

La heredabilidad faltante y el epigenoma: retos en la era post-genómica

O. González-Recio^{1*}, L.A. García-Cortés¹, B. Villanueva¹, E. Ugarte²

¹Dpto. Mejora Genética Animal-INIA. 28040 Madrid (Spain)

²NEIKER-TECNALIA. Campus Agroalimentario de Arkaute. Apdo 46 E-01080 Vitoria-Gasteiz (Araba)

Resumen. La variación epigenómica y los mecanismos de control epigenéticos se están revelando en el siglo XXI como un factor de gran importancia en la regulación genética de los caracteres complejos. La definición moderna de la epigenética es el estudio de los cambios, posiblemente heredables, en la expresión génica que no son debidos a cambios de la secuencia del DNA. Una de las marcas epigenéticas más conocidas es la metilación del DNA (DNAm). Avances recientes en las técnicas de secuenciación permiten ya obtener información sobre metilaciones del ADN a lo largo de todo el genoma que pueden tener algún impacto sobre la expresión de caracteres de interés productivo (e.g. producción de leche, crecimiento, resistencia a enfermedades, etc...). En este trabajo se evalúa la repercusión que la metilación puede tener sobre las evaluaciones de mérito genético (tanto tradicionales como genómicas) y se desarrolla un método *ad hoc* para incorporar la información de metilación del ADN en dichas evaluaciones. Se utilizó el genotipo de 2104 individuos Holstein para simular un carácter productivo de naturaleza infinitesimal, y su predisposición a la metilación del ADN. En base a esta predisposición genética la metilación y a circunstancias ambientales, se simuló un epigenotipo para cada individuo. Los datos se analizaron con un modelo BLUP tradicional, un BLUP genómico (G-BLUP) y un LASSO bayesiano que incorpora información epigenómica. La precisión de las evaluaciones de mérito genético aditivo se vio afectada cuando existía influencia epigenómica sobre el carácter, independientemente del método usado, con un descenso en la predicción entre un 15 y un 20% con respecto a la ausencia de influencia epigenómica. El LASSO Bayesiano epigenómico no mejoró la capacidad predictiva para el efecto genético aditivo, pero si predijo con mayor exactitud la expresión fenotípica de los individuos dada la metilación del ADN. La metilación del ADN constituye un ruido externo que no es posible controlar con los modelos usados actualmente en la predicción de mérito genético. Es necesario desarrollar métodos que puedan controlar este ruido generado por la información epigenómica para no perjudicar la precisión de las evaluaciones de mérito genético. El conocimiento de la causa y repercusión de las marcas epigenómicas sobre la expresión de los caracteres de interés en producción animal, pueden contribuir a la obtención de evaluaciones genéticas más fiables y al control ambiental de circunstancias productivas perjudiciales, con el objetivo de generar conocimiento en la interacción genotipoambiente y obtener animales más productivos, sanos y eficientes.

Palabras clave: Epigenética, evaluaciones genéticas; metilación del ADN;

* Corresponding author: gonzalez.oscar@inia.es

Introducción

Durante la era genómica se ha constatado que existe una parte de la heredabilidad que no es posible detectar con los modelos genómicos utilizados actualmente. Es lo que se ha llamado heredabilidad faltante, o “missing heritability” en inglés (Manolio et al., 2009). Las causas se han asociado a diferentes aspectos:

- Variantes genéticas raras, con frecuencias alélicas muy bajas.
- La utilización de modelos de un solo efecto (marcador) ignorando el resto del genoma.
- Efectos de epistasis y complejidad de rutas biológicas.
- Variaciones estructurales en el genoma como los “copy number variation” (CNV).
- Falta de información familiar o datos adecuadamente recogidos (genealogías faltantes o definición de fenotipo no sistemática).
- Epigenética o factores de regulación de la secuencia del genoma.

