

ESTIMACIÓN DE VARIANZA EPIGENÉTICA TRANSGENERACIONAL

Varona¹, L., Munilla, S., Mouresan, E. F., González-Rodríguez, A., Moreno, C. y Altarriba, J.

¹Unidad de Genética Cuantitativa y Mejora Animal. Facultad de Veterinaria. Universidad de Zaragoza. c/Miguel Servet 177. 50013. ZARAGOZA. lvarona@unizar.es

INTRODUCCIÓN

La epigenética estudia las variaciones en la expresión génica que no son causadas por modificaciones de la secuencia de ADN (Eccleston *et al.*, 2007). La causas de esta variabilidad se atribuyen, entre otras, a metilación del ADN o modificaciones de la histonas (Jablonka y Raz, 2009). En ocasiones, estas modificaciones pueden afectar a las células germinales y transmitirse a la descendencia. (Youngson y Withelaw, 2008). Pese a todo, la magnitud de este fenómeno no se conoce (Heard y Martienssen, 2014).

Desde la perspectiva de la genética cuantitativa, la presencia de herencia epigenética transgeneracional implica una redefinición del parecido entre parientes. Tal *et al.* (2010) desarrollaron un modelo que permite calcular el parecido entre parientes en función de la heredabilidad epigenética (γ^2) y el coeficiente de borrado (ν), a partir del cual se estima el coeficiente de transmisión epigenética ($1-\nu$). El objetivo de este trabajo es presentar un análisis Bayesiano que permite estimar la varianza epigenética transgeneracional, el coeficiente de borrado, y el coeficiente de transmisión epigenética a partir de la información genealógica y fenotípica.

MATERIAL Y MÉTODOS

Se define un modelo mixto $\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{Zw} + \mathbf{e}$, donde \mathbf{y} es el vector de datos fenotípicos, \mathbf{b} es el vector de efectos sistemáticos, \mathbf{u} es el vector de efectos genéticos aditivos, \mathbf{w} es el vector de efectos epigenéticos transmisibles y \mathbf{e} es el vector de residuos. Las asunciones para las distribuciones a priori de \mathbf{u} , \mathbf{w} y \mathbf{e} son las siguientes distribuciones normales multivariantes (MVN):

$$\mathbf{u} \sim MVN(\mathbf{0}, \mathbf{A}\sigma_u^2) \quad \mathbf{w} \sim MVN(\mathbf{0}, \mathbf{T}\sigma_w^2) \quad \mathbf{e} \sim MVN(\mathbf{0}, \mathbf{I}\sigma_e^2)$$

donde σ_u^2 , σ_w^2 y σ_e^2 son las varianzas genética aditiva, epigenética y residual, respectivamente.

\mathbf{A} es la matriz de parentesco numerador y \mathbf{T} es la matriz de relaciones epigenéticas.

La estructura de la matriz \mathbf{T} se define a partir de la relación recursiva entre el efecto epigenético de un individuo (w_i) con respecto al efecto epigenético de su padre (w_{β_i}) y de su madre (w_{mi}): $w_i = \lambda w_{\beta_i} + \lambda w_{mi} + \varepsilon_i$ donde $\lambda = 0.5(1-\nu)$, ν es el coeficiente de borrado y $(1-\nu)$ el coeficiente de transmisión epigenética. Además, ε_i es el efecto epigenético residual del individuo i . En notación matricial:

$$\mathbf{w} = \mathbf{Pw} + \boldsymbol{\varepsilon}$$

donde la matriz \mathbf{P} define la relación recursiva entre los efectos epigenéticos del padre y de la madre. Si se utiliza esta estructura es posible definir

$$\mathbf{T}^{-1} = [\mathbf{I} - \mathbf{P}']\mathbf{Q}_T^{-1}[\mathbf{I} - \mathbf{P}]$$

donde \mathbf{Q}_T :

$$\mathbf{Q}_T = \mathbf{V}(\boldsymbol{\varepsilon})\frac{1}{\sigma_w^2} = [\mathbf{I} - \mathbf{P}]\mathbf{T}[\mathbf{I} - \mathbf{P}']$$

