

ANÁLISIS DE ESTIMADORES DE PARENTESCO BASADOS EN INFORMACIÓN MOLECULAR EN UNA POBLACIÓN COMERCIAL DE RODABALLO (*S. MAXIMUS*).

Rodríguez-Ramilo, S.T.¹, Toro, M.A.¹, Martínez, P.², Castro, J.², Bouza, C.², Fernández, J.¹

1. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (ramilo@inia.es)

2. Dpto. de Genética, Facultad de Veterinaria, Universidad de Santiago de Compostela

INTRODUCCIÓN

El conocimiento de las relaciones de parentesco entre los individuos que integran una población de animales o de plantas es importante en numerosas áreas de investigación en genética básica y aplicada. Esta información es necesaria, por ejemplo, para estimar componentes de varianza y valores mejorantes (Lynch y Walsh, 1998), y también sirve de criterio para el manejo de poblaciones con el objetivo de evitar la pérdida de diversidad y el incremento de la consanguinidad (Caballero y Toro, 2000 y 2002). Habitualmente, el parentesco se calcula a partir de la información genealógica. Sin embargo, la ausencia de esta información es frecuente en poblaciones naturales e incluso en poblaciones comerciales, especialmente en acuicultura. Por ejemplo, sería interesante obtener algún tipo de conocimiento sobre la estructura familiar de la población en estudio para elegir los individuos que contribuirán a la siguiente generación de forma que no estén emparentados, y así mantener la máxima diversidad genética posible en la población en cuestión.

El desarrollo y aplicación de marcadores moleculares altamente polimórficos (especialmente los microsatélites) ha permitido que se puedan estimar las relaciones entre individuos cuando parte o la totalidad de la genealogía no está disponible (Awise, 1994; Frankham *et al.*, 2002). Consecuentemente, se han desarrollado numerosos estimadores de parentesco para inferir la genealogía indirectamente a partir de la información molecular proporcionada por los distintos tipos de marcadores. Una descripción de la mayoría de estos estimadores puede encontrarse en Toro *et al.* (2002) y Fernández y Toro (2006), pudiendo clasificarse en aquéllos que estiman parentescos por parejas de individuos (métodos *pairwise*) y aquéllos basados en la reconstrucción explícita de la genealogía.

El objetivo del presente estudio es evaluar el comportamiento de diversos estimadores de parentesco basados en la información molecular en una población comercial de rodaballo (*Scophthalmus maximus*) cuya genealogía es conocida.

MATERIAL Y MÉTODOS

Se genotiparon un total de 560 individuos procedentes de una población comercial de rodaballo empleando 11 microsatélites con un elevado grado de polimorfismo (Castro *et al.*, 2004). En la población evaluada el número medio de alelos fue de 14,18 y la diversidad génica (heterocigosidad esperada) promedio de 0,80. La población analizada estaba constituida por 140 familias de 4 hermanos. Algunas de las familias compartían un progenitor por lo que se pudieron encontrar 15 familias de medios hermanos, englobando en total 28 de las familias de hermanos.

El coeficiente de parentesco entre los individuos i y j (probabilidad de que un alelo elegido al azar del individuo i sea idéntico por descendencia a uno elegido al azar del individuo j ; Malécot, 1948) se obtuvo a partir de la genealogía. Este coeficiente toma valores de 0, 0,125 y 0,25 si i y j no están relacionados (NR), son medios hermanos (MH) o son hermanos (H), respectivamente. De forma análoga, también es posible definir el parentesco molecular como la probabilidad de que dos alelos elegidos al azar, uno de cada individuo, sean iguales. Esta probabilidad puede calcularse para cada alelo de un locus y sumarse para todos los alelos de un locus marcador. Considerando varios marcadores, el parentesco molecular estimado será la media aritmética de los valores obtenidos con cada uno de ellos.

A partir de la información que proporciona el parentesco molecular se han propuesto distintos estimadores que intentan inferir el parentesco genealógico. En la Tabla 1 se muestra un resumen de los estimadores considerados en este estudio.