Este trabajo se centra en este último aspecto: La epigenética. La definición moderna de epigenética es « *el estudio de cambios, posiblemente heredables, en la expresión de los genes que no son debidos a variaciones en la secuencia del DNA* » (Richards, 2006). La información epigenética se puede considerar como la gramática del DNA, ya que son mecanismos asociados a la regulación y expresión de las proteínas. El estudio realizado por Fraga et al. (2005) muestra cómo gemelos homocigóticos, con idéntica secuencia de DNA, presentan importantes diferencias en sus marcas epigenéticas y que éstas pueden ser la causa de la diferente expresión de sus genes a lo largo de la vida. La epigenética, está adquiriendo cada vez más importancia como una herramienta potencial en los procesos de selección artificial. Sin embargo, no existe mucha información sobre su funcionamiento y en ocasiones se utiliza con planteamientos equivocados que por ejemplo, presentan las metilaciones como un factor genético transmisible de la forma clásica. Los efectos de la herencia epigenética son aún desconocidos en muchas especies (Henderson y Jacobsen, 2007; Jablonka y Raz, 2009; Migicovsky y Kovalchuk, 2011; Morgan et al., 1999). Independientemente de la herencia epigenética transgeneracional, una determinada causa ambiental puede afectar a tres generaciones diferentes (i.e. la hembra productiva, su feto, y las células germinales del feto). Si estas circunstancias ambientales provocan cambios en el epigenoma de los individuos, los efectos pueden observarse durante al menos a tres generaciones (Heijmans et al., 2008; Nijland et al. (2008).

Por otro lado, la epigenética puede contribuir a explicar la interacción de los sistemas de producción con la genética, ya que la variabilidad epigenética es principalmente inducida por el ambiente como la dieta o el estrés (Petronis, 2010). De este modo se espera que animales producidos bajo diferentes circunstancias ambientales presenten diferentes patrones de metilación en sus respectivos genomas y que no pueden ser detectado a través de un análisis de ADN. Animales alimentados con concentrado o unifeed pueden presentar patrones de metilación diferentes a aquellos animales con alimentación basada en pastoreo aunque tuvieran el mismo patrón genético. Del mismo modo, la metilación de aquellos animales bajo condiciones intensivas de manejo serán diferentes a la de aquellos animales en sistemas semi-extensivos o de pastoreo. Coolen et al. (2011) y Rosenfield (2010) demuestran que además existen factores genéticos intrínsecos que regulan que algunos genotipos sean más susceptibles a adquirir grupos metilo que otros.

Según el modelo de Fisher (1918) el fenotipo observado puede descomponerse en genotipo más ambiente ($P=G+E$). Los modelos tradicionales realizan esta descomposición en la forma $y_i = \mathbf{X}\mathbf{b} + \mathbf{Z}u_i + e_i$, donde y_i es el fenotipo del individuo i , \mathbf{b} el vector de efectos ambientales, u_i es el efecto genético, y e_i el residuo, con \mathbf{X} y \mathbf{Z} siendo las matrices de incidencia correspondientes. La información epigenómica no está considerada en esta descomposición, y se desconoce el efecto que esto puede tener sobre la predicción del mérito genético de los individuos.

El objetivo de este trabajo es conocer que efecto puede tener la metilación del ADN sobre los métodos usados actualmente en las evaluaciones genéticas y realizar una primera aproximación metodológica para incorporar la información epigenómica en especies domésticas.

Materiales y métodos

Simulación

La simulación de epigenotipos es aún un tema por desarrollar y resolver. Como una primera aproximación en este estudio se usó el genotipo de los 4 primeros cromosomas de 2104 toros Holstein cedidos por CONAFE ($p=8887$ SNPs).

