

ANÁLISIS DE PATERNIDAD CON (POCOS) DATOS DE SECUENCIACIÓN

Casellas¹, J., Martín de Hijas-Villalba¹, M., Vázquez-Gómez¹, M. e Id-Lahoucine², S.

¹Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, 08193 Bellaterra, España. ²Animal and Veterinary Science Group, Scotland's Rural College, Edinburgh EH9 3JG, Reino Unido; joaquim.casellas@uab.cat

INTRODUCCIÓN

La legislación europea actual para las razas de ganado autóctonas pone un énfasis especial en el control de genealogías, requiriéndose de test genéticos de paternidades en la mayoría de razas, sobretodo aquellas con manejo más extensivo. Esto implica un gasto económico importante para muchas asociaciones de criadores, y resulta conveniente explorar alternativas que permitan reducir ese coste, e incluso generar datos que puedan ser usados con otros objetivos además del análisis de paternidades. Esta investigación se centra en desarrollar un test de paternidades a partir de datos genómicos con poca profundidad de secuenciación, así como evaluar su potencia estadística.

MATERIAL Y MÉTODOS

Los análisis se realizaron sobre datos simulados en base a genomas de 30 pares de cromosomas, con 100 cM y 3 millones de SNPs por cromosoma. Cada población evolucionó durante 1.000 generaciones no solapadas con un tamaño efectivo de 100 individuos, una tasa de mutación de 10^{-4} , y recombinación de acuerdo con la función de Kosambi (Kosambi, 1944). Alcanzada la generación 1.000, solo se mantuvieron aquellas poblaciones con $1.000.000 \pm 10\%$ SNPs polimórficos ($MAF > 0$), y se expandieron a 10.000 (generación 1.001) y 25.000 individuos (generación 1.002). Se usaron estas dos últimas generaciones de 10 poblaciones distintas para realizar los test de paternidad mediante inferencia evidencial (Edwards, 1972; Bickel, 2012). Padres, madres e hijos se secuenciaron a distintas profundidades (0,01, 0,05, 0,1 y 0,5X) y con distintas longitudes de lectura (100, 1.000 y 10.000 nucleótidos), con una tasa de error por SNP que osciló entre 10^{-5} y 10^{-2} . Las funciones de verosimilitud desarrolladas para el análisis de paternidades dependían de la tasa de error de secuenciación (TE) y la frecuencia alélica (FA) para los distintos SNPs implicados.

RESULTADOS Y DISCUSIÓN

Usando los valores de TE y FA de la simulación, la secuenciación a una profundidad de 0,05X era suficiente para alcanzar un 100% de sensibilidad y especificidad en los test de paternidad (1.000 réplicas), mientras que cuando se realizaban con datos de secuenciación a 0,1X, la sensibilidad bajaba al 98,9% (99,0%) y la sensibilidad al 99,5% (99,4%), asumiendo que la madre era conocida (desconocida). No obstante, estos resultados se basaban en el conocimiento certero de TE y FA, lo cual resulta difícil en muestras de campo.

Para evaluar un escenario más realista, se repitieron los test asumiendo TE distintas y homogéneas en todos los SNPs (0%, 1% y 10%). Tanto el 0% como el 1% de TE seguían generando un 100% de sensibilidad y especificidad a una profundidad de secuenciación de 0,05%, mientras que una TE del 10% (un orden de magnitud superior a la TE máxima de simulación) generaba un 24,5% (madre conocida) y un 22,8% (madre desconocida) de falsos negativos, sin presencia de falsos positivos. En cuanto a las FA, se evaluó su impacto simulando que se estimaban a partir de distintos tamaños muestrales. Al igual que en el caso anterior, no se generaron falsos positivos a 0,05X, y sí un 0,1% (0,1%), 0,6% (0,3%) y 10,2% (14,8%) de falsos negativos con madre conocida (desconocida) para FA muestreadas a partir de 100, 10 y 5 individuos, respectivamente.

CONCLUSIÓN

La secuenciación de baja densidad se sugiere como una alternativa prometedora a los paneles de SNPs para el análisis de paternidades en las especies de ganado doméstico, generándose datos genómicos acumulativos y que podrían reutilizarse con otros fines científicos.

REFERENCIAS BIBLIOGRÁFICAS

- Bickel D.R. 2012. *Statist. Sinica*. 22: 1147-1198
- Edwards A.W.F. 1972. *Likelihood*. Cambr. Univ. Press, Cambridge, Reino Unido
- Kosambi, D. 1944. *Ann. Eugen.* 12: 172-175

Agradecimientos: Proyecto CGL2016-80155-R y beca BES-2017-080596 (Ministerio de Economía y Competitividad).