Esta matriz tiene una estructura diagonal con valor uno para los individuos de la población base y $1-\lambda^2$ y $1-2\lambda^2$ para aquellos con uno o dos ancestros conocidos. Si se utiliza esta estructura, se puede definir el siguiente algoritmo para construir \mathbf{T}^{-1} mediante la lectura secuencial de la información genealógica:

- Si el padre y la madre son conocidos sumar 1 a la diagonal (i,i)

b) Si solo un progenitor (p) es conocido sumar $\frac{1}{(1-\lambda^2)}$ a la diagonal (i,i) , $\frac{-\lambda}{(1-\lambda^2)}$ a los elementos (i,p) , (p,i) y $\frac{\lambda^2}{(1-\lambda^2)}$ al elemento (p,p)

c) Si los dos ancestros (p and m) son conocidos sumar $\frac{1}{(1-2\lambda^2)}$ a la diagonal (i,i) , $\frac{-\lambda}{(1-2\lambda^2)}$ a los elementos (i,p) , (i,m) , (m,i) , (p,i) y $\frac{\lambda^2}{(1-2\lambda^2)}$ a los elementos (p,p) , (p,m) , (m,p) , (m,m)

Ejemplo: Población Vacuna Pirenaica.

Se ha utilizado, como ejemplo, la base de datos de la población vacuna Pirenaica, para el carácter peso al nacimiento. Se utilizaron, 78.209 registros de peso al nacimiento y una genealogía de 125.974 entradas.

El modelo de análisis fue:

$$y = \mathbf{Xb} + \mathbf{Z}_1\mathbf{u} + \mathbf{Z}_2\mathbf{m} + \mathbf{Z}_3\mathbf{p} + \mathbf{Z}_4\mathbf{h} + \mathbf{Z}_5\mathbf{w} + \mathbf{e}$$

donde \mathbf{b} incluyó sexo y edad de la madre, \mathbf{u} y \mathbf{m} son los vectores de los efectos directos y maternos, \mathbf{p} es el vector de efectos ambientales maternos, \mathbf{h} es el vector de efectos rebaño-año-estación, \mathbf{w} es el vector de efectos epigenéticos, y \mathbf{e} es el vector de residuos.

Se utilizaron dos modelos (I y II). El modelo I incluyó varianza epigenética transgeneracional y el Modelo II no, con un algoritmo de muestreo de Gibbs (Gelfand y Smith, 1990). Los modelos se compararon mediante la probabilidad pseudo-log-marginal de los datos (Gelfand, 1996).

RESULTADOS Y DISCUSIÓN

En la Figura 1 se presenta un ejemplo de genealogía con la matriz de relaciones epigenéticas (\mathbf{T}) y su inversa (\mathbf{T}^{-1}). El procedimiento de cálculo es similar al propuesto por Quaas (1976) para la matriz de parentesco numerador o por Quintanilla et al. (1999) para la matriz de relaciones ambientales maternas. El único limitante es que exige ser calculado en cada iteración del proceso de muestreo de Gibbs. Pese a todo, la implementación de este procedimiento es sencilla, ya que sólo involucra muestreo de distribuciones normales univariantes y chi-cuadrado invertidas.

En la Tabla 1 se presentan los resultados del análisis. En ella se observa que el modelo seleccionado como más plausible no incluyó varianza epigenética.

El modelo utilizado asumió que el coeficiente de borrado es el mismo en machos y en hembras. El modelo propuesto se puede generalizar de manera sencilla a coeficientes de borrado diferentes entre sexos y ambientes, de manera que se pueden estudiar los efectos de estas situaciones en los coeficientes de transmisión epigenética.

Por otra parte, el modelo de transmisión epigenética se puede ver en un sentido más amplio, ya que permite modelizar mecanismos de herencia no genética, como la transmisión cultural.

