ITEA (2000), Vol. **96A N.° 3,** 282-289

MEDIDAS DE DIVERSIDAD GENÉTICA EN POBLACIONES DE CABALLOS CELTAS ESPAÑOLES

D. García, M.L. Checa, P. García-Atance, S. Dunner, J. Cañón

Laboratorio de Genética, Dpto. de Producción Animal, Universidad Complutense de Madrid, España

RESUMEN

Con el objetivo de estudiar la diversidad genética existente en un conjunto de razas equinas españolas y el aporte de cada raza a la diversidad global de cara a la aplicación de políticas de gestión de recursos genéticos, se comparan en este artículo dos procedimientos alternativos. El propuesto por Weitzmann se basa en distancias genéticas y su medida de diversidad posee propiedades teóricas que hacen de ella una medida robusta frente a nuevas aportaciones poblacionales y variaciones en las distancias. Produce además un árbol jerárquico de fácil interpretación en lo que a aportaciones genéticas de cada raza se refiere. El análisis de correspondencias es un método estadístico clásico que aplicado en este contexto permite que un concepto análogo al de la tasa de inercia se equipare al de la diversidad aportada por cada raza, además de generar una representación gráfica de las relaciones entre las razas.

Palabras clave: Diversidad genética, Procedimiento de Weitzman, Análisis de Correspondencia.

SUMMARY

In order to study the genetic diversity in a set of spanish horse breeds and the contribution of each breed to the whole diversity, with a posterior use in genetic resources management policies in mind, two different approaches are compared in this paper. Weitzmann's is based upon genetic distances, and the theoretical properties associated to its measure of diversity make it a robust measure with respect to breed additions and variations in the distances. Furthermore, it produces an easily interpretable hierarchical tree for individual breeds genetic contributions studies. The analysis of correspondences is a clasical statistical method which, applied to this context, allows an analogous concept to that of inercy to be translated into the diversity added by each breed. Appart from that, it generates a plot of the relationship among the breeds.

Key words: Genetic diversity, Weitzman's approach, Analysis of correspondence.

Introducción

El término "diversidad genética" aparece, como es lógico, con mucha frecuencia en el contexto de conservación de recursos genéticos. Generalmente aceptamos que la tasa de evolución en poblaciones naturales está limitada por la cantidad de variabilidad genética (Teorema Fundamental de la Selección Natural de Fisher, 1930). No resulta extraño, por lo tanto, que muchos de los esfuerzos en los programas de gestión de recursos genéticos se hayan dedicado al mantenimiento de la diversidad genética, aunque no parece claro que, en el contexto de las especies de animales domésticos, este factor tenga un impacto similar sobre la probabilidad de extinción de una determinada población.

Dependiendo de las características de las poblaciones en estudio y de los objetivos se han propuesto, para definir el concepto de diversidad, diferentes parámetros que combinan la distribución de frecuencias génicas, genotípicas o gaméticas. Entre los parámetros propuestos, los que más frecuentemente aparecen son: el número de loci polimórficos, el número de alelos por locus, la heterocigosis (LEWONTIN, 1974) el Contenido de Información Polimórfica (BOTSTEIN *et al.*, 1980), y con mucha frecuencia, la manera de identificar la excepcionalidad de una población con fines de asignación prioritaria de recursos para su conservación ha sido la utilización de las medidas de distancia genética entre parejas de poblaciones. La capacidad de resolución de estas medidas de distancia no parecen muy evidentes cuando, entre otros problemas, existe la posibilidad de obtener muy diferentes topologías para un mismo conjunto de distancias genéticas y, viceversa, árboles de similar parsimonia originan diferentes estructuras genéticas.

Un aspecto central en programas de conservación genética es la definición de una medida de diversidad y su representación gráfica. El objeto de esta nota es aplicar a un conjunto de razas de la especie equina, dos procedimientos alternativos: a) un procedimiento desarrollado por WEITZMAN (1992, 1993) precisamente para medir la diversidad con fines de su utilización en la gestión de recursos genéticos, b) un análisis de correspondencia que permite estudiar una variable cualitativa (raza) en función de otras variables cualitativas (alelos) y en el que el término clásico de inercia puede ser asimilado al de diversidad.

