

ESTIMAS GENÓMICAS DE PARENTESCO Y CONSANGUINIDAD EN UNA ANTIGUA PIARA CONSANGUÍNEA DE CERDO IBÉRICO

Saura¹, M., Fernández¹, A., Rodríguez¹, M.C., Toro², M.A., Barragán¹, C., Fernández¹, A.I. y Villanueva¹, B.

¹Departamento de Mejora Genética Animal, INIA, Carretera de la Coruña Km 7.5, 28040 Madrid, Spain, email: saura.maria@inia.es

²Departamento de Producción Animal, ETS Ingenieros Agrónomos, Ciudad Universitaria 28040 Madrid, Spain

INTRODUCCIÓN

El mantenimiento de la diversidad genética y el incremento de la consanguinidad son requerimientos cruciales en los programas de conservación. El marco de actuación general para controlar la pérdida de variabilidad y evitar la depresión consanguínea se basa en controlar las tasas de consanguinidad y parentesco, o lo que es lo mismo, manejar el censo efectivo poblacional (N_e).

Dentro de las razas de porcino, el cerdo ibérico representa la más emblemática del Mediterráneo y una de las más importantes del mundo desde un punto de vista económico. A lo largo de la década de los sesenta, las poblaciones de cerdo ibérico sufrieron un acusado declive, aunque la expansión de los mercados ha conseguido revertir esta tendencia. La estirpe Guadyerbas, una de las más antiguas y actualmente la única representante del tipo negro lampiño, se encuentra en serio peligro de extinción. La estirpe se ha conservado en una piara experimental que se ha mantenido aislada desde 1944, con un exhaustivo control de la genealogía, constituyendo un recurso genético de incalculable valor.

En general, la evolución de la diversidad genética se ha medido utilizando la información del pedigrí, o utilizando marcadores microsatélites en los casos en que éste no estuviera disponible. Este tipo de información molecular es limitada, y su aplicación para mantener la variabilidad genética ha sido cuestionada en diversos estudios. El reciente desarrollo de nuevas tecnologías de secuenciación y genotipado masivo, permite identificar miles de polimorfismos nucleotídicos sencillos (SNPs) teniendo en cuenta información relevante que los microsatélites ignoran.

En este trabajo se ha utilizado el chip porcino PorcineSNP60 BeadChip (Illumina) para obtener estimas genómicas de parentesco, consanguinidad y censo efectivo para la estirpe Guadyerbas. Estas estimas moleculares se compararon con las equivalentes obtenidas a partir de la genealogía y a partir de marcadores microsatélites. Por último, se ha evaluado la eficiencia de la información genómica aplicada a programas de manejo para maximizar la diversidad genética como alternativa al uso de la genealogía.

MATERIAL Y MÉTODOS

Se genotiparon 227 animales Guadyerbas nacidos entre 1992 y 2011 con el chip de Illumina que, de acuerdo con la última versión del mapa porcino, comprende 62163 pruebas distribuidas a lo largo de los 18 autosomas y 2 cromosomas sexuales. El filtrado de los datos se realizó con el programa GenomeStudio (Illumina). Tras aplicar los controles de calidad, el set de datos incluyó 219 muestras con genotipos fiables para 35519 SNPs. De estos individuos, 30 disponían también de genotipos para 57 microsatélites (Toro y col. 2002).

El coeficiente de parentesco molecular entre dos individuos i y j (f_M) se calculó como:

$$f_{Mij} = (1/L) \sum_{l=1}^L \left[\left(\sum_{k=1}^2 \sum_{m=1}^2 I_{lk(i)m(j)} \right) / 4 \right]$$

donde L es el número de marcadores y $I_{lk(i)m(j)}$ es la identidad del alelo k -ésimo del individuo i con el alelo m -ésimo del animal j en el locus l , que toma valor de 1 si los alelos son idénticos y de 0 si no lo son. El coeficiente de consanguinidad molecular para el individuo i se calculó como $F_{Mi} = 2f_{Mii} - 1$ (es decir, como la proporción de genotipos homocigotos). Los coeficientes de parentesco (f_G) y consanguinidad (F_G) genealógicos se calcularon utilizando el algoritmo de Meuwissen y Luo (1992). Las tasas de parentesco (Δf_y) y consanguinidad (ΔF_y) (molecular y genealógico) anuales se computaron regresando el logaritmo neperiano de $(1 - f)$ o $(1 - F)$ sobre el año de nacimiento. Las pendientes de las rectas de regresión equivalen a $-\Delta f_y$ y $-\Delta F_y$, que multiplicadas por $-L$ (siendo L el intervalo generacional igual a 3

