



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 341 387**

(51) Int. Cl.:
G06F 19/00 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **07871975 .4**

(96) Fecha de presentación : **19.12.2007**

(97) Número de publicación de la solicitud: **2102774**

(97) Fecha de publicación de la solicitud: **23.09.2009**

(54) Título: **Métodos correctores de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial.**

(30) Prioridad: **19.12.2006 FR 06 55660**

(45) Fecha de publicación de la mención BOPI:
18.06.2010

(45) Fecha de la publicación del folleto de la patente:
18.06.2010

(73) Titular/es:
GALDERMA RESEARCH & DEVELOPMENT
Les Templiers, 2400 route des Colles
06410 Biot, FR

(72) Inventor/es: **Fogel, Paul;**
Zugaj, Didier;
Le Goff, Jean-Marc;
Aubert, Jerome y
Deret, Sophie

(74) Agente: **Ungría López, Javier**

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

Métodos correctores de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial.

La presente invención se relaciona con métodos correctores de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial. Se relaciona más particularmente con el tratamiento de tales resultados en el caso de experimentos realizados sobre chips de ADN. Los experimentos transcriptómicos tienen por objeto la identificación de genes de interés o de grupos de genes de interés.

Generalmente, el nivel de expresión de estos genes de interés o de estos grupos de genes de interés varía significativamente, por ejemplo, en respuesta a una señal. Cuando se realiza el análisis de resultados de experimentos transcriptómicos, por ejemplo por medio de chips de ADN, es habitual seleccionar los genes que presentan la mayor modulación, es decir, la mayor variación de su nivel de expresión. El nivel de esta modulación, también llamado coeficiente de modulación, se define como la razón del nivel de expresión observado en un experimento, por ejemplo en una condición llamada de tratamiento, con respecto al observado en otro experimento, por ejemplo en una condición llamada de referencia.

El examen de los resultados muestra que, cuanto más se utiliza un nivel de modulación elevado para restringir el número de genes seleccionados, más se favorece, en la selección realizada, la emergencia de genes cuyo nivel de expresión está, en la condición de referencia, próximo al límite de detección. Ahora bien, no existe ningún argumento biológico para explicar la razón por la cual los genes menos expresados en la condición de referencia serían los genes más intensamente modulados durante un tratamiento. Esta selección introduce, pues, un sesgo y lleva a ignorar genes que presentan un nivel de modulación menos elevado por el simple hecho de expresarse más en la condición de referencia.

Si se estima el coeficiente de modulación del nivel de expresión a partir de varias observaciones de un mismo gen sobre varios chips correspondientes a la misma condición, es decir, a partir de réplicas de la condición de referencia o de la condición de tratamiento, se demuestra que el coeficiente de modulación y el nivel medio de expresión de los genes evolucionan en sentido inverso [R. Mansourian *et al.*, The Global error assessment (GEA) model for the selection of differentially expressed genes in microarray data, Bioinformatics Advance Access, 2004]. En otras palabras, cuanto más bajo es el nivel de expresión de un gen en varias réplicas de una condición de referencia, más elevado es su coeficiente de modulación en respuesta al tratamiento, calculado a partir de varias réplicas. Este fenómeno se explica en parte por la presencia de un ruido de fondo de medición que muestra ser tanto más preponderante en el cálculo del coeficiente de modulación cuanto más baja sea la expresión de los genes.

El análisis diferencial según el método llamado de Global Error Assessment (GEA), divulgado en el documento al que se ha hecho antes referencia, permite corregir este sesgo. Consiste en reagrupar los genes según un criterio estadístico, llamado significatividad (o p-value, en inglés), teniendo en cuenta la variabilidad del coeficiente de modulación en función del nivel de expresión en la condición de referencia para cada gen. La variabilidad del coeficiente de modulación representa, para un gen dado, la desviación típica de los coeficientes de modulación con el coeficiente de modulación medio. Este p-value traduce la significatividad de un valor de coeficiente de modulación. Esto permite obtener grupos de genes correspondientes a un p-value dado y equilibrar, en la lista de los genes seleccionados, la proporción de los genes débilmente expresados en la condición de referencia.

Sin embargo, el p-value no tiene sentido biológico. Por ello, el biólogo, que razona en el universo de las modulaciones, no puede apoyarse en este valor para identificar los genes diferenciados. Por consiguiente, no puede utilizar el método llamado de GEA para encontrar los genes diferenciados.