Métodos de cálculo

Procedimiento de Weitzman

WEITZMAN define una medida de diversidad a partir de, por ejemplo, distancias genéticas, con una serie de propiedades empíricas. Utilizando una notación similar a la de WEITZMAN (1992), supongamos que S representa el conjunto de poblaciones (especies, razas, etc) y que la medida de la distancia genética entre una pareja (i, j) de esas poblaciones, d(i,j) satisface las condiciones: a) $d(i,j) \geq 0$, b) $d(i,i) = 0$, c) $d(i,j) = d(j,i)$. La medida de diversidad V, del conjunto S de poblaciones, V(S) es la solución de:

$$V(S) = \max_{i \in S} \{ V(S \backslash i) + d(i, S \backslash i) \} \quad [1]$$

donde V(S\i) denota la diversidad del conjunto de poblaciones S tras la eliminación de i.

WEITZMAN (1992) considera razonable imponer a la función de diversidad las siguientes condiciones:

• *Monotonicidad de las poblaciones:* cuando se añade un elemento j al conjunto de poblaciones Q, $Q \subseteq S$ la diversidad Q se

incrementa al menos en la cantidad d(j,Q), donde d(j,Q) representa la distancia del elemento j a Q.

$$V(Q \cup j) \geq V(Q) + d(j,Q)$$

• *La propiedad del enlace*: existe al menos una población j que satisface

$$V(S) = V(S\backslash j) + d(j,S\backslash j)$$

Es decir, existe alguna población cuya eliminación reducirá la diversidad en la misma magnitud que la distancia de esa población a la más próxima. Esta propiedad garantiza una conexión natural entre la diversidad medida y las distancias genéticas en las que esa diversidad se basa.

• *Propiedad de los idénticos*: añadir una población idéntica a otra existente en S no proporciona ningún incremento de la diversidad.

• *Continuidad de las distancias:* pequeños cambios en las distancias genéticas dan lugar a pequeños cambios en la diversidad.

• *Monotonicidad de las distancias*: si se incrementan todos los pares de distancias también se producirá un incremento de la diversidad.

En lugar de utilizar [1] para calcular V(S), lo que supone n! cálculos, Weitzman prueba que resolver [1] es equivalente a resolver:

$$V(S) = d[g(S),h(S)] + \max\{V[S\backslash g(S)], V[S\backslash h(S)]\}$$

En la que el elemento g(S) es lo que en el cuadro 1 llamamos el elemento enlace y satisface que $g(S) = \max [V(S\backslash g), V(S\backslash h)]$ y h(S) se denomina representante en esa misma tabla.

Representación gráfica

La información contenida en las distancias genéticas puede representarse gráficamente en forma de un árbol jerárquico cuya topología es única si identificamos las poblaciones que actúan de enlace. Este árbol así construido permite una fácil visualización del efecto a que daría lugar en términos de diversidad la pérdida de una población o un conjunto de ellas.

Debemos tener en cuenta que este árbol es una representación de la diversidad observada a través de las distancias genéticas y, por lo tanto, matrices de distancias genéticas escasamente correlacionadas darán lugar a topologías muy diferentes. Además, también debemos ser conscientes de que, salvo raras excepciones, la topología observada no puede ser interpretada como el reflejo del verdadero proceso evolutivo sufrido por las poblaciones objeto de estudio.

Análisis de correspondencia

El análisis se efectúa a partir de una tabla de contingencia por raza en el que las filas están constituidas por la variable a explicar (la raza) y las columnas el conjunto de variables explicativas (los alelos). El análisis se trata de la modalidad de correspondencia múltiple en el que los diferentes loci se consideran no ligados. Los resultados nos permiten obtener la importancia relativa (inercia) tanto de las razas como de los alelos, una representación gráfica mediante un sistema de puntos situados en un espacio euclidiano y la verificación de su grado de robustez.

Material utilizado

Poblaciones de caballos celtas

Las poblaciones, marcadores genéticos y las medidas de distancia genética que han

sido utilizados se encuentran descritos en CAÑON *et al.* (2000).

Resultados y discusión

La matriz de distancias genéticas puede ser transformada en otra de distancias con propiedades ultramétricas con la ventaja de poder ser utilizadas por programas estándares (ej. NTSYS, RHOLF, 1988) para la construcción de la única topología posible. Uno de los inmediatos aspectos de interés de esta representación gráfica es la visualización de la reducción que se produce en diversidad cuando alguna población o un conjunto de ellas desaparece.

