD4), GLM5, (MOD5) y

GLM6, (MOD6), y los efectos aleatorios genotipo del animal (G) y efecto permanente (Ep). Para todos los caracteres se observa un aumento del coeficiente de determinación y una disminución de la varianza residual del MOD5 y del MOD6 con respecto al resto de los modelos.

Tabla 2: Coeficientes de determinación (CD) y Varianzas residuales (VE) obtenidos con el programa BLUPAM-VERROR para los modelos:

MOD(1) : y = reb + ap + est + np + tp + an + G + Ep + e. **(26.668 lactaciones y 23.043 animales)**
MOD(4) : y = rae + nped + tp + b(int) + G + Ep + e. **(26.668 lactaciones y 23.043 animales)**
MOD(5) : y = raem + nped + tp + b(int) + G + Ep +e. **(26.522 lactaciones y 22.937 animales)**
MOD(6) : y = raem + nped + tp + int + G + Ep + e. **(26.522 lactaciones y 22.937 animales)**

CARACTER	CD				VE			
	MOD1	MOD4	MOD5	MOD6	MOD1	MOD4	MOD5	MOD6
LN	0.5842	0.6458	0.6505	0.6506	1468.6793	1275.5947	1262.7940	1262.5088
LT	0.6539	0.6999	0.7034	0.7034	943.5389	834.0571	826.3087	826.2310
PG	0.6570	0.7659	0.7745	0.7739	0.7918	0.5509	0.5312	0.5328
PP	0.6528	0.7170	0.7237	0.7223	0.1227	0.1020	0.0997	0.1002

Efectos fijos especificados en Tabla 1.

El MOD5 presenta un coeficiente de determinación algo superior para los caracteres PG y PP, que el resto de los modelos. La diferencia con el MOD6 no es significativa, y teniendo en cuenta que la convergencia mediante el proceso iterativo Gauss-Seidel con el MOD5 es mucho más lenta, debido a la covariable, parece más interesante la utilización del MOD6.

Parece por tanto que la inclusión de los efectos rebaño-año-estación móvil y número de parto-edad al parto mejoran de forma notable tanto el modelo de efectos fijos como el modelo mixto, y sobre todo con respecto al modelo fijo GLM2 donde todos estos efectos fijos estaban contemplados pero de forma independiente, manifestándose así la importancia de considerar las interacciones entre dichos factores. Dados los resultados obtenidos, será el modelo 6 el utilizado en las futuras evaluaciones de los animales de esta raza.

Bibliografía:

BMDP STATISTICAL SOFTWARE MANUAL, 1990: W.J. Dixon. University of California Press. Berkeley. California.

JURADO, J.J. y HERNANDEZ, D., 1990: Catálogo de Software de Interés en Agricultura. M.A.P.A.. Programa nº 113, pag: 81.

ROZZI, P., 1992: Comunicación personal.

SAS LANGUAGE, 1990: Reference V.G. (1042 pp.) Cary, N.C. SAS Institute Inc.

SERRANO, M.; JURADO, J.J.; PEREZ-GUZMAN, M.D. y MONTORO, V., 1993: Estudio de la tendencia genética y fenotípica de diversos caracteres en la raza ovina Manchega. ITEA. 12, Tomo1: 233-235.

UGARTE, E.; ALENDA, R. y CARABAÑO, M.J., 1992: Fixed or random contemporary groups in genetic evaluations. J. Dairy. Sci. 75: 269.

WIGGANS, G.R.; MISZTAL, I. y VAN VLECK, L.D., 1988: Implementation of an Animal Model for genetic evaluation of dairy cattle in the United States. Proceeding of the Animal Model Workshop. J. Dairy. Sci., Vol 71. Suppl. 2: 54-69.

Variación cuantitativa y cualitativa de las caseínas de la leche de cabra

Díaz Carrillo, Esteban; Angulo Heras, Conrado; Jiménez Gamero, Inmaculada; Baena Manzano, Francisca; Alonso Moraga, Angeles; Muñoz Serrano, Andrés; Serradilla Manrique, Juan Manuel.

Fisiogenética, ETSIAM, Apdo. 3048, 14080 Córdoba.

El equipo que suscribe viene realizando desde 1990 estudios para el desarrollo de esquemas de selección en ganado caprino lechero y más recientemente intenta elaborar nuevos criterios de selección. Para ello se está estudiando la variación del contenido en caseínas de la leche de tres razas de cabras españolas (Murciano-Granadina, Malagueña y Payoya), y su relación con otros parámetros productivos, con el fin de sentar las bases para el establecimiento de nuevos criterios de selección.

Se ha estudiado el efecto de la raza y de los factores no-genéticos considerados fijos sobre: a) la evolución a lo largo de la lactación de los contenidos en caseínas, de la producción y de los tres componentes principales de la leche después del agua (grasa, lactosa y proteína); b) la producción y contenidos medios por lactación de las caseínas y de estos tres componentes de la leche. También se ha estudiado las correlaciones fenotípicas entre las caseínas, la producción y los tres principales componentes de la leche. Por último se han estudiado las frecuencias de las variantes polimórficas de las caseínas caprinas en las razas mencionadas y la relación entre las variantes polimórficas y los parámetros productivos.

Las caseínas de la leche de cabra suponen del 75 al 80% de las proteínas totales, existiendo cuatro tipos de caseína distintos, cada una de ella codificada por un gen: α_{s1} (con 7 variantes polimórficas), α_{s2} (con dos alelos), ß y k (con dos alelos, descritos en las razas españolas por nosotros por primera vez).

Resultados

- Existen grandes diferencias en cuanto a las curvas de lactación para la producción total y cada uno de los componentes de la leche. En general, la producción de leche y el contenido en lactosa disminuyen a lo largo de la lactación, mientras que los contenidos en grasa proteína, caseína total y las distintas fracciones caseínicas tienden a aumentar al final de la lactación.

- Nuestros resultados nos permiten afirmar que los contenidos medios de los componentes lácteos (grasa, lactosa, proteína caseína y fracciones caseínicas) se ven influenciados por los factores raza, rebaño, año de parto, estación de parto, número de lactación y tipo de parto, siendo menor la influencia de estos dos últimos factores sobre los contenidos de caseínas, y particularmente sobre los contenidos de caseínas α_s ($\alpha_{s1} + \alpha_{s2}$) y k. Los contenidos medios de estos componentes de la leche no se ven significativamente afectados por el número de ordeños (Tabla 1).