En la práctica, tras uno o más análisis diferenciales, el biólogo utiliza más frecuentemente técnicas de clasificación y de visualización con el fin de identificar genes que presentan perfiles de modulación de expresión similares entre varias condiciones. Se trata, por ejemplo, de la técnica de clasificación jerárquica o la clasificación por Descomposición Robusta en Valores Singulares divulgada en el documento L. Liu *et al.*, Robust singular value decomposition analysis of microarray data, PNAS, 2003.

Sin embargo, en estas técnicas el biólogo se ve inducido, debido a las limitaciones ligadas a la visualización, o por concentrarse en análisis más complejos, tales como análisis de ontología o relativos a rutas metabólicas, a limitar el tamaño de las listas de genes seleccionados. También se apoya en los niveles de modulación de expresión medidos en cada condición, y no tiene, pues, en cuenta la significatividad asociada. Se pierde así la información relativa a esta significatividad durante la visualización de los niveles de modulación tras la clasificación. En otras palabras, el biólogo considera simplemente la razón del nivel de expresión de genes entre dos condiciones, clasificados por orden decreciente según su coeficiente de modulación. Se trata de la modulación estándar.

Deseando generalmente visualizar los genes más modulados en la condición de tratamiento, el biólogo aplica entonces una clasificación decreciente según el nivel de modulación y no conserva más que los primeros genes. Haciendo esto, no tiene en cuenta la significatividad y reintroduce el sesgo de selección que había sido eliminado por el cálculo del p-value.

Teniendo en cuenta lo que antecede, un problema que la invención se propone resolver es realizar un método corrector de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial, que tiene en cuenta la significatividad asociada a los valores de coeficiente de modulación, y cuyo resultado es además explotable a partir de valores que tienen un sentido biológico.

La solución propuesta por la invención a este problema tiene por objeto un método corrector de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial, consistente en las etapas siguientes:

- obtención de los resultados del nivel de expresión de genes en una condición de referencia y cálculo del nivel medio de expresión de cada uno de dichos genes;
- obtención de los resultados del nivel de expresión de dichos genes en una condición de tratamiento y cálculo del nivel medio de expresión para cada uno de dichos genes;
- cálculo del coeficiente de modulación del nivel de expresión para cada uno de dichos genes;
- cálculo de un p-value asociado a cada coeficiente de modulación; y
- cálculo de curvas isobáricas de p-value en función del nivel medio de expresión de cada uno de dichos genes en la condición de referencia;

caracterizado por incluir además una etapa de cálculo y de asociación de un coeficiente de modulación mediano sobre la curva isobárica de cada p-value observado.

Ventajosamente, dichas etapas de cálculo de un p-value asociado a cada coeficiente de modulación y de cálculo de un coeficiente de modulación mediano según la curva isobárica de cada p-value observado son realizadas por medio del método llamado GEA; dichas etapas de obtención de los resultados del nivel de expresión de genes en una condición(es) de referencia, de obtención de los resultados del nivel de expresión de genes en una condición(es) de tratamiento, de cálculo del coeficiente de modulación (variación) del nivel de expresión, de cálculo del p-value y de cálculo de un coeficiente de modulación mediano sobre la curva isobárica de cada p-value observado son realizadas para una pluralidad de condiciones de tratamiento diferentes; dicha etapa de cálculo de curvas isobáricas de p-value en función del nivel medio de expresión de cada uno de dichos genes en la condición de referencia comprende una representación mediante un punto de cada gen estudiado sobre un gráfico que presenta, en la abscisa, el logaritmo del nivel medio de expresión en la condición de referencia indicado como x y, en la ordenada, el logaritmo del nivel medio de expresión en la condición de tratamiento indicado como y , correspondiendo una curva isobárica de nivel p a los puntos teóricos para los cuales el p-value es igual a p ; un usuario selecciona genes de interés en base a un valor que tiene un sentido biológico; dicho valor que tiene un sentido biológico es un valor de coeficiente de modulación; un usuario selecciona genes de interés en base a un valor que tiene una significatividad; los experimentos transcriptómicos son realizados sobre chips de ADN.

La invención se relaciona además con un ordenador para la realización de un método corrector de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial según la invención.

