

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 000 012**

51 Int. Cl.:

G16B 20/20 (2009.01)

G16B 40/20 (2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.01.2019** **PCT/CN2019/071514**

87 Fecha y número de publicación internacional: **02.07.2020** **WO20133588**

96 Fecha de presentación y número de la solicitud europea: **14.01.2019** **E 19903054 (5)**

97 Fecha y número de publicación de la concesión europea: **16.10.2024** **EP 3905253**

54 Título: **Método rápido y estable de evaluación de los valores genómicos de cría para animales individuales**

30 Prioridad:

28.12.2018 CN 201811620927

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.02.2025

73 Titular/es:

HUAZHONG AGRICULTURAL UNIVERSITY
(50.00%)
No.1 Shizishan Street Hongshan District
Wuhan, Hubei 430070, CN y
GUANGZHOU YINGZI TECHNOLOGY CO., LTD.
(50.00%)

72 Inventor/es:

ZHAO, SHUHONG;
LIU, XIAOLEI;
YANG, XIANG;
LI, XINYUN;
ZHU, MENGJIN;
XIANG, TAO;
MA, YUNLONG;
YU, MEI;
WANG, ZHIQUAN y
YIN, LILIN

74 Agente/Representante:

GONZÁLEZ POVEDA, Sara

ES 3 000 012 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método rápido y estable de evaluación de los valores genómicos de cría para animales individuales

5 ANTECEDENTES DE LA INVENCION

1. Campo de la invención

La invención se refiere al campo técnico de la cría de animales, en particular, a un método de evaluación del valor genómico de cría rápido y estable para individuos animales.

2. Las técnicas anteriores

Con el desarrollo de una técnica de genotipado de polimorfismos de un nucleótido único (SNP por sus siglas en inglés) de alta densidad que cubre todo el genoma, la selección genómica (predicción), como herramienta de análisis estadístico del genoma, se aplica ampliamente para predecir y estimar el valor genético (valor de cría) de rasgos complejos en la mejora vegetal y la cría de animales, de manera creciente en los estudios de genética humana. La estimación de los componentes de varianza es probablemente la parte del proceso de selección genómica que más tiempo consume. Los algoritmos populares anteriores de estimación de componentes de varianza utilizados en la selección genómica, tal como EMAL, necesitan un cálculo iterativo y la complejidad computacional de cada iteración es muy elevada. Los programas anteriores de selección genómica necesitan calcular la matriz inversa de una matriz de relación de afinidad genómica y el tiempo de cálculo aumenta rápidamente con el aumento del tamaño de la muestra genotipada. El documento de ALILOO H ET AL: "Including nonadditive genetic effects in mating programs to maximize dairy farm profitability", JOURNAL OF DAIRY SCIENCE, vol. 100, n.º 2, 1 de enero de 2017 (01-01-2017), páginas 1203-1222, se refiere a la inclusión de efectos no aditivos para maximizar la rentabilidad de las explotaciones lecheras y divulga los efectos aditivos de resolución inversa y el valor del efecto dominante del marcador SNP.

SUMARIO DE LA INVENCION

La invención tiene por objeto resolver el problema técnico de proporcionar un método de evaluación del valor genómico de cría rápido y estable para individuos animales para la deficiencia existente en la técnica anterior. El algoritmo HE-AI basado en BLUP (en concreto, la mejor predicción lineal insesgada) se denomina "HIBLUP", los algoritmos de regresión de Haseman-Elston (HE) y de información promedio (AI) basados en el programa de mejor predicción lineal insesgada (HIBLUP) utilizan completamente la información sobre el pedigrí, el fenotipo y el genotipo para predecir los efectos genéticos (aditivos y dominantes) de cada animal individual, así como los valores de los efectos de cada marcador SNP, y el algoritmo más avanzado de predicción del valor genómico de cría y estimación del componente de varianza se realizan para realizar la selección genómica.