Para cada individuo se simularon dos caracteres cuantitativos: 1) predisposición genética a la metilación del genoma (λ) y 2) un carácter de interés económico (e.g. producción de leche, tasa de crecimiento o resistencia a una determinada enfermedad) afectado por la metilación, que en lo sucesivo llamaremos objetivo de selección (t). Se asumió que ambos caracteres fueron infinitesimales, con una gran cantidad de genes, 1000, con pequeño efecto, y distribuidos homogéneamente a lo largo de los cromosomas, con correlación genética igual a cero entre ellos. Se asumió únicamente dos álelos por gen, en desequilibrio de ligamiento completo con uno de los p SNPs genotipados. Los efectos de sustitución alélica, $\beta_{\lambda j}$ y $\beta_{t j}$, se muestrearon de un distribución normal con media 0 y varianza 1. El verdadero mérito genético de los individuos para ambos caracteres (u_{ij} y $u_{\lambda i}$) se calculó como el sumatorio para cada gen del producto del número de alelos mutante y el efecto de sustitución del correspondiente carácter para cada respectivo alelo.

El epigenotipo de cada individuo en cada locus se simuló siguiendo el proceso aquí descrito:

Se asumió una cantidad de posiciones a lo largo del genoma que pueden estar metiladas (en este caso el número fue igual al número de SNPs). Estas posiciones de variable metilada (MVP _{j}), regularían la transcripción del posible gen sobre el que estuviese situadas. Dicha MVP se asoció al SNP j más cercano del individuo i (MVP _{ij}), y se codificó inicialmente como 0 si ninguno de los cromosomas presenta metilación en esa posición, 0.5 si heredó la marca epigenética de uno de los progenitores (padre o madre) y 1 si ambos cromosomas heredados de su padre y su madre mantuvieron la marca de metilación en dicha posición. Cuando una MVP se hereda transgeneracionalmente, esta aparece en todas las células del individuo. Esta

probabilidad de herencia epigenética transgeneracional (t) fue fijada a 0.05, e indica el porcentaje de marcas de metilación que no son eliminadas durante el borrado epigenético que ocurre durante la meiosis.

Posteriormente, la probabilidad de que un determinado locus j en el individuo i esté metilado dependió de los siguiente factores:

- 1) La predisposición genética del individuo a la metilación de su genoma (u_{2i}).
- 2) Al ambiente, con probabilidad de adquirir metilación igual a v y de perderla igual a l .
 - a. donde v (tasa de ganancia) es la probabilidad de que el locus j adquiera una marca de metilación condicionada a unas determinadas condiciones ambientales (e.g. dieta, estrés,...).
 - b. Y l (tasa de pérdida) es la probabilidad de que el locus j pierda la marca de metilación, dado que está metilada, condicionada a unas determinadas condiciones ambientales (e.g. dieta, estrés,...).

Nuestra simulación asumió que el epigenotipo y el fenotipo de los individuos se midieron en un mismo tiempo t_0 , idéntico para todos los individuos. De este modo, a cada MVP se le asignó un valor entre 0 y 1, que indica el porcentaje de células del tejido diana del carácter que presentan metilación del ADN en dicha posición.

Finalmente el fenotipo del individuo i para el objetivo de selección se simuló como:

$$y_i = \sum_{j=1}^p x_{tj} \beta_{tj} (1 - p(x_{tj}^m)) + \sum_{j=1}^p x_{tj}^m \beta_{tj}^m p(x_{tj}^m) + e_i$$

donde el primer término corresponde al efecto aditivo para el carácter objetivo teniendo en cuenta el porcentaje, $p(x_{tj}^m)$, de células del tejido diana con metilación asociada al locus j . El término $\sum_{j=1}^p x_{tj}^m \beta_{tj}^m$ es el efecto del locus j en el fenotipo cuando está metilado, asumiendo que ni el alelo A ni el a se expresan. Este efecto fue muestreado de una distribución normal con media 0 y varianza 1, independientemente del efecto aditivo de dicho locus.

El residuo e_i se asumió distribuido como $N(0, \sigma_e^2)$, donde σ_e^2 es la varianza residual.