Tabla 1. Estimadores de parentesco analizados

Abreviatura	Tipo	Referencia
f_{QG}	parejas de individuos	Queller y Goodnight (1989)
f_R	parejas de individuos	Ritland (1996)
f_{LR}	parejas de individuos	Lynch y Ritland (1999)
f_W	parejas de individuos	Wang (2002)
f_{FT}	reconstrucción de genealogías	Fernández y Toro (2006)

Debido a que los estimadores de parentesco por parejas de individuos proporcionan estimas que se distribuyen de forma continua, se establecieron dos puntos de corte para discriminar entre las tres clases de relaciones genealógicas disponibles (H, MH y NR). El umbral para distinguir entre NR y MH se estableció en 0,0625 y el correspondiente entre MH y H en 0,1875, que son puntos equidistantes de los respectivos valores de parentesco (0 para NR, 0,125 para MH y 0,25 para H). En los estimadores de parentesco por parejas de individuos las frecuencias alélicas se estimaron empleando la información molecular disponible de la población actual en estudio. El estimador f_{FT} consiste básicamente en simular genealogías y aceptar aquella que muestre mayor correlación con el parentesco molecular.

Para analizar el ajuste entre las relaciones genealógicas reales y las estimadas se evaluó la proporción de asignaciones correctas e incorrectas (Thomas y Hill, 2002). Al considerar individuos H, MH y NR, existen 9 situaciones dependiendo de la relación estimada y la real. Tres de ellas se corresponden con relaciones estimadas correctamente, otras 3 son sobreestimaciones y las 3 últimas se corresponderían con subestimaciones.

También se evaluó el porcentaje de tripletes incongruentes de H. Este tipo de incongruencias se producen en situaciones en las que los individuos i y j son clasificados como H, los individuos j y k también son clasificados como H, pero la relación estimada entre los individuos i y k no es H.

RESULTADOS Y DISCUSIÓN

La Tabla 2 muestra el número de asignaciones correctas e incorrectas correspondientes a los distintos estimadores de parentesco analizados. Los estimadores que detectan un mayor número de parejas de individuos NR asignados correctamente son f_R y f_{LR} . Sin embargo, los demás estimadores de parentesco por parejas de individuos (f_{QG} y f_W) presentaron un éxito más reducido en la asignación de parejas NR. El estimador que emplea la reconstrucción explícita de la genealogía (f_{FT}) presentó un comportamiento intermedio en este caso. En la asignación de MH y H los estimadores de parentesco por parejas de individuos presentaron un comportamiento más adecuado.

Se ha demostrado (Toro *et al.*, 2002) que los estimadores de parentesco por parejas de individuos presentan un comportamiento adecuado en casos en los que se dispone de una estima precisa de las frecuencias alélicas en la población de referencia, como podría ser el caso de la población analizada en este estudio que se remonta una sola generación y con un censo elevado por lo que la deriva genética es muy reducida.

Tabla 2. Número de asignaciones correctas (coincidencia entre el parentesco actual y el estimado) e incorrectas. En columnas se indican las relaciones actuales y en filas las estimadas. Numero total de parejas de individuos: 156.520

Estimado	f_{QG}			f_R			f_{LR}			f_W			f_{FT}		
	NR	MH	H	NR	MH	H	NR	MH	H	NR	MH	H	NR	MH	H
NR	126.431	75	9	146.645	116	47	142.213	87	21	126.805	80	6	129.079	162	252
MH	27.504	130	172	8.093	97	388	12.134	114	226	27.234	123	170	26.063	84	384
H	1.473	51	675	670	43	421	1.061	55	609	1.369	53	680	266	10	220

El porcentaje de asignaciones correctas e incorrectas, así como el porcentaje de tripletes de hermanos incongruentes se muestra en la Tabla 3. El estimador f_{FT} presentó un

comportamiento intermedio con respecto a los estimadores de parentesco por parejas de individuos en lo que refiere al porcentaje de asignaciones correctas y sobreestimadas. El porcentaje de subestimadas en todos los casos fue menor que el de sobreestimadas, posiblemente debido a que en la población evaluada, la inmensa mayoría de las parejas de individuos no tienen ningún tipo de parentesco en común.