Para resolver el problema técnico mencionado anteriormente, la solución técnica adoptada por la invención de acuerdo con la reivindicación 1 es un método de evaluación del valor genómico de cría rápido y estable para individuos animales, que se caracteriza por que se adoptan los algoritmos de regresión de Haseman-Elston (HE) y de información promedio (IA) basados en el programa de mejor predicción lineal insesgada (HIBLUP) para realizar la predicción del valor genómico de cría utilizando la información sobre el fenotipo, el genotipo y el pedigrí, y el resultado final incluye el valor genético estimado de los individuos, los valores del efecto aditivo y del efecto dominante de cada individuo, y los valores de la resolución inversa de cada efecto de marcador genético utilizado en los chips de genotipado. El método de evaluación del valor genómico de cría rápido y estable para individuos animales incluye específicamente las siguientes etapas:

etapa 1: realizar la numeración de los genotipos, codificar genotipos de AA, AB y BB como 0, 1 y 2, respectivamente, construir una matriz de relación A (correlación de afinidad) y una matriz G (correlación de estado) entre individuos a partir de la información sobre el pedigrí mediante un método tabular de Henderson e información genómica mediante un método de VanRaden, respectivamente, y a continuación construir una matriz H de correlación de mezclas entre los individuos animales según la información de una matriz A y una matriz G, con la siguiente ecuación:

$$H = \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} + A_{12}A_{22}^{-1}GA_{22}^{-1}A_{21} & A_{12}A_{22}^{-1}G \\ GA_{22}^{-1}A_{21} & (1-\alpha)G + \alpha A_{22} \end{pmatrix};$$

asignar los individuos a dos grupos diferentes en función de si los individuos animales de los grupos tienen o no información de genotipado, en donde el grupo con pie de página "1" representa a los individuos que sólo tienen información sobre el pedigrí pero no tienen información sobre la tipificación genómica, el grupo con pie de página "2" representa a los individuos que tienen información tanto sobre el pedigrí como sobre la tipificación genómica,

A_{11} y A_{22} representan matrices de correlación de afinidad entre individuos dentro del grupo "1" y el grupo "2", A_{12} representa las matrices de correlación de afinidad entre individuos entre el grupo "1" y el grupo "2", A_{21} es una matriz de transposición de A_{12} , G representa la matriz de correlación de estado y α es un porcentaje de reconciliación de relaciones para combinar la matriz G y la matriz A_{22} ;

5 etapa 2: derivar la varianza genética y la varianza residual a partir de la matriz H y el valor del fenotipo utilizando el algoritmo de regresión Haseman-Elston (HE) con la siguiente ecuación:

$$E(\mathbf{y}^T \mathbf{A}_j \mathbf{y}) = \sum_{i=1}^n \text{tr}(\mathbf{A}_j \mathbf{K}_i) \sigma_i^2 + \text{tr}(\mathbf{A}_j) \sigma_{n+1}^2 ;$$

10 en donde E representa el valor esperado, $\mathbf{Y}$ representa el vector de valores fenotípicos, σ_i^2 es una i -ésima varianza genética explicada por efectos aleatorios, σ_{n+1}^2 es la varianza residual, n es el número de efectos aleatorios en un

modelo, $\mathbf{A}_j$ es una matriz simétrica no negativa, $\hat{\mathbf{A}}_j$ es un valor óptimo de estimación de $\mathbf{A}_j$, $\hat{\mathbf{A}}_j = \mathbf{H}^{-1} \mathbf{K}_j \mathbf{H}^{-1}$ y, $\mathbf{K}_i$ es la i -ésima matriz de covariables de efecto aditivo, y $\mathbf{K}_j$ es la j -ésima matriz de covariables de efecto aditivo;

15 etapa 3: establecer la varianza genética y la varianza residual a partir de la regresión Haseman-Elston (HE) como valores previos de la iteración posterior de información promedio (IA), a continuación derivar la varianza genética y la varianza residual utilizando el algoritmo de iteración de información promedio (IA) hasta el estándar convergente, y obtener los parámetros genéticos estimados, en donde los parámetros genéticos incluyen la varianza genética y la varianza residual;

20 El algoritmo de iteración de información promedio (IA) puede describirse como partes:

a. Algoritmo de Newton-Raphson:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - (\mathbf{Hes}^{(k)})^{-1} \frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}} \boldsymbol{\theta}^{(k)} ;$$

25 en donde $\boldsymbol{\theta}$ son parámetros genéticos que se deben estimar, k es el número de iteraciones, $\frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}}$ es la primera derivada de la función de máxima verosimilitud logarítmica para cada parámetro que se debe estimar, y $\mathbf{Hes}$ es una matriz hessiana, que es la segunda derivada de la función de verosimilitud logarítmica máxima para cada varianza;

30 b. Método de puntuación de Fisher, en donde la matriz inversa de $\mathbf{Hes}$ se sustituye por su matriz de expectativas $\mathbf{F}$, obteniendo