La invención será mejor comprendida a la lectura de la descripción no limitativa que se da a continuación y de los dibujos que la acompañan, en los cuales:

- la figura 1 representa en forma de gráfico los resultados de experimentos transcriptómicos, así como las curvas isobáricas de p-value, tal como se pueden obtener según un modo de realización de la presente invención;
- las figuras 2A y 2B representan en forma de histogramas el reparto de los genes débilmente expresados, es decir, cuya expresión es inferior a 50 (2A) o inferior a 20 (2B)(unidad arbitraria ligada a la tecnología Affymetrix™), entre los genes más modulados, resultantes del análisis de experimentos transcriptómicos, por una parte, según el método de modulación estándar y, por otra, según la invención, llamado método de modulación corregido.

El método de tratamiento según la invención es un método corrector que permite el tratamiento de resultados de experimentos transcriptómicos y, en particular, el tratamiento de resultados de experimentos transcriptómicos realizados sobre chips ADN.

Los experimentos transcriptómicos realizados sobre chips de ADN son susceptibles de ser llevados a cabo de la manera siguiente.

Se cultivan células en al menos dos condiciones diferentes. Se llama a la primera condición de referencia. Sirve de control. Se llama a la segunda condición de tratamiento. En la condición de tratamiento, se cultivan las células ya sea en presencia de un agente particular, por ejemplo una proteína o un antibiótico, ya sea en condiciones experimentales particulares, por ejemplo, de luminosidad, de oxigenación, de pH o de presión.

En la práctica, se realiza cada cultivo ventajosamente varias veces para una misma condición. Se tiene entonces una pluralidad de réplicas. Por pluralidad de réplicas, se entiende preferentemente tres o cuatro cultivos realizados en una condición de referencia y dos cultivos realizados en una misma condición de tratamiento. Los resultados obtenidos a partir de varias réplicas permiten un análisis estadístico.

Independientemente para cada réplica, se lisan las células y se solubilizan los ácidos nucleicos de estas células. Se purifican entonces los ARNm mediante pase sobre una columna que contiene perlas en las que están fijados oligómeros oligo-dT. En efecto, los ARNm, que poseen una cola poliadenilada, quedan retenidos por las perlas de la columna, mientras que se eliminan los otros ácidos nucleicos.

Después de la elución, se recuperan los ARNm purificados para cada réplica. Se sintetizan entonces ADNc fluorescentes a partir de los ARNm obtenidos. Para hacerlo, se hibridan cebadores, ya sean oligo-dT, ya sean específicos de cada ARNm, ya sean aleatorios, con los ARNm purificados, y una enzima, la transcriptasa inversa, permite la síntesis de la hebra de ADNc. La utilización de nucleótidos fluorescentes permite obtener ADNc marcados. Los ARNm son degradados para no conservar ya más que las hebras simples de ADNc fluorescentes.

Los niveles de expresión de los genes para cada réplica son medidos por medio de un chip de ADN. Dicho chip lleva clásicamente varios millares de pocillos. Se fijan varios millares de ejemplares de una misma secuencia codificante en cada pocillo y esta secuencia codificante es diferente de un pocillo a otro.

Se pone el chip de ADN en presencia de la solución que lleva los ADNc fluorescentes. Estos ADNc se hibridan específicamente sobre la secuencia que les es complementaria.

Tras una o más etapas de lavado, se coloca el chip en un instrumento de detección y se escanea con ayuda de un láser que excita los fluorocromos.

Para cada pocillo, se determina entonces una intensidad de fluorescencia. Esta intensidad es proporcional a la cantidad de ADNc fluorescente fijada. Cada intensidad es proporcional al nivel de expresión del gen considerado. Por consiguiente, a partir de las intensidades de fluorescencia, se obtiene para cada gen analizado y para cada experimento un valor de nivel de expresión.

La presente invención se relaciona con un método corrector de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial de los valores de expresión de genes, cuyo fin es seleccionar los genes más modulados, es decir, los más intensamente sobreexpresados, y/o los más intensamente inhibidos, en la o las condiciones de tratamiento con respecto a la condición de referencia.

En particular, el método según la presente invención incluye una etapa de cálculo del nivel de expresión medio de cada gen analizado a partir del conjunto de las réplicas de una misma condición. Dicho método incluye luego, para cada gen, el cálculo de un coeficiente de modulación de expresión, es decir, de la razón entre su nivel de expresión medio en la condición tratada y en la condición de referencia.