Si observamos la figura 1, en la que se representa la situación de los caballos españoles tipo pony podemos comprobar que, desde la perspectiva de pérdida de diversidad, la situación de las dos poblaciones de las islas Baleares es la más crítica ya que, en el supuesto de que se perdiera la raza mallorquina, la diversidad se vería reducida en un 22%, mayor porcentaje que la desaparición de cualquier otra de las razas incluidas en el análisis, incluso superior al supuesto en el que las dos razas más próximas entre si, el pottoka y el caballo gallego, desaparecieran. Pero la gravedad sería mucho mayor si desaparecieran las dos razas de las islas Baleares ya que ese cluster que desaparece representa el 50% de la diversidad global, es decir, de la diversidad del conjunto de las razas incluidas en este estudio. La situación es completamente diferente si nos fijamos, por ejemplo, en las poblaciones de pottoka y caballo gallego ya que su posición indica que la desaparición de cualquiera de ellas representa una escasa pérdida de diversidad, 9 y 7% respectivamente; incluso la desaparición de ambas razas no resulta tan crítico como la desaparición de una de las de las islas baleares, un 18% frente al 50%. La pérdida de variabilidad genética cuando se elimina un subconjunto de razas no necesariamente tiene que ser igual a la suma de las pérdidas

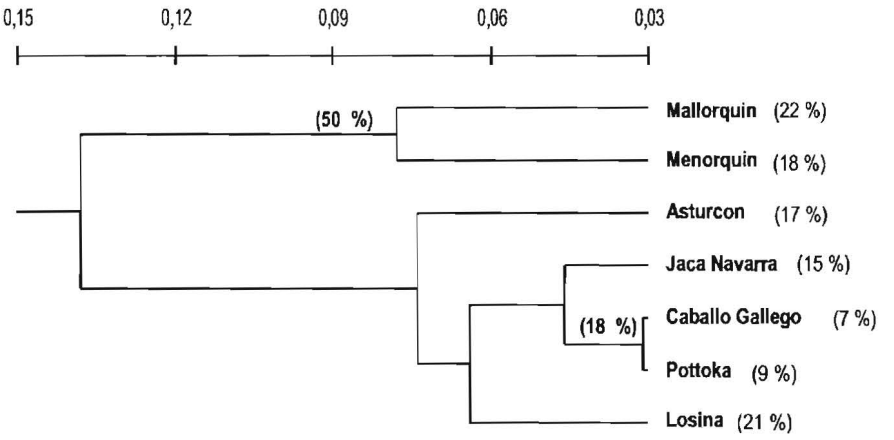


Figura 1. Dendrograma representando la posición de diferentes razas de caballos de origen celta. Las cifras entre paréntesis al lado de cada raza representa el porcentaje de diversidad que se pierde por la desaparición de esa población. Las cifras en los nodos representan el porcentaje de pérdida de diversidad por la eliminación de ese cluster

de cada una de las razas individualmente consideradas que constituyen ese cluster. La aditividad se cumple cuando las razas son independientes.

Otra aplicación de interés resulta de la posibilidad de conocer el rendimiento, en términos de diversidad marginal, que proporciona la inversión en mejorar la probabilidad de supervivencia de una población. En nuestro caso lo que hemos hecho es calcular el incremento marginal en diversidad como consecuencia del incremento de la heterocigosis. Hablar de probabilidad de extinción en poblaciones no naturales como es el caso de las razas de animales domésticos no resulta fácil, pero podríamos pensar en que las acciones de conservación de recursos genéticos animales tengan como uno de sus objetivos tratar de incrementar el nivel de heterocigosis o, más propiamente, reducir el incremento en el nivel de homocigosis, suponiendo que una situación de homocigosis es sinónimo de extinción desde el punto de vista genético. Podríamos, por lo tanto, analizar el rendimiento de la inversión en las poblaciones en función del nivel de homocigosis en la que se encuentran. Para ello hay que utilizar una función que transforme homocigosis en probabilidad de extinción. En ausencia de otra mejor se ha elegido la transformación directa de homocigosis a probabilidad de extinción, es decir, P[extinción] (raza k) = Homocigosis (raza k).