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - (\mathbf{F}^{(k)})^{-1} \frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}} \boldsymbol{\theta}^{(k)} ;$$

35 la matriz $\mathbf{AI}$ puede calcularse mediante

$$\mathbf{AI} = (-\mathbf{Hes} + \mathbf{F}) / 2 ;$$

40 y los parámetros se estiman con la siguiente ecuación:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - (\mathbf{AI}^{(k)})^{-1} \frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}} \boldsymbol{\theta}^{(k)} ;$$

45 etapa 4: utilizar los parámetros genéticos estimados en la etapa 3 para resolver una ecuación de modelo mixto utilizando el método de Henderson 3, y obtener un valor de cría estimado para cada individuo, en donde la ecuación del modelo mixto se describe como:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \lambda K^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}, \quad V(u) = \sigma_u^2 K, \quad V(e) = \sigma_e^2 I, \quad Cov(u, e') = 0,$$

$\lambda = \sigma_e^2 / \sigma_u^2$, X representa una matriz de diseño correspondiente a los efectos fijos, Z representa una matriz de diseño correspondiente a los efectos aleatorios, I representa una matriz unitaria, K^{-1} representa una matriz inversa de una matriz de relación de afinidad, $\hat{b}$ representa un vector de efectos fijos estimados, y $\hat{u}$ representa un vector de valores de cría estimados;
etapa 5: calcular los efectos aditivos de cada marcador de polimorfismo de nucleótido único (SNP) en chips de genotipado con un método de resolución inversa, en donde una ecuación de cálculo puede describirse como:

$$\hat{a} = \frac{M'K^{-1}\hat{u}}{\sum_{i=1}^m 2p_i q_i};$$

en donde $\hat{a}$ es un vector de valor de efecto aditivo de marcadores SNP, m es el número de marcadores SNP, M' es la matriz de covariables de marcadores aditivos, y p_i y q_i son la frecuencia alélica del i ésimo marcador genético SNP; y
etapa 6: cuando los genotipos de los alelos AA, AB y BB se codifican como 0, 1 y 0, respectivamente, procesar un modelo dominante utilizando el mismo procedimiento de la etapa 2 a la etapa 5, y resolver inversamente el valor del efecto dominante de cada marcador SNP.

El esquema técnico se adopta para generar los siguientes efectos beneficiosos: según el método de evaluación del valor genómico de cría rápido y estable para individuos animales, que proporciona la invención, se utiliza una estrategia combinada de algoritmos de regresión de Haseman-Elston (HE) y de información promedio (AI) para obtener eficazmente una estimación estable de los componentes de la varianza, el algoritmo HE-AI basado en BLUP (en concreto, la mejor predicción lineal insesgada) se denomina "HIBLUP", los algoritmos de regresión de Haseman-Elston (HE) y de información promedio (AI) basados en el programa de mejor predicción lineal insesgada (HIBLUP) utilizan completamente la información sobre el pedigrí, el fenotipo y el genotipo para predecir los efectos genéticos (aditivos y dominantes) de cada animal individual, así como los valores de los efectos de cada marcador SNP, y el algoritmo más avanzado de predicción del valor genómico de cría y estimación del componente de varianza se realizan para realizar la selección genómica.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La FIG. 1 muestra un diagrama de flujo de un método de evaluación del valor genómico de cría rápido y estable para individuos animales de acuerdo con realizaciones de la invención.

DESCRIPCIÓN DETALLADA DE LA REALIZACIÓN PREFERIDA

Las implementaciones específicas de la invención se describen con más detalle a continuación con referencia a los dibujos y realizaciones adjuntos. Las siguientes realizaciones tienen por objeto ilustrar la invención, en lugar de limitar el alcance de la invención.

Como se muestra en la FIG. 1, el método de la realización se describe de la siguiente manera.