En otra etapa del método según la invención, para cada gen analizado se calcula un p-value, asociado al coeficiente de modulación de dicho gen. Este valor estadístico traduce la posibilidad de observar una diferencia, al menos de un cierto nivel, entre los niveles de expresión de un gen en dos condiciones diferentes, mientras que, en realidad, estos niveles de expresión son idénticos. En otras palabras, este p-value traduce la medida en la que el valor de un coeficiente de modulación es significativo o se debe, por el contrario, a un ruido de fondo. Así, este p-value tiene en cuenta la variabilidad del coeficiente de modulación del nivel de expresión de dicho gen. Son aplicables diferentes métodos, como por ejemplo el método llamado de GEA, al presente método para determinar este p-value.

Otra etapa del método de tratamiento según la invención incluye el cálculo de curvas isobáricas del p-value en función del nivel medio de expresión en la condición de referencia. Esta etapa puede ser representada gráficamente, como se muestra en la figura 1. Para ello, se representa mediante un punto cada gen estudiado, en un gráfico que presenta, en la abscisa, el logaritmo del nivel medio de expresión en la condición de referencia indicado como x y, en la ordenada, el logaritmo del nivel medio de expresión en la condición de tratamiento indicado como y . Una curva isobárica de nivel p corresponde a los puntos teóricos para los cuales el p-value es igual a p . El proceso de construcción de las curvas isobáricas es el siguiente:

1) Se construye una cuadrícula suficientemente fina del plano xy delimitado por los valores extremos de expresión observados.

2) Se determina para cada uno de los puntos de la cuadrícula la diferencia $y-x$ entre la condición de tratamiento y la condición de referencia, que se divide por la desviación típica asociada al nivel de expresión medio $(x+y)/2$. Para estos fines, se habrá establecido previamente la relación existente entre nivel de expresión y varianza de un gen en base a las réplicas de la condición de referencia, según el método de GEA.

3) Se trazan por un método habitual de obtención de contornos (por ejemplo, referencia Matlab) las curvas isobáricas correspondientes a los triplicados $(x, y, p\text{-value}(x, y))$. Para un mismo p-value, se reúnen los genes sobreexpresados sobre una curva, mientras que se reúnen sobre otra curva los genes reprimidos.

La última etapa del método consiste, para cada curva isobárica del p-value, en asociarle la modulación mediana de los puntos contenidos en esta curva. Con estos fines, se considera el conjunto de los puntos de la cuadrícula en la abscisa, a los cuales se asocian en la ordenada los puntos correspondientes sobre la curva isobárica. A cada uno de los puntos (x,y) así obtenidos le corresponde una modulación $\exp(y-x)$. En el cálculo de la mediana propiamente
 5 dicho, cada punto (x,y) es ponderado por la densidad de genes observada en la condición de referencia alrededor del valor x.

Al ser la modulación corregida por construcción una función monótona del p-value, permite obtener las mismas clasificaciones de genes que los métodos según el estado de la técnica, y, ventajosamente, con niveles de modulación
 10 corregida mejor distribuidos. En particular, las clasificaciones no incluyen valores aberrantes provocados por expresiones que alcanzan, en una de las condiciones, un nivel muy bajo, y por lo tanto poco fiable, puesto que es inferior al umbral de detección de los chips de ADN. Así, el método según la invención permite distribuir mejor los genes débilmente expresados en la condición de referencia entre los genes más modulados.

Ventajosamente, los resultados son analizados en el universo de los niveles de modulación y las selecciones, o clasificaciones de listas, son efectuadas en el mismo universo y no en el universo de los p-value. En particular, el método según la invención tiene en cuenta la variabilidad de los coeficientes de modulación en función del nivel de expresión para cada gen, la significatividad asociada y el nivel de modulación mediano, para identificar perfiles de modulación similar, sin perder ninguna información significativa.