- Con excepción del tipo de parto, que es menos significativo, los factores que afectan a las producciones totales por lactación son los mismos que los del apartado anterior. La incidencia de muchos de los factores considerados y sus interacciones sobre las producciones de leche y de sus componentes se modifican cuando se tienen en cuenta las diferencias de duración de las lactaciones, lo que indica que muchas de las diferencias detectadas son consecuencia de diferentes duraciones de la lactación y no de diferencias de producción media.

En el único caso de ganadería con doble ordeño, perteneciente a la raza Murciano-Granadina, el número de ordeños fue significativo cuando se consideran lactaciones tipificadas a 210 días, teniendo las cabras de un ordeño las lactaciones más largas.

- Las correlaciones fenotípicas obtenidas para las producciones totales fueron positivas con todos los componentes. Las correlaciones fenotípicas obtenidas para los porcentajes entre los contenidos medios indican que la proteína, grasa, y caseínas se encuentran correlacionadas positivamente y todos se encuentran correlacionadas negativamente con la producción de leche y con el contenido en lactosa. Los contenidos de las caseínas α_s y k se correlacionan negativamente con el contenido en caseína ß.

- Existe una aparente relación entre las variantes genéticas de la caseínas α_{s1}, α_{s2}, y k y las producciones o contenidos de los componentes lácteos (tabla 2). Las cabras que presentan el alelo B de la caseína α_{s1} se correlacionan con los mayores contenidos en caseína total. Las cabras que poseen el genotipo BB de la caseína α_{s2} son las que dan una mayor producción diaria de leche y menores contenidos en proteína, caseínas (total y fracciones) grasa y lactosa. Las cabras con el genotipo BB de la caseína k presentan un mayor producción de leche diaria y por lactación, y las mayores producciones por lactación de caseína total y caseína α_s. Estas relaciones no se mantiene cuando se corrige para los efectos fijos (año de parto, tipo de parto, número de lactación, estación de parto), debido posiblemente al bajo número de datos de polimorfismos confirmados.

Los resultados que hemos obtenido no nos permiten sugerir la selección para un mayor contenido en caseínas a través de la elección de determinados genotipos para las distintas fracciones de caseína (α_{s1}, α_{s2}, k), pero sí que existe suficiente variabilidad fenotípica, y posiblemente genética, como para poder comenzar a aplicar un criterio de selección por la cantidad de caseína total en la leche.

Tabla 1. Valores F y niveles de significación correspondientes a los análisis de varianza de los datos obtenidos con los controles mensuales. N = 5221.

	Prod	Prot	Cas	Gras	Lact	Alfa	Beta	Kappa
Raza(1)(2)	1.01ns	0.02ns	0.17ns	1.34ns	4.57ns	0.84ns	2.29ns	2.04ns
Reb(Raza)(5)	219***	27.6***	62.2***	32.7***	12.3***	53.5***	72.1***	32.1***
Año(3)	6.3**	111.6***	84.2***	125***	392***	228***	97.9***	77.7***
Estación(3)	6.7***	2.6*	14.9***	14***	16.3***	9.7***	4.8**	1.8ns
Nº Lact(9)	13.7***	1.4ns	2.7**	2.51**	2.8**	0.8ns	2.3*	1.0ns
Tipo parto(4)	5.9***	4.0**	8.01***	3.1*	3.4**	1.3ns	3.7*	1.3ns
R^2	0.26	0.1	0.11	0.15	0.20	0.14	0.16	0.1

(1) Se ha utilizado el cuadrado medio correspondiente al factor Reb(Raza) como término de contraste.
Los números entre paréntesis indican los grados de libertad
*: 0,01 < P < 0,05; **: 0,001 < P < 0,01; ***: P < 0,001
ns: no significativo

Tabla 2. Valores F y niveles de significación correspondientes al análisis de varianza (GLM) de los datos de producciones totales y contenidos medios por lactación incluyendo en el modelo el genotipo de la caseína $\alpha s1$, y la duración de la lactación como covariable. Los datos son procedentes de la raza Payoya. Los números entre paréntesis indican los grados de libertad. Número total de observaciones: 64

	αs1-CN	αs2-CN	k-CN
Producción	0.4ns	0.1ns	3.5*
Proteína	0.8ns	0.1ns	2.5ns
Caseína	0.9ns	0.1ns	3.6*
Grasa	0.9ns	0.1ns	2.7ns
Lactosa	0.3ns	0.1ns	2.8ns
αs-CN	0.9ns	0.1ns	4.4*
ß-CN	1.5ns	0.1ns	2.3ns
k-CN	0.2ns	0.1ns	0.9ns
% Proteína	1.1ns	0.4ns	2.4ns
% Caseína	4.4**	2.1cs	5.3*
% Grasa	1.2ns	1.7cs	0.3ns
% Lactosa	0.2ns	3.2cs	0.3ns
% αs-CN	3.8*	4.6*	0.2ns
% ß-CN	4.1*	0.8ns	0.5ns
% k-CN	0.5ns	0.1ns	0.4ns

*: 0,01 P <0,05; **: 0,001 P< 0,01; ***: P<0,001; c.s.: 0,09<P<0,05n.s.: no significativo

COSTE ECONOMICO DEL PROGRAMA DE MEJORA GENETICA Y SELECCION DE LAS OVEJAS DE RAZA LATXA Y CARRANZANA EN LA COMUNIDAD AUTONOMA DEL PAIS VASCO

E. Ugarte, E. Urarte, F. Arrese, J. Arranz, I. Beltrán de Heredia. CIMA- Granja Modelo. Arkaute. 01080 VITORIA-GASTEIZ

INTRODUCCION

El programa de selección y mejora genética de la oveja Latxa y Carranzana se inició en 1984 con la elección, por ascendencia, de 35 machos para ser sometidos a prueba de progenie mediante el uso de la Inseminación Artificial (IA). Dicho programa se enmarca dentro de un programa mas general de mejora del ovino que se inició en 1982 con la creación de las asociaciones de criadores de ganado Latxo y Carranzano y la puesta en marcha del control lechero (CL).

El programa está financiado por diferentes entidades: Gobierno Vasco, Diputaciones Forales y Asociaciones de ganaderos. Existe una asociación de ganaderos de raza latxa en cada territorio histórico: (ELE en Gipuzkoa, AGORALA en Araba y ACOL en Bizkaia. En esta última están también integrados los ganaderos de raza Carranzana). Las tres asociaciones forman junto con la asociación navarra (ASLANA) la confederación de asociaciones: CONFELAC.