Según el método de evaluación del valor genómico de cría rápido y estable para individuos animales, se adoptan los algoritmos de regresión de Haseman-Elston (HE) y de información promedio (IA) basados en el programa de mejor predicción lineal insesgada (HIBLUP) para realizar la predicción del valor genómico de cría utilizando la información sobre el fenotipo, el genotipo y el pedigrí, y el resultado final incluye el valor genético estimado de los individuos, los valores del efecto aditivo y del efecto dominante de cada individuo, y los valores de la resolución inversa de cada efecto de marcador genético utilizado en los chips de genotipado. El método de evaluación del valor genómico de cría rápido y estable para individuos animales incluye específicamente las siguientes etapas:

etapa 1: realizar la numeración de los genotipos, codificar genotipos de AA, AB y BB como 0, 1 y 2, respectivamente, construir una matriz de relación A (correlación de afinidad) y una matriz G (correlación de estado) entre individuos a partir de la información sobre el pedigrí mediante un método tabular de Henderson e información genómica mediante un método de VanRaden, respectivamente, y a continuación construir una matriz H de correlación de mezclas entre los individuos animales según la información de una matriz A y una matriz G, con la siguiente ecuación, en donde la matriz incluye información de la matriz A y de la matriz G, como se muestra en la

siguiente ecuación:

$$H = \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} + A_{12}A_{22}^{-1}GA_{22}^{-1}A_{21} & A_{12}A_{22}^{-1}G \\ GA_{22}^{-1}A_{21} & (1-\alpha)G + \alpha A_{22} \end{pmatrix};$$

asignar los individuos a dos grupos diferentes en función de si los individuos animales de los grupos tienen o no información de genotipado, en donde el grupo con pie de página "1" representa a los individuos que sólo tienen información sobre el pedigrí pero no tienen información sobre la tipificación genómica, el grupo con pie de página "2" representa a los individuos que tienen información tanto sobre el pedigrí como sobre la tipificación genómica, A_{11} y A_{22} representan matrices de correlación de afinidad entre individuos dentro del grupo "1" y el grupo "2", A_{12} representa las matrices de correlación de afinidad entre individuos entre el grupo "1" y el grupo "2", A_{21} es una matriz de transposición de A_{12} , G representa la matriz de correlación de estado y α es un porcentaje de reconciliación de relaciones para combinar la matriz G y la matriz A_{22} ;

etapa 2: derivar la varianza genética y la varianza residual a partir de la matriz H y el valor del fenotipo utilizando el algoritmo de regresión Haseman-Elston (HE) con la siguiente ecuación:

$$E(y^T A_j y) = \sum_{i=1}^n tr(A_j K_i) \sigma_i^2 + tr(A_j) \sigma_{n+1}^2;$$

en donde E representa el valor esperado, Y representa el vector de valores fenotípicos, σ_i^2 es una i -ésima varianza genética explicada por efectos aleatorios, σ_{n+1}^2 es la varianza residual, n es el número de efectos aleatorios en un

modelo, A_j es una matriz simétrica no negativa, $\hat{A}_j$ es un valor óptimo de estimación de A_j , $\hat{A}_j = H^{-1} K_j H^{-1}$ y, K_i es i -ésima matriz de covariables de efecto aditivo, y K_j es j -ésima matriz de covariables de efecto aditivo;

etapa 3: establecer la varianza genética y la varianza residual a partir de la regresión Haseman-Elston (HE) como valores previos de la iteración posterior de información promedio (IA), a continuación derivar la varianza genética y la varianza residual utilizando el algoritmo de iteración de información promedio (IA) hasta el estándar convergente, y obtener los parámetros genéticos estimados, en donde los parámetros genéticos incluyen la varianza genética y la varianza residual;

El algoritmo de iteración de información promedio (IA) puede describirse como partes:

a. Algoritmo de Newton-Raphson:

$$\theta^{(k+1)} = \theta^{(k)} - (Hes^{(k)})^{-1} \frac{\partial L}{\partial \theta} \theta^{(k)};$$

en donde θ son los parámetros genéticos que se deben estimar, k es el número de iteraciones, $\frac{\partial L}{\partial \theta}$ es la primera derivada de la función de máxima verosimilitud logarítmica para cada parámetro que se debe estimar, y Hes es una matriz hessiana, que es la segunda derivada de la función de verosimilitud logarítmica máxima para cada varianza;

b. Método de puntuación de Fisher, en donde la matriz inversa de Hes se sustituye por su matriz de expectativas F , obteniendo

$$\theta^{(k+1)} = \theta^{(k)} - (F^{(k)})^{-1} \frac{\partial L}{\partial \theta} \theta^{(k)};$$

la matriz AI puede calcularse mediante;

$$AI = (-Hes + F) / 2;$$

y los parámetros se estiman con la siguiente ecuación:

$$\theta^{(k+1)} = \theta^{(k)} - (AI^{(k)})^{-1} \frac{\partial L}{\partial \theta} \theta^{(k)} ;$$

etapa 4: utilizar los parámetros genéticos estimados en la etapa 3 para resolver una ecuación de modelo mixto utilizando el método de Henderson 3, y obtener un valor de cría estimado para cada individuo, en donde la ecuación del modelo mixto se describe como:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \lambda K^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}, \quad V(u) = \sigma_u^2 K, \quad V(e) = \sigma_e^2 I, \quad Cov(u, e') = 0,$$