Aún ventajosamente, el método según la invención permite al biólogo razonar en un único y mismo espacio, conservar sus hábitos y vencer el prejuicio según el cual no es posible trabajar teniendo en cuenta el p-value y confiriendo al mismo tiempo un sentido biológico a los resultados. En efecto, el biólogo puede en adelante escoger un valor biológico, y no ya un valor estadístico, como umbral de selección de los genes de interés. En efecto, puede a partir de ahora
 25 seleccionar los genes de interés suficientemente modulados, según su juicio, directamente a partir de un valor umbral de coeficiente de modulación que tiene un sentido biológico y una significatividad asociada. Al haber reagrupado los genes el método según la invención en subconjuntos, a la vez en función del valor mediano de su coeficiente de modulación y de la significatividad de estos coeficientes de modulación, el biólogo selecciona entonces genes en función de la significatividad de su coeficiente de modulación.

Las figuras 2A y 2B representan análisis diferenciales de resultados de experimentos transcriptómicos obtenidos a partir de una condición de tratamiento con respecto a una condición de referencia (cultivo de sebocitos prepuciales de rata en presencia (condición de tratamiento) o no (condición de referencia) de un agonista de PPAR-gamma, el compuesto CD4700).

En la figura 2A, se seleccionan los genes más intensamente modulados y se representan en forma de histogramas los porcentajes de genes que tienen una intensidad de expresión inferior a 50 (unidad arbitraria ligada a la tecnología Affymetrix™) entre los primeros genes seleccionados. Estas selecciones son efectuadas, por una parte, con el método de modulación corregido según la invención y, por otra, con el método estándar del estado de la técnica. Comparando
 40 estos dos métodos, se constata que el método según la invención uniformiza la distribución de estos genes débilmente expresados. En otras palabras, aplicando el método de modulación corregido según la invención, se obtiene aproximadamente un 15% de genes que tienen una intensidad de expresión inferior a 50 durante la selección de los 10 genes más modulados o de los 100 genes más modulados. Por el contrario, el método antes citado del estado de la técnica no da resultados uniformes. Por ejemplo, con el método estándar, entre los 10 genes más modulados se encuentra un
 45 80% de genes que tienen una intensidad de expresión inferior a 50.

En la figura 2B, se seleccionan los genes más intensamente modulados y se representan en forma de histogramas los porcentajes de genes que tienen una intensidad de expresión inferior a 20 (unidad arbitraria ligada a la tecnología Affymetrix™) entre los primeros genes seleccionados, según el método de modulación corregido y según el método
 50 de modulación estándar. De forma similar a la figura 2A, se obtiene una uniformización de la distribución de los genes que tienen una intensidad de expresión inferior a 20 entre los genes más modulados seleccionados con el método de modulación corregido según la invención.

En otro ejemplo, no representado, se realizó un análisis estadístico diferencial de la expresión de genes por medio de un chip de ADN cuya referencia es RAE230A, estando estos genes modulados por un agonista PPAR. Se siguieron dos metodologías con el fin de estudiar los genes, correspondientes a los identificadores Affymetrix™ contenidos en los pocillos del chip, en función del nivel de expresión de los genes: la metodología llamada estándar, correspondiente al método llamado GEA, y la metodología llamada corregida de la invención. Se clasificaron las dos listas obtenidas en función del nivel de expresión de los genes, por orden decreciente. Los 50 primeros identificadores Affymetrix™ de
 60 cada lista fueron analizados con la aplicación web GOTM™ descrita en la publicación siguiente: *Zhang B, Schmoyer D, Kirov S, Snoddy J.* (2004), BMC Bioinformatics, 18; 5(1): 16.

Si existe un interés por una ontología relativa al “metabolismo de los lípidos”, que constituye una de las ontologías más significativamente tratadas en el plano estadístico, parece ser que, con el método estándar, se extraen 7 identificadores Affymetrix™ correspondientes a 5 genes, mientras que, con el método según la invención, se identificaron 10
 65 identificadores Affymetrix™ correspondientes a 7 genes.

ES 2 341 387 T3

Se entiende que el método corrector según la invención puede, en la práctica, ser utilizado mediante un ordenador de tipo personal equipado al menos con un microprocesador y con memorias ROM y RAM asociadas. Constituye entonces un logicial cuya ejecución es controlada por un biólogo u otro operador, para la obtención de resultados corregidos de experimentos transcriptómicos. Este logicial puede, por otra parte, ser registrado en cualquier soporte de memoria de tipo no volátil o no, tal como CD-ROM o disquetes.