En el presente trabajo se expone el coste económico del programa de mejora genética y selección en la Comunidad Autónoma del País Vasco (CAPV). Esto constituye un ejemplo de los costes que conlleva la puesta en práctica de un programa de estas características

MATERIALES

Los datos que se van a presentar corresponden a 215 rebaños de la campaña 1992-1993. En esa campaña se controlaron, en la CAPV, 58.392 cabezas de ovino con 36.277 partos y se estimaron en total 21.985 lactaciones aunque se tomaron datos de 31.070 ovejas. El número de ovejas controladas supone el 25 % del censo total de la población. El CL se realiza mediante el sistema AT (control lechero mensual alternando un control por la mañana y el siguiente por la tarde). La campaña de CL se inicia el 1 de septiembre y termina el 31 de agosto del año siguiente. Anualmente y antes del inicio de la paridera se realiza un recenso identificando todos los animales del rebaño. Un elemento básico de la recogida de información es el libro de partos donde se anotan: identificación de la oveja, fecha de parto, número de identificación transitorio de los corderos, sexo, incidencias, etc.

La campaña de IA se inicia a finales de mayo y termina a principios de octubre. La IA se realiza con semen refrigerado por vía cervical sobre celos inducidos por tratamiento hormonal. Dicho tratamiento se realiza con esponjas vaginales impregnadas de progestágeno colocadas durante 12 días e inyectando PMSG en el momento de la retirada. La inseminación se realiza 56 horas después de la retirada. El número de inseminaciones realizadas fue de 8.946 en 147 rebaños.

ANALISIS ECONOMICO

Se ha dividido el coste total en cuatro capítulos:

1. Coste de las asociaciones
2. Coste del centro de selección
3. Subvenciones a la autoreposición
4. Apoyo tecnológico y de investigación

COSTE DE LAS ASOCIACIONES

Las asociaciones de ganaderos criadores de raza Latxa y Carranzana son asociaciones abiertas y libres y están reconocidas como entidades colaboradoras en materia de producción y sanidad (vegetal y animal). Cada asociación establece un convenio cuatrianual con su respectiva Diputación y en el mismo se comprometen a realizar las acciones sanitarias que se estimen oportunas por el Gobierno y/o Diputaciones; gestionar el libro genealógico; controlar producciones, aportar los datos necesarios, etc.

Los ganaderos asociados pagan en cada asociación una cuota por rebaño y oveja. Además también pagan una cuota por inseminación y una cuota por los diferentes servicios que se les ofrece desde la asociación como alimentación y gestión técnico-económica, realización de proyectos de construcción, etc... Todos estos servicios están integrados junto con el programa de mejora genética en el programa de mejora del ovino y están íntimamente ligados entre sí. Las cuotas son diferentes en cada territorio en función de lo que cada asociación haya decidido.

El coste que supone para las asociaciones el control lechero y el programa de mejora genética es el siguiente:

CONCEPTO	COSTE (PESETAS)
PERSONAL 6 controladores a tiempo total y 4 a tiempo parcial	28.900.000
INFORMATIZACION	1.900.000
MATERIAL crotales, papel, etc.	2.000.000
PERSONAL TECNICO 3 técnicos y 2 coordinadores	15.600.000
TRATAMIENTOS HORMONALES	5.000.000
COMPRA DE CORDEROS	2.600.000
GASTOS DE FUNCIONAMIENTO	3.765.000
TOTAL	59.765.000

En el apartado de personal técnico y gastos de funcionamiento se han contemplado partidas que no son directamente adjudicables al programa de mejora genética como son el asesoramiento al ganadero en temas de alimentación, sanidad e higiene, manejo de animales,

gestión forragera, reuniones informativas, juntas y asambleas de asociaciones, etc... Sin embargo, se han tomado en cuenta porque el programa de mejora genética independiente de estas acciones no tendría sentido y no se obtendrían resultados.

Estos costes son pagados entre las Diputaciones Forales y los propios asociados en una proporción del 80 % y del 20 % respectivamente.

COSTE DEL CENTRO DE SELECCION

ARDIEKIN, S. L. es la sociedad que gestiona el centro de selección de machos de IA de las razas Latxa y Carranzana. La función principal de dicho centro es impulsar y potenciar el programa de selección y mejora genética. Uno de los puntos en los que trabaja es el testaje de machos y difusión de los machos mejorantes mediante el uso de la IA. Las distintas fases que siguen los machos son: selección, por ascendencia, de los corderos que entran al centro de selección artificial; eliminación de machos por no ser utilizables en IA; testaje sobre descendencia y difusión de machos mejorantes con tres objetivos fundamentales: producción de corderas de reposición, producción de machos para ser utilizados en monta natural y producción de machos para entrar al centro de selección.

Actualmente existen en el centro de selección 297 machos. De ellos 39 son machos mejorantes, 114 son machos que están a la espera de pasar la comisión de selección para poder ser testados este año, 99 son machos que están a la espera de que salgan sus pruebas y 45 son corderos que han entrado este año para ser testados el año que viene. Está previsto que entren otros 75 corderos. Los costes que ocasiona el centro son los siguientes:

CONCEPTO	COSTE (PESETAS)
PERSONAL	11.000.000
INSEMINACIONES	400.000
ALIMENTACION DE MACHOS	1.500.000
GESTION	500.000
CONTROL DE PATERNIDADES	1.200.000
SANIDAD	300.000
VARIOS	500.000
TOTAL	15.100.000

Este coste se financia en un 80 % a través de instituciones públicas directamente. El resto se financia a través de diferentes proyectos en los que el centro está involucrado (CDTI, AQUITANIA-EUSKADI-NAVARRA)

SUBVENCIONES A LA AUTOREPOSICION

Este es un punto clave en el programa de mejora genética y selección ya que dichas subvenciones están diseñadas como un incentivo para que la reposición del rebaño se haga de forma lógica y la misma se deje de los mejores animales. En un primer momento los mejores se definían desde el punto de vista productivo, pero a partir de 1993 son definidos según los valores genéticos. Dentro del grupo de ganaderos que realizan CL existe un grupo que, además de realizar CL, realiza IA. Este grupo es el grupo de ganaderos seleccionadores. Las subvenciones que reciben los ganaderos para la autoreposición son las siguientes:

SUBVENCIONES A LA AUTOREPOSICION (PESETAS)	HEMBRAS		MACHOS
	HIJAS DEL 10% MEJOR DE LAS OVEJAS	HIJAS DEL 50% MEJOR DE LAS OVEJAS	HIJOS DEL 10% MEJOR DE LAS OVEJAS
GANADERO SELECCIONADOR	6.000	5.000	10.000
GANADERO NO SELECCIONADOR	4.000	3.000	6.000

No se subvenciona más del 30 % de las hembras ni más del 3 % de los machos. En la campaña pasada el montante total en este apartado fue de 25.500.000 de pesetas

APOYO TECNOLOGICO Y DE INVESTIGACION

En este apartado entran los salarios del personal técnico que desde el CIMA-IKT (organismo dependiente del Departamento de Agricultura del Gobierno Vasco) trabaja de forma directa, aunque no a tiempo total, (1 coordinador y dos técnicos) en el programa de mejora genética y selección. Los proyectos que tienen relación directa con el programa de mejora genética y selección son:

.- Desarrollo de la selección (proyecto externo del Gobierno Vasco)
.- Optimización de los parámetros del esquema de selección (proyecto CICYT)
.- Efecto del fotoperiodo sobre la calidad seminal en un centro de selección e IA (proyecto CDTI)
.- Comparación genética con la raza Manech (equivalente francesa de la Latxa, proyecto transfronterizo)

En este apartado se han imputado 10.000.000 de pesetas

Sumando los cuatro apartados nos da un coste total de:

COSTE TOTAL	110.365.000 PESETAS

ANALISIS DE RESULTADOS

Es necesario indicar que los ganaderos de ovino de raza Latxa y Carranzana de la CAPV tienen una gran importancia no sólo económica (debida a la venta de leche y corderos) sino también social y ecológica. Esto último se debe a que el tipo de alimentación de los animales aprovechando los recursos naturales de las zonas marginales mas desfavorecidas ha mantenido la productividad y diversidad de la tierra, evitando así la despoblación y erosión de dichas zonas.

¿Qué se obtiene con este dinero? El programa de mejora incide directamente sobre los ganaderos en CL pero también repercute en el resto de la población. La conexión entre estos dos grupos de ganaderos se hace a través de la venta por ARDIEKIN, S. L. de los machos que no pueden ser utilizados en IA y a través de la venta de animales (machos y hembras) entre ganaderos de los dos grupos. Sin embargo, no se tiene información sobre los datos productivos de los ganaderos que están fuera del CL por lo que el análisis lo efectuaremos con los dados de CL teniendo siempre en cuenta que el dinero invertido en ellos también repercute en el otro grupo de ganaderos.

El programa de mejora del ovino ha obtenido resultados importantes desde su puesta en marcha en 1982. Así el incremento en lactación real durante estos 13 años se cifra en un 55 % aproximadamente. Esto se ha debido tanto a un aumento de la producción como a un aumento en la duración total de la lactación que ha pasado de 136 días a 150 días. En este sentido el aumento de producción diaria ha sido de un 37 %. El aumento también se ha dado en la lactación tipificada a 120 días que se utiliza como criterio de selección siendo este aumento de un 54 %. Además de conseguir resultados productivos se cumplen otro tipo de objetivos que se pretendían con el uso de la IA. Así, de las hembras que nacen de IA se deja el 65 % para reposición. En machos este porcentaje es del 8%.

Si tomamos los datos del CL, en esa campaña se han producido 4.469.236 litros de leche (no se contabiliza la leche de las ovejas que crían a las corderas y cuya producción no se registra) lo que supone que el coste aproximado del litro de leche que es debido al programa de mejora se cifra en 25 pesetas. Esa cantidad de leche supone un aumento de 238.494 litros respecto al año 1992. Si tenemos en cuenta que el litro de leche se está pagando aproximadamente a 100 pesetas obtenemos que con la inversión realizada, principalmente por organismos oficiales, los ganaderos asociados están obteniendo una rentabilidad de 23.849.400 pesetas. Analizando estos datos teniendo en cuenta el número de ovejas controladas se obtiene que este coste económico supone una inversión de 1.890 pesetas por oveja controlada.

Estos datos se refieren únicamente a datos productivos. Sin embargo, este programa también ha repercutido en una mejora de la calidad y condiciones de vida de los ganaderos. Se han mejorado sistemas de manejo, tipos de alimentación, racionalización del trabajo, mejora de las condiciones de habitabilidad de los apriscos en el monte, mejora de los sistemas de comunicación, etc.

DESARROLLO Y PERSPECTIVAS DEL ESQUEMA DE SELECCION DE LA RAZA OVINA SEGUREÑA

ANALLA M., MUÑOZ-SERRANO A., SANCHEZ A. y SERRADILLA J.M.

Laboratorio de Fisiogenética. ETSIAM. Universidad de Córdoba.

INTRODUCCION

Uno de los objetivos del convenio firmado entre la Consejería de Agricultura y Pesca de la Junta de Andalucía, la Asociación Nacional de Criadores de Ovino Segureño (ANCOS) y el Laboratorio de Fisiogenética de la Universidad de Córdoba, estipula la mejora de la producción de carne de ovino segureño. Se pretende mejorar la productividad numérica de las ovejas; pero, para que no haya un deterioro del peso al sacrifico (≅ 90 días) debido a la reducción de peso en animales nacidos en partos múltiples, se intentará mejorar también el crecimiento de los corderos. Actualmente hay un núcleo de selección formado por doce ganaderías. También se dispone de un rebaño experimental propiedad de la Diputación de Granada, que está en control de producciones desde 1981.

El objetivo que se ha planteado el equipo del laboratorio de Fisiogenética es establecer las bases para el esquema de selección, esto es:

- Estimar los parámetros genéticos y definir los modelos para la valoración de animales para crecimiento y prolificidad.

Estudiar las condiciones óptimas de selección, en un núcleo pequeño, considerando la conectividad y la consanguinidad.