$\lambda = \sigma_e^2 / \sigma_u^2$, X representa una matriz de diseño correspondiente a los efectos fijos, Z representa una matriz de diseño correspondiente a los efectos aleatorios, I representa una matriz unitaria, K^{-1} representa una matriz inversa de una matriz de relación de afinidad, $\hat{b}$ representa un vector de efectos fijos estimados, y $\hat{u}$ representa un vector de valores de cría estimados;

etapa 5: calcular los efectos aditivos de cada marcador de polimorfismo de nucleótido único (SNP) en chips de genotipado con un método de resolución inversa, en donde una ecuación de cálculo puede describirse como:

$$\hat{a} = \frac{MK^{-1}\hat{u}}{\sum_{i=1}^m 2p_i q_i};$$

en donde $\hat{a}$ es un vector de valor de efecto aditivo de marcadores SNP, m es el número de marcadores SNP, M es la matriz de covariables de marcadores aditivos, y p_i y q_i son la frecuencia alélica del i -ésimo marcador genético SNP; y

etapa 6: cuando los genotipos de los alelos AA, AB y BB se codifican como 0, 1 y 0, respectivamente, procesar un modelo dominante utilizando el mismo procedimiento de la etapa 2 a la etapa 5, y resolver inversamente el valor del efecto dominante de cada marcador SNP.

La aplicación de los algoritmos de regresión de Haseman-Elston (HE) y de información promedio (IA) basados en el programa de mejor predicción lineal insesgada (HIBLUP) en la selección genómica porcina puede servir para acortar el ciclo de cría (intervalo de tiempo), aumentar la precisión de la selección y acelerar la ganancia genética de los rasgos objeto de selección. La solicitud incluye principalmente las siguientes etapas: obtener datos sobre el genotipo, datos sobre el pedigrí y datos sobre el fenotipo; preparar los conjuntos de datos descritos anteriormente en los algoritmos de regresión de Haseman-Elston (HE) y de información promedio (AI) basados en el formato de datos de entrada del programa de mejor predicción lineal insesgada (HIBLUP); utilizar los algoritmos de regresión de Haseman-Elston (HE) y de información promedio (IA) basados en el programa de mejor predicción lineal insesgada (HIBLUP) para obtener el valor de cría estimado (VEB) de cada individuo; calcular el índice de selección utilizando los VEB de múltiples rasgos; y clasificar a los individuos según un índice de selección exhaustivo y proporcionar una lista de candidatos a la selección.

Por último, cabe señalar que las realizaciones anteriormente descritas están previstas únicamente para describir las soluciones técnicas de la invención, en lugar de limitar la invención.

REIVINDICACIONES

1. Un método de evaluación del valor genómico de cría rápido y estable para individuos animales realizado por ordenador, en donde se adoptan los algoritmos de regresión de Haseman-Elston (HE) y de información promedio (IA) basados en el programa de mejor predicción lineal insesgada (HIBLUP) para realizar la predicción del valor genómico de cría utilizando la información sobre el fenotipo, el genotipo y el pedigrí, y el resultado final incluye el valor genético estimado de los individuos, valores del efecto aditivo y del efecto dominante de cada individuo, y valores de la solución inversa de cada efecto de marcador genético utilizado en los chips de genotipado; el método incluye las siguientes etapas:

etapa 1: realizar la numeración de los genotipos, codificar genotipos de AA, AB y BB como 0, 1 y 2, respectivamente, construir una matriz de relación A (correlación de afinidad) y una matriz G (correlación de estado) entre individuos a partir de la información sobre el pedigrí mediante un método tabular de Henderson e información genómica mediante un método de VanRaden, respectivamente, y a continuación construir una matriz H de correlación de mezclas entre los individuos animales según la información de una matriz A y una matriz G, con la siguiente ecuación:

