INFORMACIÓN MOLECULAR VS. INFORMACIÓN GENEALÓGICA EN LA GESTIÓN DE POBLACIONES

de Cara, M. A. R. [1], Fernández, J., Toro, M.A. y Villanueva, B.
[1] Departamento de Mejora Genética Animal, Instituto Nacional de Investigación y Tecnologías Agrarias y Alimentarias (INIA), Ctra. de La Coruña, km. 7.5, Madrid 28040. Email: rodriguez.angeles@inia.es

INTRODUCCIÓN

Tanto en el campo de la conservación como en el de la mejora genética, se han propuesto diversos métodos para gestionar una población controlando la pérdida de diversidad genética. En poblaciones no subdivididas, el método aceptado es determinar la contribución de cada posible padre (i.e., el número de descendientes que cada individuo deja a la siguiente generación), minimizando el parentesco global de los padres ponderado por estas contribuciones (Meuwissen, 1997; Grundy et al., 1998; Fernández et al., 2003). Los coeficientes de parentesco se obtienen normalmente de la genealogía, y en dicho caso, se optimiza el parentesco global genealógico. Sin embargo, la información genealógica no está siempre disponible, en cuyo caso se pueden usar marcadores moleculares para calcular el parentesco molecular o estimar el parentesco genealógico (Toro et al., 2009). Así, cuando no se dispone de genealogías, se puede minimizar el parentesco molecular global o el parentesco global genealógico estimado con los marcadores.

Fernández et al. (2005) estudiaron mediante simulaciones la capacidad de la información molecular de reemplazar a la información genealógica, y concluyeron que el uso exclusivo de información molecular era claramente insuficiente. En dicho estudio, los autores se basaron en un número limitado de marcadores microsatélites, del orden de decenas. En la actualidad, gracias a los métodos de secuenciación de última generación, disponemos de miles de marcadores de tipo SNP, lo que hace necesaria una revisión de aquellas investigaciones que concluían que la utilidad de la información molecular era limitada e inferior a la genealógica. En este estudio, reevaluamos vía simulaciones la capacidad de la información genómica de reemplazar a la información genealógica para mantener diversidad genética en programas de conservación, usando un número de SNPs en línea con los datos actualmente disponibles.

MATERIAL Y MÉTODOS

Nuestras simulaciones comprenden dos partes: en la primera generamos una población base de censo efectivo de 100 individuos en equilibrio mutación-deriva, y en la segunda procedemos a la gestión de la población para mantener la máxima diversidad genética, en poblaciones de 20 ó 100 individuos.

Para generar la población base, simulamos durante 5000 generaciones una población de 100 individuos (50 machos y 50 hembras) que se aparean al azar con reemplazamiento. El genoma está compuesto por 1 ó 20 cromosomas, de 1 M cada uno. Cada cromosoma tiene 11000 loci bialélicos equidistantes de dos tipos: 1000 marcadores que usamos para la gestión, intercalados entre los 10000 loci no marcadores que usamos para medir diversidad. En cada generación ocurren mutaciones con tasa por locus y por generación de 0.0025, sigue con recombinación y apareamiento al azar. Inicialmente, todos los loci están fijados, y tras 5000 generaciones la población en equilibrio mutación-deriva presenta un 93% de los loci polimórficos.

La población se gestiona entonces durante 10 generaciones, optimizando las contribuciones para minimizar el parentesco global ponderado por dichas contribuciones (Fernández et al., 2003). Este método maximiza la diversidad genética (heterocigosidad esperada) de la población, el censo efectivo y la heterocigosidad esperada. La solución óptima se obtiene mediante un algoritmo de “simulated annealing”.

Podemos clasificar las estrategias de gestión en función de los individuos usados para la optimización, así como si usamos información molecular o genealógica.

Respecto al grupo de individuos utilizados para la optimización, podemos distinguir dos casos: optimizar las contribuciones de los potenciales padres, o bien, actuar sobre los descendientes. En el primer caso, actuando sobre los padres, calculamos parentescos esperados, mientras que en el segundo, cuando actuamos sobre los descendientes, usamos los parentescos observados.

