

## APLICACION DE UN METODO MODIFICADO DE CALCULO DEL COEFICIENTE DE CONSANGUINIDAD EN UNA MUESTRA DEL GANADO VACUNO FRISON ESPAÑOL.

COMPUTING INBREEDING COEFFICIENT IN A SPANISH HOLSTEIN'S SAMPLE.

Gutiérrez, J.P. \*, J. Cañón \* y M. Rico \*\*.

\* Departamento de Producción Animal. Facultad de Veterinaria. Universidad Complutense. 28040 Madrid. España.

\*\*Departamento de Producción Animal. Escuela Técnica Superior de Ingenieros Agrónomos. Madrid. España.

Palabras clave: Consanguinidad. Endogamia. Método iterativo. Holstein.

Keywords: Inbreeding. Iterative method. Holstein.

### Summary

Average inbreeding coefficient from a sample of the Spanish Holstein population is computed using an iterative method. Although the positive result represented by the average inbreeding value obtained is close to zero ( $4.03 \times 10^{-4}$ ), more detailed analysis of the data shows that the results must be confirmed with more pedigree information.

### Resumen

Mediante un método iterativo se estudia la consanguinidad media de una muestra de la población del vacuno Frísón español. La endogamia media obtenida resultó despreciable ( $4,03 \times 10^{-4}$ ). Aunque este resultado parece positivo, un análisis detallado de los datos sugiere que los resultados no son definitivos, y que es necesario un estudio con un mayor volumen de información de pedigrí.

### Introducción

El cálculo de los coeficientes de consanguinidad en poblaciones de animales domésticos puede interesar para valorar de forma más precisa a los reproductores y para evitar el aumento de la misma.

Dichos coeficientes han sido calculados durante mucho tiempo mediante la detección de antepasados comunes por comparación de líneas ancestrales (Wright, 1922; Hazel y Lush, 1950; Abe y Nishida, 1971; Schaap y Cohen, 1977) o por detección de antepasados comunes por medio de algoritmos de búsqueda de caminos de parentesco (Rehfeld *et al.*, 1967; Alfonso Ponce, 1971 y 1985; Stevens, 1975; Nadot, 1971; Nadot y Vaysseix, 1973).

Quaas (1976) ha desarrollado un método iterativo que permite, al tiempo, construir la inversa de la matriz de relaciones aditivas para ser incluida en

las ecuaciones de los modelos mixtos, mediante la búsqueda de los padres de cada individuo repetidas veces.

Mediante un método iterativo, con una estrategia de computación ligeramente diferente de la propuesta por Quaas (Gutiérrez *et al.*, 1988), se estudia la consanguinidad de una muestra de la población de vacuno Frisón español.

### Material y métodos

**Datos.** El archivo de datos fue proporcionado por la D.G.P.A. y el archivo de pedigrís por A.N.F.E., siendo su estructura la siguiente:

La población sobre la que se calcula la consanguinidad está constituida por 22100 individuos (vacas y toros) de los cuales se consideran población base (no se conocen los padres) 8996.

De los 636 toros que aparecen en el análisis, sólo se conocen ascendientes de 335.

De las 21464 hembras existentes, únicamente 12769 tienen al menos un padre conocido.

El número de generaciones se calculó buscando el antecesor más lejano por cualquiera de las vías, paterna o materna, en cada uno de los individuos. No se pudo estimar el número de generaciones mediante la edad de los animales por carecer un gran número de ellos de la fecha de nacimiento.

El número máximo de generaciones existentes en esta población, calculado de la forma descrita, es de 6, y se distribuyen como se indica en

la Tabla I.

Tabla I. Distribución del número de generaciones en la población estudiada.

Individuos	número	p.100
con 1 generación	8997	40,71
con 2 generaciones	933	4,22
con 3 generaciones	8307	37,59
con 4 generaciones	3598	16,28
con 5 generaciones	263	1,19
con 6 generaciones	2	0,01

El número medio de generaciones encontradas por la vía más larga, sea esta la vía materna o la paterna, es de 3,24443.

**Método.** Si definimos  $u$  como el vector de valores genéticos aditivos de cada uno de los individuos de una población, es evidente que:

$$u_i = \frac{1}{2} u_j + \frac{1}{2} u_k + \Phi_i \quad [1]$$

siendo:

- $j$  y  $k$  los padres de  $i$ .
- $u_j$ ,  $u_k$  y  $u_i$  los valores aditivos de los individuos  $j$ ,  $k$  y  $i$ .
- $\Phi_i$  la desviación debida al efecto mendeliano de los genes en el individuo.

Si por otra parte, definimos  $P$  como la matriz de padres, es decir aquella en la que aparece un uno en el lugar donde se encuentra el padre del individuo  $i$ , podemos generalizar la

expresión [1] de la forma siguiente:

$$u = \frac{1}{2} P u + \Phi \quad [2]$$

o:

$$u = (I - \frac{1}{2} P)^{-1} \Phi$$

La varianza de u resulta ser igual a:

$$\text{Var}(u) = A \sigma_u^2$$

siendo:  $\sigma_u^2$  la varianza genética aditiva entre los individuos suponiendo a sus padres sin relación de parentesco entre ellos.