$$H = \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} + A_{12}A_{22}^{-1}GA_{22}^{-1}A_{21} & A_{12}A_{22}^{-1}G \\ GA_{22}^{-1}A_{21} & (1-\alpha)G + \alpha A_{22} \end{pmatrix};$$

asignar los individuos a dos grupos diferentes en función de si los individuos animales de los grupos tienen o no información de genotipado, en donde el grupo con pie de página "1" representa a los individuos que sólo tienen información sobre el pedigrí pero no tienen información sobre la tipificación genómica, el grupo con pie de página "2" representa a los individuos que tienen información tanto sobre el pedigrí como sobre la tipificación genómica, A_{11} y A_{22} representan matrices de correlación de afinidad entre individuos dentro del grupo "1" y el grupo "2", respectivamente, A_{12} representa las matrices de correlación de afinidad entre individuos entre el grupo "1" y el grupo "2", A_{21} es una matriz de transposición de A_{12} , G representa la matriz de correlación de estado y α es un porcentaje de reconciliación de relaciones para combinar la matriz G y la matriz A_{22} ;

etapa 2: derivar la varianza genética y la varianza residual a partir de la matriz H y el valor del fenotipo utilizando el algoritmo de regresión Haseman-Elston (HE) con la siguiente ecuación:

$$E(y^T A_j y) = \sum_{i=1}^n tr(A_j K_i) \sigma_i^2 + tr(A_j) \sigma_{n+1}^2;$$

en donde E representa el valor esperado, Y representa el vector de valores fenotípicos, σ_i^2 es una i -ésima varianza genética explicada por efectos aleatorios, σ_{n+1}^2 es la varianza residual, n es el número de efectos aleatorios en un

modelo, A_j es una matriz simétrica no negativa, $\hat{A}_j$ es un valor óptimo de estimación de A_j , $\hat{A}_j = H^{-1} K_j H^{-1}$ y K_i es la i -ésima matriz de covariables de efecto aditivo, y K_j es la j -ésima matriz de covariables de efecto aditivo;

etapa 3: establecer la varianza genética y la varianza residual a partir de la regresión Haseman-Elston (HE) como valores previos de la iteración posterior de información promedio (IA), a continuación derivar la varianza genética y la varianza residual utilizando el algoritmo de iteración de información promedio (IA) hasta el estándar convergente, y obtener los parámetros genéticos estimados, en donde los parámetros genéticos incluyen la varianza genética y la varianza residual;

etapa 4: utilizar los parámetros genéticos estimados en la etapa 3 para resolver una ecuación de modelo mixto utilizando el método de Henderson 3, y obtener un valor de cría estimado para cada individuo, en donde la ecuación del modelo mixto se describe como:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \lambda K^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}, \quad V(u) = \sigma_u^2 K, \quad V(e) = \sigma_e^2 I, \quad Cov(u, e') = 0,$$

$\lambda = \sigma_e^2 / \sigma_u^2$, X representa una matriz de diseño correspondiente a los efectos fijos, Z representa una matriz de diseño correspondiente a los efectos aleatorios, I representa una matriz unitaria, K^{-1} representa una matriz inversa

de una matriz de relación de afinidad, $\hat{\mathbf{b}}$ representa un vector de efectos fijos estimados, y $\hat{\mathbf{u}}$ representa un vector de valores de cría estimados;

etapa 5: calcular los efectos aditivos de cada marcador de polimorfismo de nucleótido único (SNP) en chips de genotipado con un método de resolución inversa, en donde una ecuación de cálculo se describe como:

$$\hat{\mathbf{a}} = \frac{\mathbf{M}'\mathbf{K}^{-1}\hat{\mathbf{u}}}{\sum_{i=1}^m 2p_i q_i};$$

en donde $\hat{\mathbf{a}}$ es un vector de valor de efecto aditivo de marcadores SNP, m es el número de marcadores SNP, $\mathbf{M}'$ es una matriz para covariables de marcadores aditivos, y p_i y q_i son la frecuencia alélica del i ésimo marcador genético SNP; y

etapa 6: cuando los genotipos de los alelos AA, AB y BB se codifican como 0, 1 y 0, respectivamente, procesar un modelo dominante utilizando el mismo procedimiento de la etapa 2 a la etapa 5, y resolver inversamente el valor del efecto dominante de cada marcador SNP,