Otro de los aspectos de interés señalado por Weitzman es la posibilidad de calcular la elasticidad de la diversidad genética en las diferentes poblaciones, es decir, en qué medida cambios porcentuales en los niveles de homocigosis provocan cambios más o menos elevados en la diversidad. Una elasticidad elevada indicaría que cambios relativamente pequeños en los niveles de heterocigosis proporcionarían una elevada respuesta en la diversidad del conjunto de las poblaciones.

Recientemente se han propuesto otros procedimientos para medir la diversidad como el Análisis de Correspondencia (LEBART, MORINEAU & WARWICK, 1984) en el que la medida de la 'inercia' de una variable (raza) puede ser equiparable a la diversidad que aporta al conjunto de razas analizadas (LALÖE *et al.*, 1999). Será interesante comprobar en que medida ambos pro-

Cuadro I
Resultados sobre la diversidad acumulada y el orden de aparición de las razas en el dendrograma

Representante	Enlace	Distancia	Variabilidad acumulada	Orden
Pot	Gal	.0310	.0310	Mal
JN	Pot	.0460	.0770	Men
Los	JN	.0640	.1410	Ast
Los	Ast	.0740	.2150	JN
Mal	Men	.0780	.2930	Gal
Mal	Los	.1380	.4310	Pot
				Los

Pot: Pottoka; JN: Jaca Navarra; Los: Losino; Mal: Mallorquina; Men: Menorquina; Gal: Caballo Gallego; Ast: Asturcón.

cedimientos proporcionan resultados parecidos.

El análisis de correspondencia que llevamos a cabo daba como resultado que la raza de caballos asturcones, con un 22%, era la que mayor inercia aportaba, y la población de pottoka la que menos (8%). El primer eje (explica un 31% de la inercia total) claramente diferenciaba al asturcón del resto de las razas. El eje 2 (explica un 24% de la inercia total) a la jaca navarra y el tercer eje (explica un 19% de la inercia total) discriminaba al losino (figura 2). El hecho de que un único marcador (HMS6) contribuyera a las dos primeras componentes con un porcentaje relativamente elevado, el alelo HMS6-171 un 13% al eje 1 y un 11% al eje 2, y los alelos HMS6-161 y HMS6-090 con un 9 y 3% al segundo eje, nos llevó a realizar el análisis después de eliminar este marcador. Los resultados de las contribuciones a la inercia reflejan la misma situación para todas las razas y los ejes discriminan de la misma manera que en el caso anterior.

En el cuadro 2 se presentan las inercias aportadas por cada una de las razas, así como los valores de la diversidad obtenidos por el procedimiento de WEITZMAN. La correlación entre ambos resultados es aproximadamente de 0,6.

La utilización de la medida de distancia genética como argumento base de la medida de diversidad mediante el procedimiento propuesto por WEITZMAN lleva a preguntarnos por la dependencia de los resultados en función del tipo de medida que pueda ser calculada. A pesar de que THAON D'ARNALDI, FOULLEY and OLLIVER (1998) con razas bovinas encuentran valores muy parecidos cuando utilizan dos distancias genéticas (Nei y Cavalli-Sforza) resulta evidente que cambios relativos importantes entre las dis-

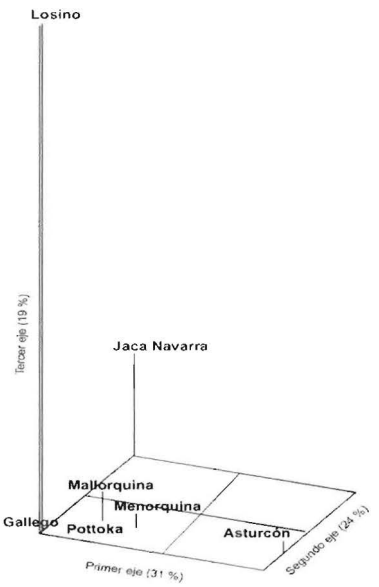


Figura 2. Proyección de las razas sobre los tres primeros ejes del análisis de correspondencia

Cuadro 2
Valores de inercias (análisis de correspondencia) y diversidad (procedimiento de Weitzman) aportadas por cada una de las razas analizadas