Los valores de v y l se fijaron a 0.05 y 0.45, respectivamente. El ratio entre la varianza genética aditiva y la residual se fijó en 1. Para el ratio entre la varianza genética de la susceptibilidad a la metilación y la varianza residual se asumieron 2 valores en simulaciones independientes (9/1 y 1/5). Se realizaron 10 réplicas del proceso de simulación descrito para cada uno de los ratios de la varianza genética de susceptibilidad a la metilación.

Métodos de evaluación

Se utilizaron 3 modelos diferentes para el análisis de los datos: un modelo animal tradicional usando un BLUP bayesiano equivalente, un modelo que incorpora la información genómica proporcionada por los SNPs, y un modelo que además incorpora la información del epigenotipo. Estos modelos se describen a continuación.

Modelo 1: BLUP tradicional

$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$, donde \mathbf{b} es el vector de efectos sistemáticos, \mathbf{u} es el efecto poligénico aditivo distribuido a priori como $N(\mathbf{0}, \mathbf{A}\sigma_a^2)$, con \mathbf{A} y σ_a^2 siendo la matriz de relaciones aditivas y la varianza genética aditiva, respectivamente. El vector \mathbf{e} es el residuo distribuido a priori como $N(\mathbf{0}, \sigma_e^2)$. \mathbf{X} y \mathbf{Z} son las correspondientes matrices de incidencia.

Modelo 2: BLUP genómico

$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$, donde la diferencia con el modelo anterior es que \mathbf{u} es el valor genómico del individuo distribuido a priori como $N(\mathbf{0}, \mathbf{G}\sigma_g^2)$, con \mathbf{G} siendo la matriz de parecido genómico calculada con la información de los SNPs y σ_g^2 la varianza genómica de la población.

Modelo 3: LASSO Bayesiano epigenómico

$\mathbf{y} = \mathbf{Xb} + \sum_p \mathbf{z}_j \beta_j + \sum_p \mathbf{z}_{ej} \mathbf{w}_j + \mathbf{e}$, donde $\sum_p \mathbf{z}_j \beta_j$ es el sumatorio de los efectos aditivos de los $p=8887$ SNPs usados en el análisis, y $\sum_p \mathbf{z}_{ej} \mathbf{w}_j$ es el sumatorio de los efectos epigenéticos debidos a la metilación de los p SNPs correspondientes. \mathbf{z} y \mathbf{z}_e son las matrices de incidencia correspondientes. Tanto para los coeficientes de los SNPs como de los efectos de metilación se asumió una distribución a priori doble Laplaciana independiente para los efectos genéticos aditivos y de metilación de la forma:

$$\rho(\beta | \sigma_e^2) = \prod_{j=1}^p \frac{\lambda}{2\sqrt{\sigma_e^2}} e^{-\lambda|\beta_j|/\sqrt{\sigma_e^2}}$$

$$\rho(\mathbf{w} | \sigma_e^2) = \prod_{j=1}^p \frac{\lambda_e}{2\sqrt{\sigma_e^2}} e^{-\lambda_e \mathbf{w}_j / \sqrt{\sigma_e^2}},$$

siendo λ y λ_e los parámetros que controlan el encogimiento a cero de las distribuciones de los respectivos coeficientes aditivos y epigenómicos (Park y Casellas, 2008; de los Campos et al., 2009).

Los análisis se realizaron en un diseño de validación cruzada de 2 pliegues, dejando el 20% de los animales más jóvenes como set de validación, y el resto como población de referencia.