en donde el algoritmo de iteración de información promedio (IA) en la etapa 3 se describe como partes:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - (\mathbf{Hes}^{(k)})^{-1} \frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}} \boldsymbol{\theta}^{(k)}$$

a. Algoritmo de Newton-Raphson:

$$\frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}}$$

; en donde $\boldsymbol{\theta}$ son parámetros genéticos

que se deben estimar, k es el número de iteraciones, $\frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}}$ es la primera derivada de la función de máxima verosimilitud logarítmica para cada parámetro que se debe estimar, y Hes es una matriz hessiana, que es la segunda derivada de la función de verosimilitud logarítmica máxima para cada varianza;

b. Método de puntuación de Fisher, en donde la matriz inversa de Hes se sustituye por su matriz de expectativas F, obteniendo

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - (\mathbf{F}^{(k)})^{-1} \frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}} \boldsymbol{\theta}^{(k)};$$

se calcula una matriz de información promedio (IA) mediante

$$\mathbf{AI} = (-\mathbf{Hes} + \mathbf{F}) / 2;$$

y los parámetros se estiman con la siguiente ecuación:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - (\mathbf{AI}^{(k)})^{-1} \frac{\partial \mathbf{L}}{\partial \boldsymbol{\theta}} \boldsymbol{\theta}^{(k)}.$$

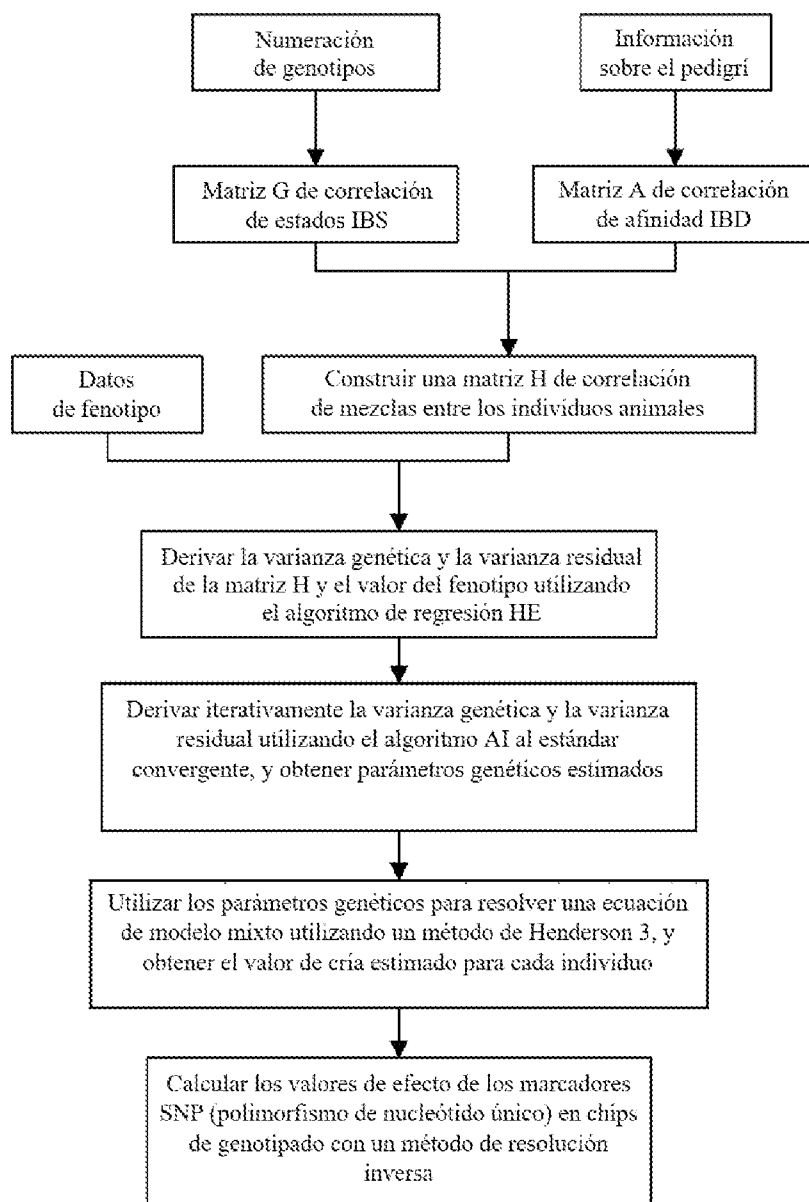


FIG. 1