Tabla 3. Porcentaje de asignaciones correctas e incorrectas (subestimadas y sobreestimadas) y porcentaje de tripletes de hermanos (H) incongruentes

Porcentaje	f_{OG}	f_R	f_{LR}	f_W	f_{FT}
Asignaciones subestimadas	0,2	0,4	0,2	0,2	0,5
Asignaciones sobreestimadas	18,5	5,6	8,5	18,3	16,8
Asignaciones correctas	81,3	94,0	91,3	81,5	82,7
Tripletes H incongruentes	83,5	67,9	72,9	83,3	0,0

En cuanto al porcentaje de tripletes de hermanos incongruentes, es importante resaltar que el estimador f_{FT} , por su naturaleza, no produce incongruencias, mientras que cualquiera de los estimadores de parentesco por parejas de individuos muestra más de un 65% de incongruencias. Hay que hacer notar que podrían darse otras situaciones de incompatibilidad en el establecimiento de familias de H teniendo en cuenta los genotipos, pero no se han evaluado en este estudio.

En resumen, los resultados indican que los métodos para inferir parentesco por parejas de individuos a partir de la información molecular (Queller y Goodnight, 1989; Ritland, 1996; Lynch y Ritland, 1999; Wang, 2002) presentaron un buen comportamiento al evaluar parejas de individuos debido, probablemente, a que la población evaluada de rodaballo dispone de una situación poblacional favorable y una información molecular disponible muy precisa.

Sin embargo, al evaluar el comportamiento de los estimadores de parentesco en más de dos individuos a la vez, es el método que implica una reconstrucción explícita de la genealogía (Fernández y Toro, 2006), el que presenta el mejor comportamiento. Esta ventaja del método de reconstrucción de genealogías podría ser de gran importancia, por ejemplo, para realizar contribuciones y apareamientos de mínimo parentesco. Además, en la fundación de un stock de peces es probable que no se disponga de una estructura poblacional tan marcada y también podría darse la posibilidad de contar con parentescos de mayor grado. Por tanto, se esperaría que en este tipo de situaciones el método que implica una reconstrucción explícita de la genealogía presentase un comportamiento más aceptable que los estimadores de parentesco por parejas de individuos.

REFERENCIAS BIBLIOGRÁFICAS

- Avise, J.C. 1994. Chapman and Hall, New York, USA.
- Caballero, A., Toro, M.A. 2000. Genet. Res. 75: 331-343.
- Caballero, A., Toro, M.A. 2002. Conserv. Genet. 3: 289-299.
- Castro, J., Bouza, C., Presa, P., Pino-Querido, A., Rianza, A., Ferreiro, I., Sánchez, L., Martínez P. 2004. Aquaculture 242: 119-135.
- Fernández, J., Toro, M.A. 2006. Mol. Ecol. 15: 1657-1667.
- Frankham, R., Ballou, J.D., Briscoe, D.A. 2002. Cambridge University Press, Cambridge, Reino Unido.
- Lynch, M., Ritland, K. 1999. Genetics 152: 1753-1766.
- Lynch, M., Walsh, B. 1998. Sinauer Associates. Sunderland, Massachusetts, USA.
- Malécot, G. 1948. Masson et. Cie., Paris, Francia.
- Queller, D.C., Goodnight K.F. 1989. Evolution 43: 258-275.
- Ritland, K. 1996. Genet. Res. 67: 175-186.
- Thomas S.C., Hill, W.G. 2002. Genet. Res. 79: 227-234.
- Toro M.A., Barragán C., Óvilo C., Rodríguez J., Rodríguez, C., Silió, L. 2002. Conserv. Genet. 3: 309-320.
- Wang, J. 2002. Genetics 160: 1203-1215.