10

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Método corrector llevado a cabo por ordenador de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial, consistente en las etapas siguientes:

- obtención de los resultados del nivel de expresión de genes en una condición de referencia y cálculo del nivel medio de expresión de cada uno de dichos genes;

- obtención de los resultados del nivel de expresión de dichos genes en una condición de tratamiento y cálculo del nivel medio de expresión para cada uno de dichos genes;

- cálculo del coeficiente de modulación del nivel de expresión para cada uno de dichos genes;

- cálculo de un p-value asociado a cada coeficiente de modulación;

- cálculo de curvas isobáricas de p-value en función del nivel medio de expresión de cada uno de dichos genes en la condición de referencia;

caracterizado por incluir además una etapa de cálculo y de asociación de un coeficiente de modulación mediano en la curva isobárica de cada p-value observado.

2. Método según la reivindicación 1, **caracterizado** por realizar dichas etapas de cálculo de un p-value asociado a cada coeficiente de modulación y de cálculo de un coeficiente de modulación mediano según la curva isobárica de cada p-value observado por medio del método llamado GEA.

3. Método según una de las reivindicaciones 1 ó 2, **caracterizado** por llevar a cabo dichas etapas de obtención de los resultados del nivel de expresión de genes en una condición de referencia, de obtención de los resultados del nivel de expresión de genes en una condición de tratamiento, de cálculo del coeficiente de modulación del nivel de expresión, de cálculo del p-value y de cálculo de un coeficiente de modulación mediano en la curva isobárica de cada p-value observado para una pluralidad de condiciones de tratamiento diferentes.

4. Método según una de las reivindicaciones 1, 2 ó 3, **caracterizado** por el hecho de que dicha etapa de cálculo de curvas isobáricas de p-value en función del nivel medio de expresión de cada uno de dichos genes en la condición de referencia comprende una representación mediante un punto de cada gen estudiado en un gráfico que presenta, en la abscisa, el logaritmo del nivel medio de expresión en la condición de referencia indicado como x y, en la ordenada, el logaritmo del nivel medio de expresión en la condición de tratamiento indicado como y , correspondiendo una curva isobárica de nivel p a los puntos teóricos para los cuales el p-value es igual a p .

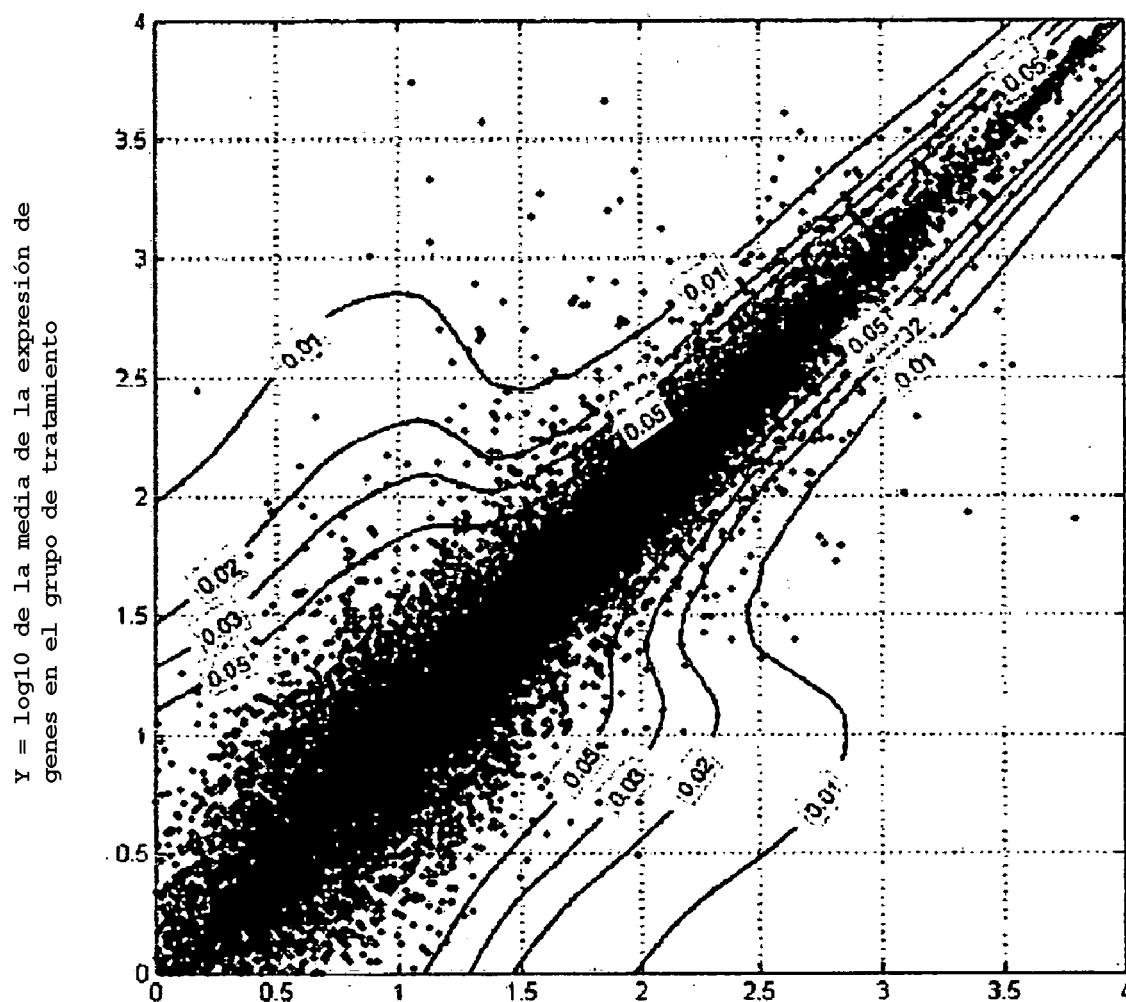
5. Método según una de las reivindicaciones 1 a 4, **caracterizado** por el hecho de que un usuario selecciona genes de interés en base a un valor que tiene un sentido biológico.