RESULTADOS

A partir de la base de datos de control de producciones entre 1981 y 1991, en el rebaño de la granja experimental de Huéscar, se han estimado los parámetros genéticos con el paquete DFREML (Meyer, 1993) para los datos de crecimiento. El análisis según el modelo de valores aditivos (Quaas y Pollak, 1980) dió

como resultado una correlación genética bastante alta (= 0.7) tanto del peso al destete (= 45 días) como del crecimiento predestete, con el peso a los noventa días. La heredabilidad más alta fue para el peso al destete. Otro análisis llevado a cabo según el modelo con efectos maternos (Quaas y Pollak, 1980), confirmó la alta heredabilidad del peso al destete y al mismo tiempo reveló una correlación aditiva directa más alta entre el peso a los noventa días y el peso al destete (tabla 1). Por todo ello, se optó por el peso al destete como criterio precoz de selección para mejorar el peso a los noventa días. En todos los casos, la proporción de varianza total explicada por efectos aditivos directos fue mucho más alta que la explicada por efectos aditivos maternos (relación de ocho a uno en el caso del peso al destete). Por lo que se pensó utilizar para la valoración un modelo simple como es el de valores aditivos. Por ello actualmente se está comparando, por simulación, la eficiencia de este modelo con el de efectos maternos.

En las doce ganaderías se han realizado dos campañas de inseminación con machos en prueba y a finales de este año se dispondrá de información de los partos de sus hijas.

Quedan por estimar la heredabilidad y la repetibilidad de la prolificidad. Para ello se piensa utilizar una metodología lineal, ya que las estimaciones con técnicas específicas para caracteres umbrales no son mejores, incluso a veces éstas están fuera del espacio parametrico (heredabilidad superior a la repetibilidad, Matos et al. 1993). También, habrá que estimar las correlaciones entre prolificidad y crecimiento. Debido a que el núcleo de selección es pequeño (doce ganaderías) convendría estudiar el comportamiento de la consanguinidad, para establecer el número de machos de inseminación artificial, necesarios para conectar los rebaños sin aumentar mucho la consaguinidad en los mismos.

Tabla 1. Heredabilidades (en la diagonal), correlaciones genéticas (encima de la diagonal) y correlaciones fenotípicas (debajo de la diagonal). Para los caracteres: peso al destete (PD), peso a los noventa días (P90) y crecimiento predestete (CPD).(m^2 corresponde a la proporción de varianza fenotípica explicada por efectos aditivos maternos).

	CPD	PD	P90	m^2
CPD	0.36 (0.29)	0.96 (0.95)	0.48 (0.72)	0.07
PD	0.93 (0.95)	0.51 (0.34)	0.74 (0.77)	0.06
P90	0.71 (0.70)	0.75 (0.74)	0.34 (0.31)	0.12

( ) valores entre paréntesis se han obtenido con un modelo de valores aditivos mientra que los demás son de un modelo con efectos maternos.

REFERENCIAS

MATOS C.A.P., D.L. THOMAS, D. GIANOLA, R.J. TEMPELMAN y L.D. YOUNG 1993. J. Anim. Sci. (en prensa).

MEYER K. 1993. DFREML users' notes.

QUAAS R.L. y E.J. POLLAK 1980. J. Anim. Sci. 51:1277-1287.

DETECCION DE PORTADORES DE LA DEFICIENCIA DE ADHESION LEUCOCITARIA BOVINA (BLAD/LAD Bovine) EN REPRODUCTORES UTILIZADOS EN ESPAÑA.

ZARAGOZA, P., ROMERO, A. y GARCIA-MURO, E.

Laboratorio de Genética Bioquímica y Grupos Sanguíneos. Facultad de Veterinaria. Miguel Servet, 177 - 50.013 ZARAGOZA (Spain).

INTRODUCCION

Un gran número de investigadores en todo el mundo están trabajando en la detección de marcadores del DNA ligados a asociados a genes de importancia económica, no sólo en producción, sino también en resistencia a enfermedades (MAS = Selección Asistida por Marcadores). El objetivo de estos programas es incrementar la respuesta a la selección, aumentando la precisión y disminuyendo el tiempo. En la actualidad, la mortalidad de terneros en granjas de leche supone graves pérdidas económicas. Algunos autores, indicaron que en ganado Holstein, un 7,7% de mortalidad se producía al nacimiento y en un 36% de los casos el agente etiológico era desconocido y atribuía la muerte a diarrea.

Recientes investigaciones del Centro Nacional de Enfermedades Animales del Servicio de Investigación Agrícola de USDA en Ames (Iowa), han descrito la base genética y molecular de este síndrome, denominando la enfermedad como BLAD (Bovine Leucoite Adhesion Deficience) ó actualmente LAD Bovino, (KEHRLI et al., 1992). El BLAD ó LAD Bovino, es una enfermedad controlada por un gen autosómico letal recesivo, que afecta hasta el momento a ganado vacuno Holstein. Los terneros afectados tienen una glicoproteína de los leucocitos alterada. La deficiencia de esta glicoproteína leucocitaria, denominada Mac-1 (Complejo CD11/CD18), impide la adhesión de los leucocitos a los receptores de membrana, impidiendo el transporte normalmente desde la sangre a los lugares de infección. Por esta razón, los terneros con BLAD son muy susceptibles a infecciones bacterianas y fúngicas y expresan menos del 2% del nivel normal de la glicoproteína CD18 (KEHRLI et al., 1990). La sintomatología característica es la siguiente: Los teneros al nacimiento son aparentemente sanos, pero a los pocos días, cursan confiebre alta, diarrea crónica, falta de cicatrización en heridas, gingivitis, graves infecciones, aumento del número de leucocitos y muerte del animal de las 7 semanas a los 8 meses de vida (NAGAHATA et al., 1987; TAKAHASHI et al., 1987).

BASE MOLECULAR EN GANADO VACUNO

El síndrome de granulocitopatía bovina se debe a un defecto de la expresión de la proteína CD18, por una secuencia anormal en el gen, que originaba la falta de la subunidad α en la integrina $\beta 2$ (proteína que atraviesa la membrana plasmática) de las membranas de los neutrófilos aislados de animales enfermos. Esto origina una alteración en el transporte de los leucocitos desde la sangre a los lugares de infección. A partir del RNA mensajero aislado de

leucocitos animales afectados y mediante una transcriptasa inversa, se obtuvo el cDNA del gen CD18 (SMITH et al., 1989). Al realizar su secuenciación se observó que en el alelo raro, aparecen dos puntos de mutación, una silente y otra que origina la sustitución de una adenina por una guanina en la posición 28. Esto origina el cambio del aminoácido glicina por ácido aspártico en la posición 128 (D128G) de la secuencia de la proteína que codifica el gen (KEHRLI et al., 1992). Así con respecto a este gen CD18, aparecen animales con tres tipos de genotipos:

D128G/D128G: Individuos omocigóticos recesivos, que padecen la enfermedad;
D128/D128G: Individuos heterocigotos, portadores de la enfermedad y que no la padecen;
D128/D128: Individuos homocigóticos dominantes, sanos.