años) dan las correspondientes tasas por generación (Δf y ΔF). El censo efectivo se calculó como $N_{e_f} = 1/2 \Delta f$ o $N_{e_F} = 1/2 \Delta F$, utilizando tanto datos moleculares como genealógicos. Desde 1982, la estrategia adoptada para mantener la diversidad genética en el programa de conservación de Guadyerbas ha sido realizar apareamientos de mínimo parentesco genealógico. En este contexto, se evaluó la eficiencia de los marcadores de alta densidad en el manejo de poblaciones en conservación como alternativa al uso de la información genealógica. Se utilizaron datos empíricos de una generación (12 machos y 47 hembras) y se evaluaron dos criterios de optimización de apareamientos: minimización del parentesco molecular (MC_M) y minimización del parentesco genealógico (MC_G). Las restricciones impuestas implicaron un diseño balanceado en el que todos los individuos contribúan a la siguiente generación y que se ajusta a lo que se lleva a cabo en la práctica. Con objeto de clarificar cuál de los dos criterios es más eficiente, se midió el parentesco molecular y genealógico promedios que se alcanzaba con cada uno de ellos.

RESULTADOS Y DISCUSIÓN

Tras la eliminación de los SNPs monomórficos para la raza de cerdo ibérico, cerca del 45% de los SNPs fueron monomórficos en Guadyerbas, en consonancia con lo que se espera de una población aislada de censo tan reducido. El promedio de los coeficientes de parentesco y consanguinidad moleculares ($0,82 \pm 0,015$ y $0,81 \pm 0,016$, respectivamente) fueron casi el doble que los coeficientes genealógicos ($0,42 \pm 0,047$ y $0,39 \pm 0,037$), aunque las tasas de parentesco molecular y genealógico fueron muy similares, al igual que ocurrió con las tasas de consanguinidad. No obstante, Δf y ΔF fueron diferentes y por tanto difirieron también las estimas de N_e calculadas a partir de Δf (10 individuos) y de ΔF (16-18 individuos). Esto se debe al tipo de manejo al que se ha sometido a la población, en el que se fuerzan apareamientos de mínimo parentesco, desviándose del apareamiento aleatorio.

La correlación entre f_M y f_G fue muy elevada (0,90), no así tanto entre F_M y F_G (0,68). El hecho de que la correlación entre parentescos fuese más alta podría tener relación con una mayor dispersión de los datos en el caso del parentesco genealógico (rango de $f_G = 0,41$; rango de $F_G = 0,16$). Estudios previos con marcadores microsatélites enfatizan la importancia de la varianza de los coeficientes genealógicos como el factor determinante para obtener una alta correlación (Slate y col. 2004). Sin embargo, en este estudio la varianza de los coeficientes genealógicos fue muy baja (0,0022 para f_G y 0,0014 para F_G), sugiriendo que en aquellos casos en que se utilizan miles de marcadores, la varianza de los coeficientes genealógicos no es el factor determinante de la magnitud de la correlación. De hecho, en el grupo de individuos con información de genealogía, SNPs y microsatélites, la correlación entre el parentesco molecular obtenido con SNPs y el parentesco genealógico resultó ser significativamente más elevada (0,86) que la equivalente utilizando microsatélites para estimar el parentesco molecular (0,67). También se observó que la predicción del parentesco genealógico a partir del molecular fue mucho más precisa utilizando SNPs (0,95) que utilizando microsatélites (0,82) (Figura 1).

El desarrollo de nuevas tecnologías de genotipado masivo ha transformado el paradigma de que la genealogía es el parámetro más apropiado para medir la diversidad genética. Contrariamente a lo sugerido en estudios basados en microsatélites (Fernández y col. 2005), estudios recientes han resaltado los beneficios de los chips densos de SNP para mantener variabilidad genética (de Cara y col. 2011). A pesar de que en este estudio los criterios de minimizar el parentesco molecular (MC_M) o minimizar el parentesco genealógico (MC_G) mostraron resultados similares en términos de parentesco global (tanto molecular como genealógico), las parejas escogidas por cada uno de los criterios sólo coincidieron en el 30% de los casos. Esto se debe a que, aunque el parentesco molecular tiene la habilidad de discriminar entre individuos con el mismo grado de parentesco genealógico, el manejo de la población ha dado lugar a una gran reducción en la varianza del parentesco, por lo que la ventaja de utilizar el parentesco molecular no se ha trasladado en la práctica.