Por otro lado, podemos tener información genealógica, a partir de la cual inferimos el parentesco genealógico, o bien, información molecular. Cuando sólo disponemos de información molecular, podemos usarla para calcular parentescos moleculares o para estimar parentesco genealógico usando los estimadores de Lynch y Ritland (1999) y el estimador WEDS (Oliehoek et al., 2006).

Para simplificar, los estimadores de parentesco sólo los usamos cuando actuamos sobre los padres, por lo que tenemos 6 estrategias de gestión: 1), actuando sobre los padres con información genealógica; 2) actuando sobre los padres con información molecular y optimización del parentesco molecular; 3) actuando sobre los padres con información molecular y optimización del parentesco genealógico estimado mediante el estimador de Lynch y Ritland (1999); 4) actuando sobre los padres con información molecular y optimización del parentesco genealógico estimado mediante el estimador WEDS de Oliehoek et al. (2006); 5) actuando sobre los hijos con información genealógica; y 6) actuando sobre los hijos con información molecular y optimización del parentesco molecular.

Estas diferentes estrategias las comparamos en términos de diversidad génica (GD), diversidad alélica (AD) y heterocigosidad observada (OH), medidas en los loci no marcadores.

RESULTADOS Y DISCUSIÓN

Generamos 50 poblaciones base sobre las que luego hacemos la gestión y promediamos. Para cada escenario arriba mencionado, usamos las mismas réplicas, de forma que en t = 0 de la gestión las medias son iguales para cada escenario.

En la Figura 1, en la columna izquierda, podemos ver como la gestión, medida a través de la diversidad génica, es mejor utilizando el parentesco molecular, que es casi equivalente a utilizar el estimador WEDS. Los resultados son consistentemente peores cuando usamos información genealógica. En la misma figura, en la columna derecha, vemos como actuar sobre los hijos mantiene mayor diversidad que actuar sobre los padres. De igual manera, se mantiene mayor diversidad usando parentescos moleculares que usando información genealógica. Los resultados son equivalentes para poblaciones de 20 individuos, y para genomas de 1 cromosoma (resultados no mostrados).

Así, nuestra revisión de los resultados de Fernández et al. (2003) ha resultado en un claro cambio sobre sus conclusiones, y creemos que estamos ahora en una posición en la que si tenemos un gran número de marcadores, es preferible usar la información molecular. Por otra parte, no creemos necesario usar los estimadores de parentesco, ya que éstos tan sólo nos acercan a la información genealógica, lo que resulta en una peor gestión. Actuar sobre los hijos es preferible a actuar sobre los padres, ya que tenemos información observada, no esperada. Con estos resultados, cabe preguntarse si los estimadores de parentesco son necesarios y si no sería preferible ignorarlos y usar información molecular sin corregir, siempre que tengamos un gran número de marcadores distribuidos a lo largo del genoma.

REFERENCIAS BIBLIOGRÁFICAS

● Grundy, B., Villanueva, B. & Wooliams, J.A. 1998. Gen. Res. 72: 159-168. ● Fernández, J. & Toro, M.A. 1999. J. Anim. Breed. Genet. 116: 447-466. ● Fernández, J., Toro, M.A. & Caballero, A. 2003. Genetics 165: 885-894. ● Fernández, J., Villanueva, B., Pong-Wong, R. & Toro, M.A. 2005. Genetics 170: 1313-1321. ● Lynch, M. & Ritland, K. 1999. Genetics 152: 1753-1766. ● Meuwissen, T.H.E. 1997. J. Anim. Sci. 75: 934-940. ● Oliehoek, P.A., Windig, J.J, van Arendonk, J.A.M. & Bijma, P. 2006. Genetics 173: 483-496. ● Toro, M.A., Fernández, J. & A. Caballero. 2009. Livest. Sci. 120: 174-195.

Agradecimientos: Este trabajo ha sido financiado por el proyecto CGL2009-13278-C02-02.