DESCUBRIMIENTO Y APARICION

Las características clínicas y patológicas en vacuno, se presentaron muy similares a las previamente descritas en pacientes humanos que tienen una alteración denominada "Deficiencia de Adhesión Leucocitaria o LAD"(TODD y FREYER, 1988). También los perros de la raza "Setters irlandeses" y que padecían el "síndrome de granulocitopatía canina" presentaban síntomas muy parecidos (GIGER et al., 1987). Tanto en humana como en perros estos síndromes están relacionados con una deficiencia de tipo genético del complejo glicoproteíco (CD11/CD18) de los leucocitos.

En humana, los pacientes de LAD tienen carencia o deficiencia a nivel estructural y funcional de las glicoproteínas Mac-1, LFA-1 y p150.95. Estas glicoproteínas hoy se denominan con el término CD11/CD18, según la organización mundial de la salud.

Cada una de estas moléculas contienen una α y una β subunidad, asociadas no covalentemente, con una estructura $\alpha_1 \beta_1$. Todas estas glicoproteínas comparten una subunidad β idéntica (CD18) y se distinguen inmunológicamente por su subunidad α, designada CD11a, CD11b y CD11c, para LFA-1α, Mac-1α y p150.95 α, respectivamente.

Una mutación del gen que codifica la subunidad β (CD18), fue identificada como la base primaria de la enfermedad en todos los casos (HIBBS et al., 1990).

La patogénesis de la enfermedad en perros Setters irlandeses, no ha sido completamente estudiada, encontrándose una expresión deficiente de Mac-1, LFA-1 y p150.95 sobre la superficie de los leucocitos, lo que parece sugerir de forma similar a la humana un defecto del gen que codifica la CD18 canina (GIGER et al., 1987).

DIAGNOSTICO

El método de diagnóstico desarrollado consiste en la amplificación específica de un fragmento del gen CD18 (58 pares de bases); mediante la técnica de Reacción en Cadena de la Polimerasa (PCR) (ZARAGOZA, et al., 1944). Tras posterior digestión con enzimas de

restricción y visualización mediante electroforesis en gel de agarosa, pueden diferenciarse los tres genotipos posibles.

El producto amplificado se puede digerir con distintos enzimas de restricción. Así, cuando el enzima de restricción es Taq I, esta enzima reconoce en el alelo normal (D128) un sitio de restricción (T / C G A), originando dos fragmentos de 33 pb y 25pb, mientras que en la mutación D128G no reconoce ningún sitio de restricción, dando en este caso un fragmento de 58pb. Si por el contrario las enzimas de restricción usadas son Hae III ó Pal I, éstas reconocen en el alelo normal (D128), un único sitio de restricción (G G / C C), originando dos fragmentos de 9pb y 49pb. En el alelo mutado (D128G), estas enzimas reconocen dos sitios de restricción, originando fragmentos de 9pb, 19pb y 30pb.

Este método además ofrece grandes ventajas:

- Rapidez: en 24-48 horas pueden ofrecerse los resultados.
- Alto rendimiento: se puede partir de cantidades mínimas deDNA y ello supone el poder diagnosticar esta enfermedad incluso en una sola célula y obtener 10^9 veces la cantidad de DNA de la que hemos partido.
- Automatización: lo que conlleva su utilidad en el servicio y en el procesado de muchas muestras.

La tipificación realizada hasta ahora en nuestro laboratorio ha sido en 123 muestras de las cuales 40 pertenecen al test de Comparación Internacional de la ISAG en individuos machos de distintas razas como Holstein, Limousin, Angus, Hereford, Brahman y Simmental, en el objetivo de standarizar la técnica. El resto 83 animales Holstein pertenecientes a programas de Inseminación Artificial o en período de prueba de Centros como C.E.N.S.Y.R.A. de Movera ó empresa como Tauste Ganadera o Aberekin.

REPERCUSION

Es importante indicar que es obvio que la presencia de esta enfermedad no supone una "catástrofe económica" puesto que ya lleva persistiendo, desde hace mucho tiempo. Sin embargo, con los métodos de selección y los métodos de reproducción usados actualmente, principalmente, en vacuno de leche, es necesario hacer el diagnóstico en todos los machos candidatos a reproductores y en las hembras usadas en lor programas de MOET, así como todas aquellas que hayan sido cruzadas con machos portadores. Estas y otras precauciones pueden permitir manejar la enfermedad de forma adecuada y con especial atención en los programas de selección.

La Asociación Holstein Americana ha decidido designar los toros portadores de BLAD (D128/D128G) con letras BL después del nombre. Los animales sanos, no portadores (D128/D128), llevarán el sufijo TL.

CONTRIBUCION AL CONOCIMIENTO DEL MAPA BOVINO (BOVMAP PROJECT): PRIMEROS RESULTADOS OBTENIDOS EN LA BUSQUEDA DE NUEVOS MARCADORES.

ELDUQUE, C., MARCOS, S. y MARTIN, I.

Laboratorio de Genética Bioquímica y Grupos Sanguíneos. Facultad de Veterinaria. Miguel Servet, 177 - 50.013 ZARAGOZA (Spain).

El mapa génico de las distintas especies de interés en la producción animal, y en particular el mapa génico bovino, ha experimentado un espectacular desarrollo en los últimos diez años.

Aunque aún lejos del nivel alcanzado en el conocimiento de los mapas humano y de ratón, es de esperar que en los próximos años el número de marcadores asignados a regiones cromosómicas determinadas, permitirá comprender el control genético de caracteres de importancia económica.

Distintos laboratorios en todo el mundo trabajan de una forma u otra en la elaboración del mapa génico bovino. De ahí que iniciativas para unir esfuerzos y realizar una investigación más efectiva, tuvieran como fruto la unión de diferentes laboratorios europeos en lo que se denominó Proyecto BOVMAP.

