

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 765 748**

51 Int. Cl.:

G06T 7/00 (2007.01)

G06T 7/11 (2007.01)

G06T 7/136 (2007.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.12.2016 PCT/EP2016/080507**

87 Fecha y número de publicación internacional: **15.06.2017 WO17098010**

96 Fecha de presentación y número de la solicitud europea: **09.12.2016 E 16822919 (3)**

97 Fecha y número de publicación de la concesión europea: **16.10.2019 EP 3387616**

54 Título: **Clasificación de objetos en imágenes digitales**

30 Prioridad:

10.12.2015 EP 15199401

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

10.06.2020

73 Titular/es:

**QIAGEN GMBH (100.0%)
Qiagen Str. 1
40724 Hilden, DE**

72 Inventor/es:

**UNTERER, THOMAS y
SCHUSTER, GUIDO**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 765 748 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

Clasificación de objetos en imágenes digitales

La presente invención se refiere a sistemas y métodos para distinguir al menos un primer objeto de al menos un segundo objeto en una pluralidad de imágenes digitales. Más particularmente, la presente invención se refiere a la clasificación de objetos vivos y distinguir los mismos de objetos en blanco, preferiblemente para secuenciación de DNA.

La biotecnología, la medicina y campos técnicos afines están basados en el análisis de moléculas. Dispositivos electrónicos pueden analizar las moléculas con precisión y especificidad altas. Especialmente en los últimos años, se han desarrollado dispositivos electrónicos automáticos para analizar grandes números de muestras por métodos rutinarios. Por ejemplo, se utilizan aparatos modernos de secuenciación de DNA para análisis rutinario de grandes números de sondas de DNA. Pueden analizarse muestras de proteínas por escrutinio de alta capacidad y métodos afines. Frecuentemente, tales dispositivos electrónicos detectan señales fluorescentes emitidas por las sondas de muestra. Esto es posible cuando las moléculas, tales como ácidos nucleicos o proteínas, se están marcadas con compuestos fluorescentes tales como colorantes.

Los aparatos de secuenciación disponibles comercialmente son capaces de secuenciar grandes números de muestras marcadas con colorantes fluorescentes en paralelo. Métodos desarrollados recientemente, a los que se hace referencia como "secuenciación de nueva generación", NGS, han revolucionado la secuenciación. La NGS permite la secuenciación paralela en gran escala de moléculas de DNA amplificadas por clonación o individuales que están separadas espacialmente en una celda de flujo o por generación de una emulsión aceite-agua. La NGS hace posible que se realicen simultáneamente miles o incluso millones a miles de millones de reacciones de secuenciación.

En la NGS, la secuenciación se realiza por ciclos repetidos de extensiones de nucleótidos mediadas por polimerasa o, en un formato, por ciclos iterativos de ligación de oligonucleótidos. Como un proceso paralelo en gran escala, la NGS genera cientos de megabases a gigabases de capacidad de secuenciación de nucleótidos en una sola operación del instrumento, dependiendo de la plataforma. La producción económica de grandes volúmenes de datos de secuencia es la ventaja principal sobre los métodos convencionales.

Las plataformas NGS y aplicaciones/campos comunes para las tecnologías NGS se revisan por ejemplo en Voelkerding et al, *Clinical Chemistry* 55:4 641-658, 2009 y Metzker, *Nature Reviews/Genetics* Volume 11, enero 2010, páginas 31-46.

En la NGS, diversos oligonucleótidos de interés se fijan covalentemente a un soporte. A continuación, un nucleótido marcado con un colorante fluorescente se fija a la cadena de oligonucleótidos en crecimiento con DNA-polimerasa. Cuando los cuatro nucleótidos están marcados con colorantes fluorescentes diferentes, las señales de fluorescencia emitidas por una sonda pueden detectarse y se puede identificar el tipo de nucleótido fijado al oligonucleótido. Después de la detección, el colorante fluorescente se escinde y se lleva a cabo el ciclo de síntesis siguiente, en el cual se fija un nuevo nucleótido marcado a la cadena en crecimiento. Por realización de ciclos múltiples, se puede determinar de manera gradual la secuencia de una cadena de oligonucleótidos en crecimiento. Los pasos de trabajo se llevan a cabo en un aparato de secuenciación automático.

Los documentos US 2010/0323350 A1 y WO 2009/117119 A1 se refieren a métodos y composiciones para determinación de la identidad de los ácidos nucleicos en secuencias de nucleótidos utilizando, por ejemplo, datos obtenidos de secuenciación por métodos de síntesis.

El documento WO 2008/097455 A1 se refiere a un sistema de obtención de imágenes para excitación y medición de la fluorescencia sobre o en muestras que comprenden materiales fluorescentes, por ejemplo marcadores, colorantes o pigmentos fluorescentes, en particular para detectar marcadores fluorescentes en ácido nucleico. Además, se describe un dispositivo que está configurado de tal manera que se detectan simultáneamente marcadores fluorescentes en una pluralidad de plantillas de DNA diferentes.