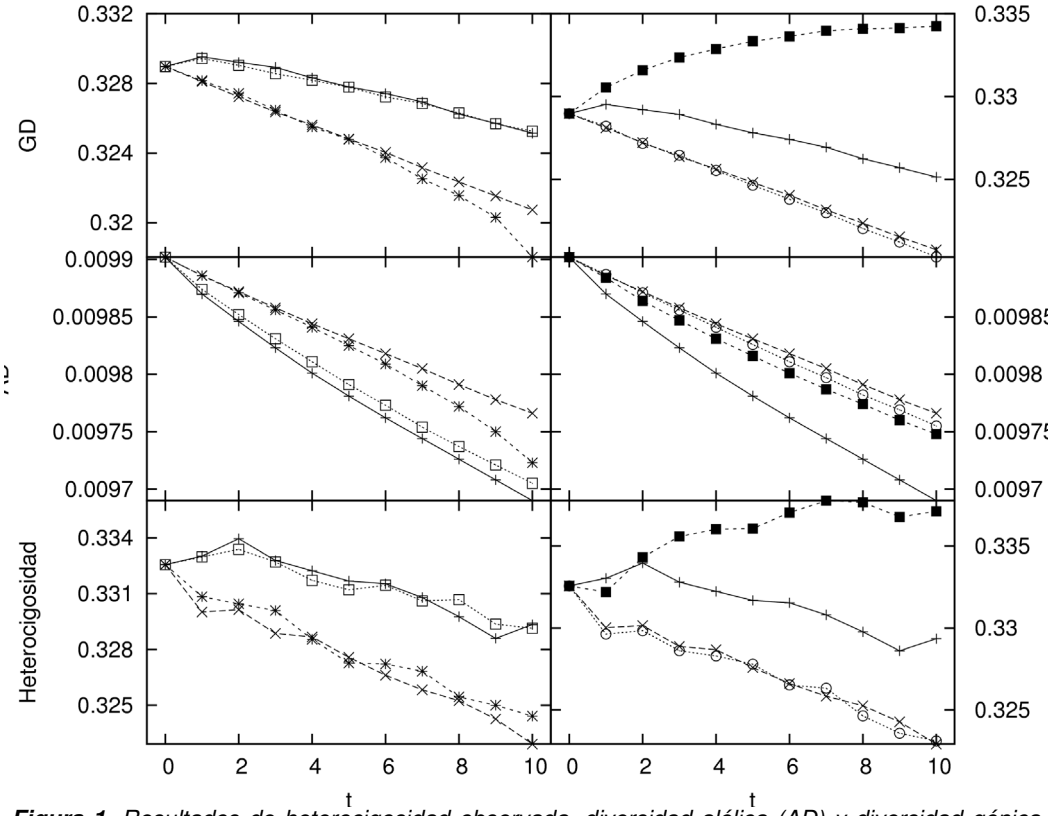


Figura 1. *Resultados de heterocigosidad observada, diversidad alélica (AD) y diversidad génica (GD) durante la gestión para una población de 100 individuos de genoma de 20 cromosomas. En la columna izquierda, comparación sobre la gestión en los padres, optimizando el parentesco genealógico (x), optimizando el parentesco molecular (+), optimizando parentesco genealógico estimado LR usando información molecular (*) y optimizando el parentesco genealógico estimado WEDS usando información molecular (□). En la columna de la derecha, comparación de la gestión resultante de actuar sobre padres optimizando el parentesco genealógico (x), optimizando el parentesco molecular (+) o bien actuando sobre los hijos optimizando el parentesco genealógico (○), optimizando el parentesco molecular (■).*

GENOME-WIDE INFORMATION VS. GENEALOGICAL INFORMATION IN POPULATION MANAGEMENT

ABSTRACT: We study here the consequences in the management of populations of using genealogical information versus genome-wide data for populations of different sizes and different genome sizes. The use of genealogical coancestry from the first type of data is compared with the use of molecular coancestry or estimated genealogical coancestry from the latter data. Similarly, we compare acting on the potential parents versus acting on the actual offspring. Using molecular coancestry performed consistently better than genealogical coancestry, while acting on the offspring maintained a larger gene diversity than acting on the parents.

Keywords: Pedigree, molecular coancestry, optimal contributions.