La metodología y estrategias utilizadas en la determinación del mapa bovino no difieren de las empleadas en la construcción de los mapas de cualquier otra especie mamífera. Es por esta razón que se utilizan los métodos y técnicas ya empleadas en las especies humana y el ratón.

Distintos métodos llevan a la construcción de dos tipos de mapas que se complementan, llamados mapa físico y mapa genético.

La construcción de los mapas físicos se realiza a partir de distintas metodologías, como son el análisis de clones de células somáticas híbridas, o mediante hibridación in situ en cromosomas metafásicos o el estudio del solapamiento de DNA clonado.

Los mapas genéticos son el resultado de los análisis de ligamiento entre loci, que son función de la frecuencia de entrecruzamientos producidos en los cromosomas durante la meiosis. Para la realización de estos estudios de ligamiento, se ha realizado un esfuerzo internacional para conseguir grandes familias de hermanos completos y distribuir muestras de DNA a los distintos laboratorios que testan los diferentes marcadores.

La comparación entre los mapas de distintas especies reveló la existencia de fragmentos cromosómicos conservados durante la evolución. El mapa génico bovino actual contiene unos 300 loci. Muchos de estos loci representan genes codificantes muy conservados entre especies, que se han denominado loci Tipo I. Estos loci son poco polimórficos en general, lo que supone un inconveniente en su utilización en análisis de

ligamiento, sin embargo permiten la identificación de segmentos de sintenia conservados entre especies.

La información de los mapas de las distintas especies es de una gran utilidad para la construcción de cualquier mapa génico, pues permite orientar el estudio de marcadores determinados en zonas poco saturadas de marcadores por comparación con grupos de sintenia conservados.

Por otra parte, los denominados loci Tipo II, son marcadores muy polimórficos - microsatélites y minisatélites- idóneos en estudios de ligamiento, que serán de gran utilidad en la búsqueda de marcadores de diferentes caracteres productivos, así como de enfermedades genéticas o de resistencia a enfermedades.

Dentro del Proyecto BOVMAP, nuestro laboratorio ha participado en la producción y testaje de marcadores, tanto Tipo I como Tipo II. Para ello se ha realizado la extracción de DNA de muestras a partir de sangre, semen, leche y embriones.

Se han recogido un total de 1200 muestras de DNA pertenecientes a varias razas autóctonas españolas: Asturiana de los Valles, Asturiana de las Montañas, Rubia Gallega, Pirenaica, Avileña y otras razas explotadas en España como Frisona, Limousin y Parda Alpina.

Hasta ahora hemos estudiado 18 marcadores de DNA: 4 de Tipo II y 14 de Tipo I, utilizando tanto PCR como técnica de Southern Blotting.

Cuatro microsatélites (marcadores tipo II) han sido estudiados:

CYP21 en el que se han detectado 14 alelos

D21S4 con 17 alelos

ILSTS005 con 5 alelos

Microsatélite de la K-caseína con 6 alelos

Dentro de los marcadores tipo I se han testado con PCR los genes κ-CN, β-CN, β-LG y **ZFX/ZFY**. Un nuevo RFLP con 2 alelos ha sido detectado en el gen **ZFX** y lo hemos denominado ZES10.

Se han realizado estudios poblacionales en el caso de los genes relacionados con proteínas de la leche llegando a las siguientes conclusiones. El alelo β de la κ-caseína es más frecuente en las razas autóctonas y no así en las dedicadas a producción láctea, como sería deseable. Este alelo está asociado con una mejor calidad de leche para derivados como el queso.

Para la β-caseína el alelo A es el más frecuente en todas las razas estudiadas. En el caso de la β-lactoglobulina se han observado frecuencias intermedias. También se han analizado 8 marcadores tipo I mediante la técnica de Southern Blotting, utilizando sondas heterólogas humanas. Los loci estudiados han sido amilasa-1 (**AMY1**), albúmina (**ALB**), hemoglobina (**HBB**), transferrina (**TF**), nucleósidofosforilasa (**NP**), proteína activadora de esfingolípido-2 (**SAP-2**) y la subunidad beta del factor de crecimiento nervioso (**NGFB**).

Asimismo, en colaboración con el laboratorio del Dr Womack, nuestro laboratorio ha contribuído en el conocimiento del mapa físico con la asignación de dos loci (la proteína activadora de esfingolípidos-2 (**SAP-2**) y la unidad beta del factor de crecimiento nervioso (**NGFB**) a los grupos de sintenia U29 y U6 respectivamente, por medio del análisis de un panel de células híbridas.

GRUPOS SANGUINEOS VS POLIMORFISMOS DEL DNA EN LA IDENTIFICACION DEL GANADO VACUNO: UNA CONTROVERSIA EN VIAS DE SOLUCION.

RODELLAR, C., VAZQUEZ, F., OSTA, R. y ZARAZAGA, I.

Laboratorio de Genética Bioquímica y Grupos Sanguíneos. Facultad de Veterinaria. Miguel Servet, 177 - 50.013 ZARAGOZA.

Los constantes avances de la técnica analítica, incorporan permanentemente a la metodología laboratorial nuevas modalidades de trabajo e investigaciones y como resultado, se logran más amplias perspectivas de información científica y de posibilidades de transferencia de resultados aplicativos a los sectores implicados.

En el campo de la genética animal, se está produciendo una verdadera revolución en todas las vertientes de investigación abocadas a la producción de calidad y a la mejora de las agrupaciones animales. Efectivamente, tenemos por una parte la Genética Cuantitativa, con una concepción y estudio integral de los factores incidentes en la mejora. En su metodología se utiliza con gran profusión la recogida y tratamiento informático de datos, -cada día más abundantes-, una gran velocidad de cálculo y mayores posibilidades de almacenamiento dinámico en bases interactivas. Por otro lado, tenemos la moderna Citogenética que permite una aportación totalmente nueva a las producciones animales. Tampoco podemos olvidar la potente y revolucionaria metodología que nos aporta la Biología Molecular y la Genética Bioquímica, que basándose en el estudio de material genético (DNA) nos han prestado y nos están proporcionando permanentemente un papel imprescindible en la identificación genética y chequeo de parentesco, sin olvidar por supuesto, el diagnóstico de enfermedades hereditarias y la aportación de marcadores genéticos que apoyan la selección de caracteres cuantitativos o económicos (por ejemplo, los relacionados con los tipos de caseínas y la calidad de la leche o con el sexaje precoz de embriones) en los efectivos ganaderos.