6. Método según la reivindicación 5, **caracterizado** por el hecho de que dicho valor que tiene un sentido biológico es un valor de coeficiente de modulación.

7. Método según una de las reivindicaciones precedentes, **caracterizado** por el hecho de que un usuario selecciona genes de interés en base a un valor que tiene una significatividad.

8. Método según una de las reivindicaciones precedentes, **caracterizado** por realizar los experimentos transcriptómicos sobre chips de ADN.

9. Ordenador para la realización de un método corrector de tratamiento de resultados de experimentos transcriptómicos obtenidos por análisis diferencial según una de las reivindicaciones precedentes.



X = log10 de la media de la expresión de
genes en el grupo de referencia

FIG. 1

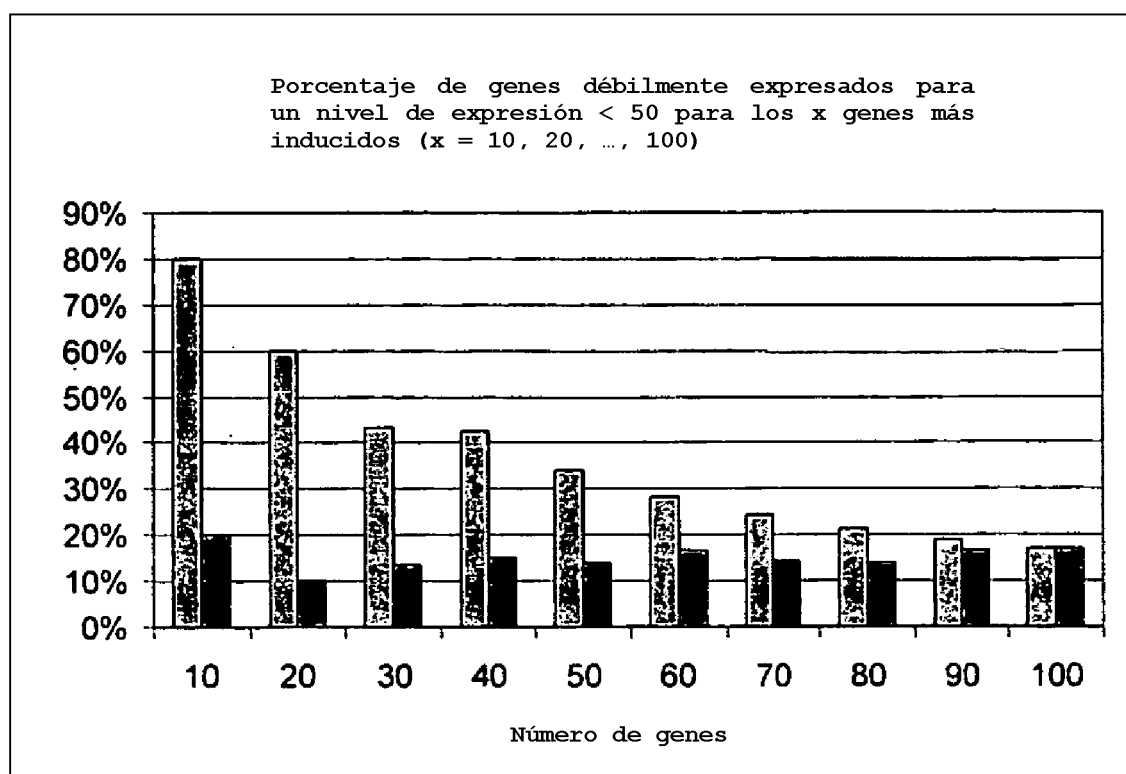
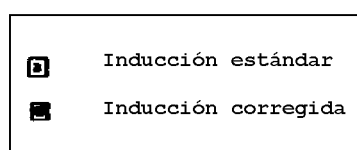