REFERENCIAS BIBLIOGRÁFICAS

- Eccleston, A., DeWitt, N., Gunter, C., Marte, B., Nath, D. 2007. *Nature* 447:395-400
- Gelfand, A. E. & Smith A. F. M. 1990. *Journal of American Statistical Association* 85:398-409
- Gelfand, A. E. 1996. En *Markov Chain Monte Carlo in Practice*, edited by W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. Chapman & Hall, London.
- Heard, E & Martienssen, R. A. 2014. *Cell* 157:95-109
- Jablonska, E & Raz, G. 2009. *Q. Rev. Biol.* 84:131-176.
- Quintanilla, R., Varona, L., Pujol, M. R., Piedrafita, J. 1999. *J. Anim. Sci.* 77: 2904-2917.
- Tal, O., Kisdí, E. & Jablonka, E. 2010. *Genetics* 184:1037-1050.
- Youngson, N. A. & Withelaw, E. 2008. *Annu. Rev. Genomics Hum. Gen.* 9:233-257.

Agradecimientos: Los autores agradecen al proyecto de investigación AGL2010-15903 y a la Confederación Nacional de Asociaciones de Ganado Pirenaico (CONASPI).

Tabla 1. Media posterior y desviación típica posterior (entre paréntesis) para los modelo I y II

	Model I	Model II
h^2	0.377 (0.035)	0.412 (0.012)
m^2	0.149 (0.007)	0.148 (0.007)
γ^2	0.036 (0.034)	-
λ	0.398 (0.103)	-
ν	0.204 (0.207)	-
$1 - \nu$	0.796 (0.207)	-
LogCPO	-287542.3	-287475.1

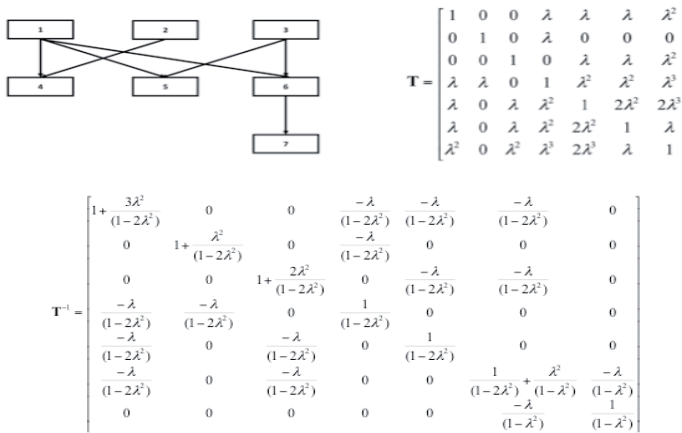


Figura 1. Ejemplo de genealogía, matriz de relaciones epigenéticas (T) y su inversa (T^{-1}).

EPIGENETIC VARIANCE COMPONENT ESTIMATION

ABSTRACT: Epigenetics has become one of the major areas of biological research. However, the degree of phenotypic variability that is explained by epigenetic processes still remains unclear. From a quantitative genetics perspective, the estimation of variance components is achieved by means of the information provided by the resemblance between relatives. In a previous study, this resemblance was described as a function of the epigenetic variance component and a reset coefficient that indicates the rate of dissipation of epigenetic marks across generations. Given these assumptions, we propose a Bayesian mixed model methodology that allows the estimation of epigenetic variance from a genealogical and phenotypic database. The methodology is based on the development of a T matrix of epigenetic relationships that depends on the reset coefficient. In addition, we present a simple procedure for the calculation of the inverse of this matrix (T^{-1}) and a Gibbs sampler algorithm that obtains posterior estimates of all the unknowns in the model. The new procedure was used with a beef cattle database. The posterior estimate of transgenerational epigenetic variability was very low and a model comparison test indicated that a model that did not included it was the most plausible.

Keywords: Epigenetics, Bayesian Analysis, Genetic Variance, Resemblance Between Relatives