Pero dentro del complejo campo del panorama de la Genética animal, nos salen al paso cuestiones muy importantes que quizá, para muchos, habían pasado hasta ahora inadvertidas. Surge, por ejemplo, la siguiente cuestión: ¿Podemos dar por buena, la entrada en prueba de reproductores o donadores de gametos, únicamente porque sus controles de fertilidad o producciones son buenos, sin recibir el apoyo inequívoco de una identificación genética y de chequeo de parentesco, que demuestren con la autenticidad de toda su constelación familiar, que estamos en el buen camino de aprovechamiento de una genética de élite?. Evidentemente no. Y además, podemos asegurar que su olvido, resultaría muy peligroso para el porvenir de la población de que se trate.

Pero es que incluso, dentro del capítulo de identificación genética y chequeo de parentesco queda todavía más por cuestionarse. El problema puede ser alarmante, si lo que se

exige como garantía de una buena metodología, base de la mejora, incumple no sólo con las exigencias científico-técnicas, sino con las normales legales y administrativas, admitidas todas ellas a nivel internacional.

Este es el clima de incertidumbre que se estaba detectando desde hacía varios años en el ambiente de los ganaderos de elite y de los técnicos especialistas en la materia. Efectivamente, se ha reconocido que el progreso científico aplicado adecuadamente, aporta novedades y posiblemente en muchos casos en el futuro, una mayor eficacia. Pero también, un malentendido progreso acarrea complicación y confusión, cuando se intenta adaptarlo sin la debida información de sus limitaciones y riesgos en su desarrollo. En dicho contexto, se ha llegado a crear en algunos sectores una especie de confrontación de la metodología de grupos sanguíneos respecto al análisis del DNA en el ganado vacuno. Afortunadamente, esta incógnita ha sido despejada tras la discusión objetiva y abierta en los medios internacionales.

¿Cuáles son los pasos que se han dado y las garantías que prestan esta definitiva postura objetiva, que admite la evolución y adaptación de las técnicas en estudio y a la vez no olvida la importancia de exigir estrictamente la norma consagrada y valedera? Podemos decir que ha sido una larga controversia aclarada y defendida al más alto nivel internacional y que se está en el mejor camino de armonización y adecuación al progreso tecnológico.

A modo de resumen indicaremos hechos significativos de esta larga trayectoria, finalmente aclarada por la International Society of Animal Genetics (ISAG), sociedad integrada por un grupo de investigadores con autoridad Internacional que respalda un Programa de Pruebas de Comparación entre Laboratorios, con generosa apertura al progreso científico y a la innovación crítica y con reconocido prestigio internacional. La ISAG, además, consta de una serie de laboratorios que figuran como miembros institucionales, y que participan activamente en la estandarización de tecnologías y en los servicios de tipificación de marcadores en especies ganaderas.

En lo referente a la especie bovina, la ISAG reconoce:

1.- Que la identidad de los individuos (como miembro concreto de una genealogía o familia) puede ser probada como errónea a través de identificación por marcadores genéticos. En consecuencia, las asociaciones de ganaderos eliminan de los libros genealógicos a dichos individuos, los erróneamente asignados a una genealogía (pedigree).

2.- Que los marcadores genéticos utilizados en la actualidad son los grupos sanguíneos (antígenos de la membrana eritrocitaria) y polimorfismos bioquímicos (variantes proteicas sanguíneas detectadas por electroforesis). Con ellos puede llegarse a un nivel de fiabilidad del 99%. Estos marcadores y las técnicas necesarias para su detección son los consesuados por la ISAG. Ello facilita el seguimiento de pedigrees en programas nacionales e internacionales de verificación de paternidades (transferencia de embriones, inseminación artificial, registro en libros genealógicos, compra-venta, importación-exportación, etc.).

3.- Que el DNA, a través de las distintas tencologías laboratoriales (PCR, soutl.ern blotting) y de los distintos tipos de marcadores genéticos (microsatélites, minisatélites, VNTR, RFLP, etc.) puede dar lugar a niveles de alta fiabilidad mayor o igual al 99%, siempre y cuando el número y tipo de marcadores estudiados sea suficiente y adecuado para llegar a dicho nivel.
4.- Que a nivel oficial, los libros genealógicos de razas han estado hasta la fecha (y lo siguen estando en la actualidad) autentificados solamente mediante grupos sanguíneos y polimorfismos bioquímicos (dicha normativa es la actualmente vigente dentro de la Comunidad Europea).
5.- Que dado el avance tecnológico surgido mediante las tencologías de DNA, conviene investigar acerca de qué marcadores de DNA podrían ser candidatos para su utilización en identificación, en base a criterios: informativos, económicos, fáciles de interpretar, etc.
6.- Que la tecnología y tipo de marcadores de DNA a utilizar debe consensuarse internacionalmente antes de ser aplicada oficialmente en identificación animal. De lo contrario, distintos laboratorios utilizarían distinta tecnología y hoy se utilizará un marcador y mañana otro. Ello derivaría en la imposibilidad de verificar genealogías (de abuelos a padres, de padres a hijos, etc.) tanto a nivel nacional como internacional, resultando en una situación caótica en los libros genealógicos de razas y en las propias explotaciones.
7.- Que una vez consensuada la tencología de DNA dentro de unos años, comenzará un período en el que el laboratorio de la ISAG encargado de la identificación animal en cada país realizará a la vez los tests de grupos sanguíneos + polimorfismos bioquímicos que tradicionalmente ha venido desarrollando y los de DNA, todos ellos correspondientes a marcadores consensuados. Con ello se logrará que los archivos de marcadores genéticos correspondientes a los libros genealógicos tengan una continuidad y puedan comprobarse las genealogías aún incorporándose la tencología del DNA.

Todo ello indica que si las cosas han de hacerse bien, en favor de la ganadería y con solidez en cada país, los protocolos y ensayos que se lleven a cabo de cara a la adopción de una nueva norma, en cuanto a las tencologías y tipos de marcadores se refiere, no han de servir en ningún caso para adoptar de manera automática como definitivos los marcadores ensayados. En nuestro país, ha de pasar en primer lugar la criba de la normativa comunitaria europea; en segundo lugar, la criba de la comunidad científica internacional (ISAG) en cuanto a tecnologías consensuadas y contrastadas; y en tercer lugar, la criba de los intereses y disponibilidades de las propias asociaciones/agrupaciones ganaderas.