	Asturcón	Jaca Nava.	Losino	Mallorquín	Menorquín	Gallego	Pottoka
Inercia	22	17	16	13	15	10	8
Inercia[a]	25	17	17	12	12	10	6
Weitzman	17	15	21	22	17	7	9

([a]) Resultados después de eliminar el marcador HMS6.

tancias genéticas de las parejas de poblaciones pueden proporcionar ordenaciones muy diferentes. En nuestro caso sólo las distancias de FST (WRIGHT, 1965) y de REYNOLDS *et al.* (1983) dieron lugar exactamente al mismo orden, aunque no exactamente a la misma topología, no existiendo correlación distinta de 0 cuando se comparaban los resultados obtenidos con estas distancias y los obtenidos al emplear distancias calculadas bajo el modelo de mutación paso a paso como la distancia cuadrática media (GOLDSTEIN *et al.*, 1995a) o la distancia delta-mu (GOLDSTEIN *et al.*, 1995b) que han sido propuestas para tener en cuenta la elevada tasa de mutación que presentan los microsatélites.

Algunas de las críticas que se hacen a este procedimiento han sido señaladas por THAOM D'ARNOLDI, FOULLEY and OLLIVIER (1998); entre ellas la de no tener en cuenta la variabilidad dentro de poblaciones, algo que sí ocurre con el análisis de correspondencia. Habrá que llevar a cabo estudios adicionales para ver cual es el grado de robustez de los resultados que proporcionan estos métodos cuando se modifican determinadas variables como el tamaño de población, el número de marcadores y su potencia de discriminación, medida de distancia genética utilizada, etc. Otro problema que obviamente no puede ser objetivo de estos métodos es el que concierne a las necesidades de conservación basadas en argumentos o valores no tangibles, como los culturales.

Agradecimientos

El material utilizado en esta presentación es consecuencia de un proyecto financiado por la CICYT (AGF95-0643) y CajaAstur.

Bibliografía

BOSTSTEIN D., WHITE R.L., SKOLNIK M., DAVIS R.W., 1980. Construction of a genetic linkage map in man using restriction frgament length polymorphism. American Journal of Human Genetics, 32, 314-331.

CAÑON J, CHECA ML., CARLEOS C., VEGA-PLA JL., VALLEJO M., DUNNER S., 2000. The genetic structure of Spanish Celtic horse breeds inferred from microsatellite data. Animal Genetics, 31, 39-48.

FISHER R.A., 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

GOLDSTEIN D.B., RUIZ LINARES A., CAVALLI-SFORZA L.L., FELDMAN M.W., 1995a. An evaluation of genetic distances for use with microsatellite loci. Genetics, 139: 463-471.

GOLDSTEIN D.B., RUIZ LINARES A., CAVALLI-SFORZA L.L., FELDMAN M.W., 1995b. Genetic absolute dating based on microsatelites and the origin of modern humans. Proceedings of National Academy of Sciences USA., 92: 6723-6727.

LALOË D., MOAZAMI-GOUDARZI K., SOUVENIR ZAFINDRAJAONA P., 1999. Analyse des correspondances et biodiversité dans les races domestiques. Société Française de Biométrie, 20 mai 1999, Grenoble, 5 pages.

LEBART L, MORINEAU A, WARWICK K, 1984. Multivariate Descriptive Statistical Analysis. J. Wiley, New York.

LEWONTIN R.C., 1974. The Genetic Basis of Evolutionary Change. Columbia U.P.

REYNOLDS J., WEIR B.S., COCKERHAM C.C., 1983. Estimation of the coancestry coefficient: Basis for a short-term genetic distance. Genetics 105, 767-79.

ROHLF F.J., 1988. NTSYS-pc: Numercial Taxonomy and Multivariate Analysis System. Exeter Software. Heritage Lane. Setauket, New York, USA.

THAON D'ARNOLDI C., FOULLEY J.L., OLLIVIER L., 1998. An overview of the Weitzman approach to diversity. Genetics, Selection and Evolution, 30, 149-161.

WEITZMAN M., 1992. On diversity. Quarterly Journal of Economics, 107, 363-405.

WEITZMAN M., 1993. What to preserve? An application of diversity theory to crane conservation. Quarterly Journal of Economics, 108, 157-183.

WRIGHT S., 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19, 395-420.