Por tanto:

$$A = (I - \frac{1}{2} P)^{-1} \frac{\text{Var } \Phi}{\sigma_u^2} (I - \frac{1}{2} P)^{-1} \quad [3]$$

Si suponemos que las desviaciones  $\Phi_i$  son independientes, se puede definir D como una matriz diagonal:

$$D = \left\{ \frac{\text{Var } \Phi}{\sigma_u^2} \right\}$$

Y llamado a T:

$$T = (I - \frac{1}{2} P)^{-1}$$

La expresión [3] se puede escribir de la forma:

$$A = T D^{\frac{1}{2}} D^{\frac{1}{2}} T'$$

Si a  $T D^{\frac{1}{2}}$  lo llamamos L la expresión [3] queda:

$$A = L L' \quad [4]$$

La inversa de la matriz A se puede expresar, por tanto, de la siguiente forma:

$$A^{-1} = (I - \frac{1}{2} P') D^{-1} (I - \frac{1}{2} P) \quad [5]$$

La única dificultad que puede aparecer al intentar construir la matriz  $A^{-1}$  es el cálculo de los valores  $d_{ii}$  para construir la matriz diagonal D

Tomando varianzas en [1] y dividiendo por  $\sigma_u^2$  se obtiene:

$$\frac{\sigma_{ui}^2}{\sigma_u^2} = \frac{1}{2} \frac{\sigma_{uj}^2}{\sigma_u^2} + \frac{1}{2} \frac{\sigma_{uk}^2}{\sigma_u^2} + \frac{1}{2} \frac{\sigma_{uj uk}}{\sigma_u^2} d_i$$

[6]

Sabiendo que:

$$a_{ii} = 1 + F_i = \frac{\sigma_{ui}^2}{\sigma_u^2}$$

$$F_i = \frac{1}{2} a_{jk} = \frac{\sigma_{uj uk}}{\sigma_u^2}$$

Sustituyendo en [6] se obtiene:

$$a_{ii} = 1 + F_i = \frac{1}{2} a_{jj} + \frac{1}{2} a_{kk} + \frac{1}{2} a_{jk} + d_i$$

$$d_i = 1 - \frac{1}{2} a_{jj} - \frac{1}{2} a_{kk}$$

Por tanto, conociendo  $a_{jj}$  y  $a_{kk}$ , tendremos todo lo necesario para calcular  $d_i$ . Para obtener el valor de  $a_{jj}$  y  $a_{kk}$  debemos conocer  $d_j$  y  $d_k$  que a su vez necesitan ser calculados. De este modo y ascendiendo en la genealogía

llegará un momento en que se desconozcan los padres, en cuyo caso se asume que son población base, y por tanto  $F_j$  y  $F_k$  son igual a cero.

Estrategia de cálculo. El archivo de datos debe estar organizado con los individuos numerados secuencialmente e incluidos en una matriz de dos filas, correspondiendo la primera al número del padre y la segunda al de la madre, y el número de la columna al del individuo.

El algoritmo de cálculo que se presenta, puede ser dividido en dos partes. En la primera se calcula la matriz diagonal  $D$  para, calcular la matriz  $A^{-1}$  mediante las sencillas reglas de Henderson (Henderson, 1976).

Si comenzamos por el primer animal (ordenado por fecha de nacimiento), para calcular  $a_{ii}$  recurrimos a la siguiente expresión, deducida fácilmente de [4]:

$$a_{ii} = \sum_{m=1,i} l_{mi}^2 \quad [7]$$

por tratarse del primer individuo se calcula sencillamente por:

$$a_{11} = l_{11}^2$$

y como sabemos que:

$$L = (I - \frac{1}{2}P)^{-1} D^{\frac{1}{2}}$$

$$d_i = 1 - \frac{1}{2} a_{jj} - \frac{1}{2} a_{kk}$$

entonces

$$d_{11} = 1$$

y

$$a_{11} = 1$$

Conocido el valor del primer animal, podemos rellenar la primera columna de la matriz triangular inferior  $L$  en un vector de trabajo de la siguiente manera: buscamos el primer hijo según el orden de nacimiento. Le asignamos a  $l_{1m}$  (siendo  $m$  el número de identificación del hijo encontrado)  $1/2$  del valor de  $a_{11}$  y pasamos a buscar un hijo del individuo  $m$ , para asignarle  $1/4$  del valor  $a_{11}$  y así hasta llegar al último individuo de la población para entonces retroceder nuevamente una generación y buscar un hermano del último individuo que se buscó. Así se continua el proceso hasta acabar con el último hijo del individuo 1.

Quaas (1976), calcula este vector, animal por animal, mediante la expresión:

$$l_{im} = 1/2 l_{jm} + 1/2 l_{km}$$

para todos los individuos de la población, lo que supone un tiempo de cálculo muy superior al que resulta con el algoritmo presentado.

Para comenzar nuevamente con otro individuo primero se eleva al cuadrado cada uno de los elementos del vector de trabajo ya mencionado para acumular el resultado en un segundo vector de trabajo con el fin de ir calculando la expresión [7].