Resultados y discusión

La tabla 1 muestra la correlación de Pearson entre el mérito genético verdadero de los individuos y su predicción en ausencia de efectos epigenómicos. Para un carácter de heredabilidad 0.50, el modelo genómico es mucho más preciso que el BLUP tradicional, y alcanza una precisión de 0.75(0.05) en el set de validación. Esta precisión se ve notablemente perjudicada cuando existen efectos epigenómicos sobre el carácter, aun manteniendo el mismo ratio entre la varianza genética aditiva y la residual (Tablas 2 y 3). Estas tablas muestran que la presencia de efectos epigenómicos empeora la precisión de las predicciones de los valores genéticos de los individuos. Estos resultados indican que los efectos epigenómicos provocan un ruido en los datos que los modelos usados

actualmente no son capaces de tener en cuenta, con un aumento de la estima de la varianza residual y perjudicando las predicciones del mérito genético. En presencia de efectos epigenéticos, los modelos que no incorporan información de metilación del ADN (BLUP, GBLUP) pueden capturar cierta cantidad de información epigenómica, sobre todo cuando la heredabilidad de la susceptibilidad a la metilación es alta, aunque no permiten diferenciar entre variabilidad genética y epigenética.

El LASSO Bayesiano epigenómico fue capaz de incorporar la información de la metilación del DNA. La correlación entre el mérito genético y el predicho fue similar que la alcanzada con el GBLUP, pero este modelo presentó entre un 4 y un 8% de mejora en la precisión de la estima de la expresión real del carácter en el individuo para animales con dato y un 15% para animales jóvenes sin dato. Por tanto, este modelo puede predecir con mayor exactitud que los modelos actuales cual será el rendimiento productivo de un animal o que probabilidad tendría de desarrollar una determinada enfermedad, dada la metilación de su ADN. Además, el LASSO Bayesiano epigenómico es capaz de detectar que metilaciones están asociadas a la expresión de genes que regulan un determinado carácter.

Tabla 1. Correlación de Pearson¹ entre el mérito genético real (TBV) y el estimado (EVGD) para los animales del set de entrenamiento y el de validación, en ausencia de efectos epigenómicos.

	TRAINING	TESTING
	TBV-EVGD	TBV-EVGD
BLUP	0.72(0.01)	0.26(0.09)
GBLUP	0.83(0.02)	0.75(0.05)

¹Promedio de las 10 réplicas. Desviación standard entre parenthesis.

Tabla 2. Correlación de Pearson¹ entre el mérito genético real (TBV) y el estimado (EVGD), y el valor epigenómico real (TEV) y el estimado (EVED) para los animales del set de entrenamiento y el de validación, y utilizando cada uno de los modelos. Ratio entre varianza genética de la susceptibilidad a la metilación y varianza residual igual a 9/1

	TRAINING		TESTING	
	TBV-EVGD	TEV-EVED	TBV-EVGD	TEV-EVED
BLUP	0.55(0.05)	0.76(0.04)	0.22(0.10)	0.22(0.09)
GBLUP	0.69(0.09)	0.81(0.04)	0.61(0.09)	0.59(0.07)
epiBLASSO	0.72(0.02)	0.84(0.04)	0.63(0.09)	0.68(0.11)

¹Promedio de las 10 réplicas. Desviación estándar entre paréntesis.

Tabla 3. Correlación de Pearson¹ entre el mérito genético real (TBV) y el estimado (EVDG), y el valor epigenómico real (TEV) y el estimado (EVED) para los animales del set de entrenamiento y el de validación, y utilizando cada uno de los modelos. Ratio entre varianza genética de la susceptibilidad a la metilación y varianza residual igual a 1/5.

	TRAINING		TESTING	
	TBV-EVDG	TEV-EVED	TBV-EVDG	TEV-EVED
BLUP	0.55(0.02)	0.49(0.02)	0.22(0.07)	0.14(0.05)
GBLUP	0.73(0.02)	0.73(0.04)	0.65(0.03)	0.47(0.07)
epiBLASSO	0.72(0.02)	0.79(0.03)	0.64(0.04)	0.54(0.07)

¹Promedio de las 10 réplicas. Desviación estándar entre paréntesis.