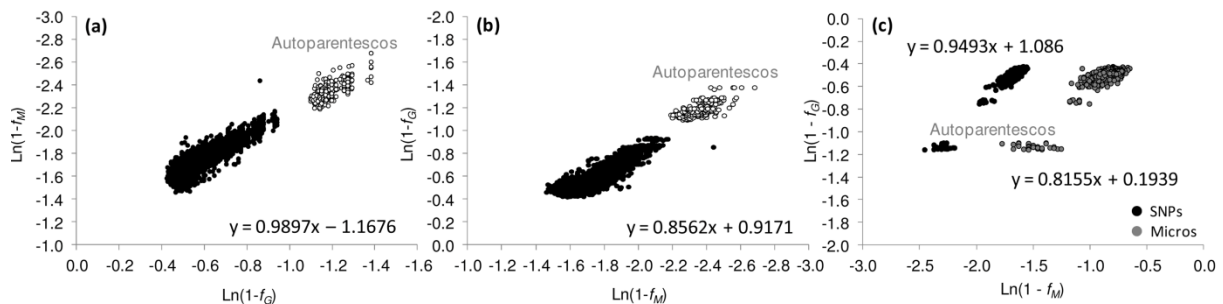
La población analizada en este estudio dispone de información completa y precisa del pedigrí, una situación muy inusual. Nuestros resultados sugieren que en ausencia de esta información, la utilidad de la información molecular a partir de chips de SNPs para mantener diversidad genética es indiscutible.

REFERENCIAS BIBLIOGRÁFICAS

- Meuwissen T., & Luo Z. 1992. *Genet Sel Evol* 24:305-313.
- Toro MA., Barragán C., Óvilo C., Rodríguez J., Rodríguez MC., & Silió L. 2002. *Cons Gen* 3:309–320.
- Fernández J., Villanueva B., Pong-Wong R., & Toro MA. 2005. *Genetics* 170:1313-1321.
- de Cara MA., Fernandez J., Toro MA., & Villanueva B. 2011. *J Anim Breed Genet* 128:456–464.
- Slate J., David P., Dodds KG., Veenvliet BA., Glass BC., Broad TE., & McEwan JC. 2004. *Heredity* 93:255–265.

Agradecimientos: Los autores agradecen a L. Silió su implicación y sugerencias en el trabajo y a R. Benítez y F. García su aporte técnico. Trabajo financiado por el proyecto INIA RZ2010-00009-00-00. M. Saura disfruta de un contrato Juan de la Cierva del MINECO.

Figura 1. Predicción del parentesco molecular (f_M ; estimado con datos de SNPs) a partir del parentesco genealógico (f_G) (a), y viceversa (b), para los 219 Guadyerbos analizados. Predicción del parentesco genealógico a partir de los dos tipos de parentesco molecular (SNPs y microsatélites), para los 30 individuos con información adicional de marcadores microsatélites (c).



GENOME-WIDE ESTIMATES OF COANCESTRY AND INBREEDING IN A CLOSED HERD OF ANCIENT IBERIAN PIGS

ABSTRACT: Maintaining genetic variation and controlling the increase in inbreeding are crucial requirements in animal conservation programs. As a general rule, the most widely accepted strategy to implement in conservation programs is to maximize the effective population size by minimizing the global coancestry obtained from the pedigree. When genealogical information is absent, microsatellites have been traditionally the markers of choice to characterize genetic variation, with questionable results. The development of high-throughput genotyping techniques states the necessity of reviewing the paradigm that genealogy is the best parameter to measure genetic diversity. In this study the Illumina PorcineSNP60 BeadChip was used to obtain genome-wide estimates of rates of coancestry and inbreeding and effective population size for an ancient strain of Iberian pigs which is now in serious danger of extinction. These estimates were compared with those obtained from microsatellite and from pedigree data. Estimates of coancestry and inbreeding computed from the SNP chip were strongly correlated with genealogical estimates and these correlations were substantially higher than those between microsatellite and genealogical coefficients. Molecular coancestry computed from SNP information was a better predictor of genealogical coancestry than coancestry computed from microsatellites. Our results indicate that genome-wide information represents a useful alternative to genealogical information for measuring and maintaining genetic diversity.

Keywords: coancestry, effective population size, genome-wide, Iberian pig.