Una vez calculados todos los elementos de la diagonal de la matriz  $D$  se obtiene el inverso de cada uno de ellos, y pasamos a la segunda parte de la subrutina en la cual se calcula la inversa de la matriz  $A$  con ayuda de las conocidas reglas de Henderson:

- En el cruce del individuo consigo mismo ponemos  $d_j$ .

## COEFICIENTE DE CONSANGUINIDAD.

- En el cruce del individuo con cada uno de sus padres se añade ( $-1 d_i$ ).
- En los lugares donde se cruzan los padres ( $+d_i$ ).

De este modo realizamos los cálculos pertinentes con cada uno de los individuos y la matriz  $A^{-1}$  queda construida.

La subrutina devuelve únicamente los elementos de  $A^{-1}$  que son distintos de cero, con su valor y la fila y columna que ocupan.

Para poder realizar una búsqueda más rápida se puede incluir otro vector que codifica el sexo del individuo.

Para terminar, sólo resaltar la importancia que tiene el hecho de construir la matriz  $A^{-1}$  sobre todo en poblaciones de tamaño pequeño, donde asumir la ausencia de consanguinidad puede afectar de forma importante a la clasificación de los reproductores por su mérito genético.

### Resultados y Discusión

La endogamia media obtenida fue de  $4,03 \times 10^{-4} \pm 7,72 \times 10^{-5}$ , siendo 0,25 y 0,0625 las consanguinidades máxima y mínima encontradas, respectivamente.

A primera vista, el nivel de consanguinidad medio es despreciable, lo cual, desde un punto de vista general, es positivo. No obstante, mediante un análisis más detallado de los datos se

puede llegar a la conclusión de que la información de pedigrí presente parece muy escasa como para considerar estos resultados como definitivos.

En primer lugar, el número de individuos cuyos padres son desconocidos, y que por tanto se consideran población base, supone el 40% del total, lo cual indica la escasa información de parentesco de que se dispone. Esto provoca que la consanguinidad calculada pueda ser más baja que la real.

Un dato que hace intuir la existencia de una endogamia superior a la hallada en este análisis es que la consanguinidad máxima obtenida (0.25) corresponde a aquella que se produce cuando los abuelos paternos son los mismos que los maternos, lo que supone la máxima endogamia posible cuando sólo tres generaciones son consideradas. Dado que el número medio de generaciones encontradas es de 3,2 y que además sólo el 17,5% de los animales superan este número de generaciones, este dato debería ser cuando menos preocupante.

Por tanto, de haber tenido mayor cantidad de información de pedigrí, los resultados tal vez hubieran sido otros muy diferentes.

Un estudio más profundo de los datos parece ser necesario para obtener conclusiones de mayor peso que las expuestas. Este sólo sería posible si se pudiese disponer de un archivo de pedigrís más completo que el actual.

### Bibliografía

Abe, T. and A. Nishida, 1971. A method of calculating coefficients of inbreeding and rela-

tionship based on pedigree lines. Jap. Poultry Sci. 8: 245-249.

- Alfonso Ponce, P. 1971. Redes genealógicas y computación de coeficientes de parentesco y consanguinidad. *Arch. Zootec.* 20: 257-299.
- Alfonso Ponce, P. 1985. Un método de cálculo de coeficientes de identidad genética sobre genealogías extensas. *An. Vet. (Murcia)* 1: 29-43.
- Boyce, A. J. 1983. Computation of inbreeding and Kinship coefficients on extended pedigrees. *J. Hered.* 74: 400-404.
- Gutiérrez, J.P., J. Cañón y M. Rico. 1988. Cálculo del coeficiente de consanguinidad en poblaciones de animales domésticos. *Actas I Congreso monográfico internacional "Mejora genética del ganado porcino"*, vol. I, c-10 a c-12. Lorca (Murcia).
- Hazel, L. N. and J.L. Lush. 1950. Computing inbreeding and relationship coefficients from punched cards. *J. Hered.* 40: 301-306.
- Henderson, C.R. 1976. A simple method for computing the inverse of a numerator relationship matrix used in the prediction of breeding values. *Biometrics* 32: 69-83.
- Nadot, R. 1971. Les coefficients d'identité. Une méthode de calcul. In *Génétique et Populations. Hommage à Jean Sutter*. Presses Universitaires de France. Paris, p. 111-117.
- Nadot, R. and G. Vaysseix. 1973. Apparentement et identité. *Biometrics* 29: 347-359.
- Quaas, R. L. 1976. Computing the diagonal elements and inverse of a larger numerator relationship matrix. *Biometrics* 32: 949-953.
- Schaap, T. and M.M. Cohen. 1977. A simple non graphic method for pedigree description and analysis. *J. Med. Genet.* 14: 25-29.
- Stevens, A. 1975. An elementary computer algorithm for the calculation of the coefficient of inbreeding. *Inf. Processing Lett.* 3:153-163.
- Wright, S. 1922. Coefficients of inbreeding and relationship. *Am. Nat.* 56: 330-338.