Conclusiones

La metilación del ADN constituye un ruido externo que no es posible controlar con los modelos usados actualmente en la predicción de mérito genético. La precisión de las predicciones de mérito genético se ven perjudicadas por estas marcas epigenéticas. Las estimas de heredabilidad obtenidas hasta ahora han ignorado las “fuerzas epigenéticas” y podrían verse afectadas por esta fuente de ruido externo. Este estudio muestra que ni los modelos basados en información de pedigrí tradicional ni aquellos basados en información genómica pueden corregir el ruido que se incorpora al fenotipo cuando hay regulación epigenómica sobre los caracteres. El LASSO Bayesiano epigenómico desarrollado ad hoc para incorporar la información de la metilación del DNA no mejora las predicciones del valor genético aditivo, pero si es capaz de predecir con mayor exactitud cuál será la expresión final del fenotipo dada la información de la metilación del ADN de un determinado individuo. Esto tiene importantes implicación en la producción animal. Un análisis de las metilaciones de ADN permitirían establecer que ambientes o prácticas de manejo producen unos patrones de metilación (des)favorables a la expresión de determinados caracteres. Sería posible comprobar si existen determinadas variantes genéticas que son más o menos susceptibles a la metilación bajo determinadas circunstancias ambientales, para evitar la propagación de esos genotipos en la población, o realizar acoplamientos dirigidos en función del ambiente de producción al que se destinará la descendencia.

La interacción genotipo ambiente ha sido uno de los retos más importantes en la mejora genética animal desde hace mucho tiempo, y la información epigenética puede abrir la puerta al conocimiento de este tipo de interacciones. Es necesario avanzar en el conocimiento de la repercusión de las marcas epigenómicas sobre la expresión de los caracteres de interés en producción animal, y en el desarrollo de métodos que permitan acomodar la varianza epigenómica en los modelos para reducir el ruido generado, así como para poder detectar que marcas epigenómicas influyen sobre la expresión de los caracteres. Sería necesario evitar detectar ambientes que generen metilaciones del DNA contraproducentes en la población.

Agradecimientos

Los autores agradecen al Dr. Mario Fraga (IUOPA) sus valiosos comentarios sobre los procesos de regulación de la metilación del ADN.

Referencias bibliográficas

- Coolen MW, Statham AL, Qu W, Campbell MJ, Henders AK, et al. (2011) Impact of the Genome on the Epigenome Is Manifested in DNA Methylation Patterns of Imprinted Regions in Monozygotic and Dizygotic Twins. *PLoS ONE* 6(10): e25590. doi:10.1371/journal.pone.0025590.
- Fisher, R. A. (1918). The correlation between relatives on the supposition of mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52, 399-433.
- Fraga MF, et al. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* 102:10604–10609.
- Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G. J., Susser, E. S., Slagboom, P. E. and Lumey L. H. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *PNAS* 105, 44.
- Henderson, I.R., and Jacobsen, S.E. (2007). Epigenetic inheritance in plants. *Nature* 447, 418-424.
- Jablonka, E., Raz G. (2009). Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology* 84(2), 131-175.
- Manolio T.A. et al. (2009) Finding the missing heritability. *Nature* 461, 747-753
- Migicovsky, Z. and Kovalchuk, I. (2011). Epigenetic memory in mammals. *Front.Gene.* 2:28. doi: 10.3389/fgene.2011.00028.
- Morgan, H.D., Sutherland, H.G., Martin, D.I., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in mouse. *Nature Genet* 23, 314-318.
- Nijland, M.J., Ford, S.P. and Nathanielsz P.W. (2008). Prenatal origins of adult disease. *Curr Opin Obstet Gynecol.* 20(2), 132-8.
- Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature* 465, 721-727.
- Richards, E.J. (2006). Inherited epigenetic variation-revisiting soft inheritance. *Nat. Rev. Genet.*, 7: 395-401.
- Rosenfeld, C. S. (2010). Animal models to study environmental epigenetics. *Biol. Reprod.* 82, 473–488.