FIG. 2A



Porcentaje de genes débilmente expresados para un nivel de expresión < 20 para los x genes más inducidos ($x = 10, 20, \dots, 100$)

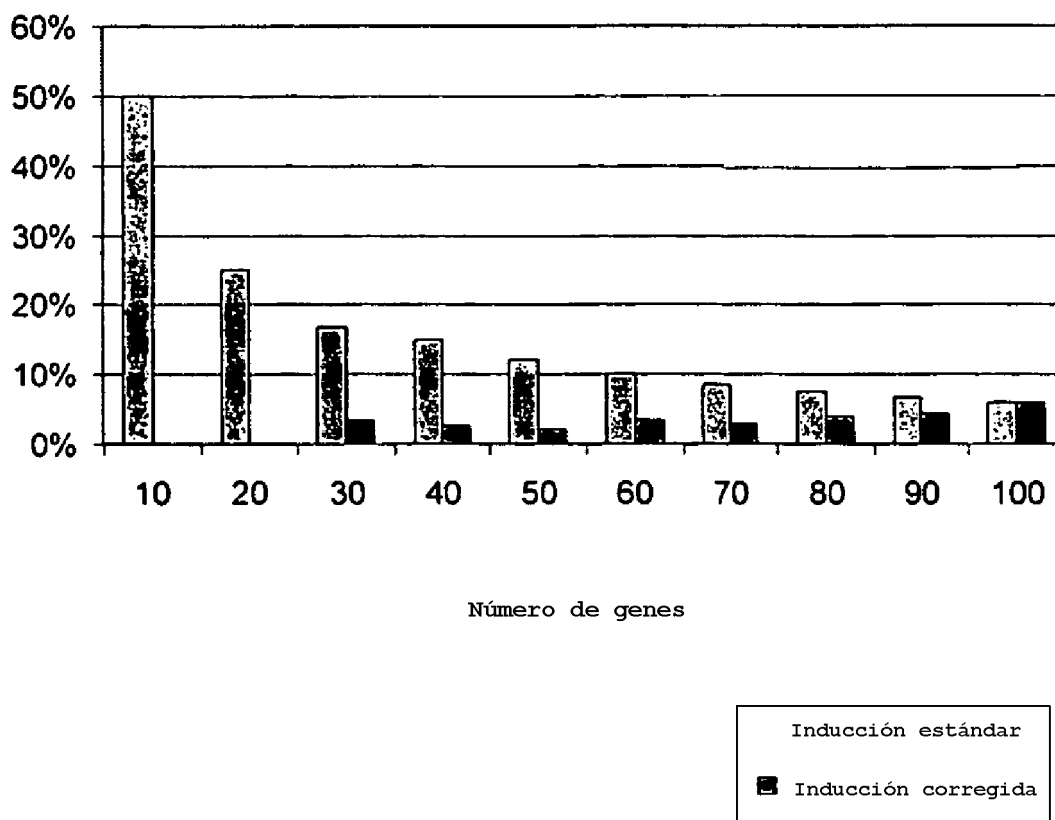


FIG. 2B