El documento WO 2014/020137 A1 se refiere a un método para enriquecimiento de secuencias diana de una biblioteca de secuenciación a fin de proporcionar una biblioteca de secuenciación enriquecida en dianas, en donde la biblioteca de secuenciación es adecuada para para secuenciación paralela en gran escala y comprende una pluralidad de moléculas de ácido nucleico bicatenario.

Lamarre, Claude et al.: "Transcriptomic analysis of the exit from dormancy of *Aspergillus fumigatus* conidia", *BMC Genomics*, 16, septiembre 2008, página 417, se refiere a datos transcriptómicos que indican que la salida del estado latente está asociada con un cambio de un metabolismo fermentativo a un metabolismo respiratorio, así como una tendencia hacia la síntesis inmediata de proteínas. Como parte del análisis de los datos, el documento describe una estimación de un umbral para identificación de puntos vacíos en imágenes de microrredes con sondas marcadas.

El documento WO 2004/045545 A2 se refiere a métodos para tratamiento del cáncer, métodos para escrutinio de compuestos que tienen actividad anticáncer, y métodos de evaluación de la progresión del cáncer. Como parte del

análisis de los datos, el documento describe una estimación de un umbral para identificación de puntos blancos, vacíos o pobres en imágenes de sondas marcadas.

5 El documento US 2010/285577 A1 se refiere a identificación, aislamiento, y diferenciación de poblaciones de células madre de líneas germinales distintas, y linajes de células generados a partir de ellas, con diferentes usos potenciales en terapia de reemplazamiento de células.

Las señales fluorescentes emitidas por sondas de muestra con moléculas marcadas son débiles, pero las señales tienen que ser detectadas con precisión y especificidad altas. Así, para tales procesos se requiere equipo óptico preciso, especialmente cámaras y tecnología de escaneo.

10 Adicionalmente, es necesaria una evaluación extensa de las imágenes digitales capturadas por un sistema óptico de obtención de imágenes del aparato de secuenciación a fin de obtener un resultado de la secuenciación preciso y fiable, por ejemplo en FASTQ.

15 En la secuenciación de DNA/RNA, las posiciones cuenta/objeto se determinan típicamente en una imagen que se conoce como iluminada con luz blanca, es decir en un rango espectral amplio. Por una parte, este procedimiento es lo suficientemente robusto para determinar las posiciones de las cuentas, pero, por otra parte, este método no permite distinguir las cuentas que contienen abundante DNA, denominadas cuentas vivas, de las cuentas en blanco que no contienen DNA alguno. Así pues, el objeto de la presente invención es la determinación de las cuentas vivas a fin de, por ejemplo, permitir la evaluación del enriquecimiento, es decir la ratio entre cuentas en blanco y cuentas vivas de una operación.

20 Es un objeto de la invención proporcionar un método y un sistema, que resuelve los problemas de la técnica anterior arriba mencionados. En particular, un objeto es distinguir cuentas vivas de cuentas en blanco.

Este objeto se resuelve con las características de las reivindicaciones independientes. Las reivindicaciones dependientes se refieren a aspectos adicionales de la invención.

El método conforme a la presente invención está implementado por computadora. Sin embargo, las personas expertas entenderán que existen también otras vías de implementación del método conforme a la presente invención.

25 En un aspecto de la presente invención, se proporciona un método para distinguir al menos un primer objeto de al menos un segundo objeto en una pluralidad de imágenes digitales. El al menos un primer objeto ha recibido al menos una molécula que comprende información genética, y el al menos un segundo objeto no ha recibido una molécula que comprende información genética. La al menos una molécula está configurada para recibir uno de una pluralidad de compuestos fluorescentes en cada uno de una pluralidad de ciclos. Las imágenes digitales se determinan por un sistema óptico de obtención de imágenes durante la emisión de radiación electromagnética por los compuestos fluorescentes, en donde la pluralidad de imágenes digitales comprende una pluralidad de series de imágenes, refiriéndose cada imagen de una serie al espectro de emisión de un compuesto fluorescente respectivo y en donde la serie de imágenes se toma repetidamente para cada uno de la pluralidad de ciclos. El método comprende los pasos siguientes, preferiblemente en el orden siguiente:

35 a) determinación de los valores de intensidad a partir de la pluralidad de imágenes digitales para cada objeto;
b) corrección de los efectos de autofluorescencia por sustracción de un valor de intensidad predeterminado a partir de los valores de intensidad determinados;

40 c) corrección de los efectos de intermodulación causados por la superposición de los espectros fluorescentes de los compuestos fluorescentes por aplicación de una matriz de intermodulación predeterminada a los valores de intensidad determinados;

d) computación de los valores de desviación estándar de los valores de intensidad corregidos para cada objeto y cada uno de la pluralidad de ciclos y computación de un valor medio de las desviaciones estándar para cada objeto;

e) normalización de los valores medios para cada objeto utilizando una transformación lineal;

45 f) determinación de los umbrales de clasificación a partir de los valores medios normalizados utilizando un método de selección de umbrales para obtener una imagen digital binaria, en donde el al menos un primer objeto se representa por un primer valor de intensidad por encima del umbral de clasificación correspondiente y el al menos un segundo objeto se representa por un segundo valor de intensidad inferior al umbral de clasificación correspondiente.

50 Preferiblemente, el método se aplica a lo largo de un intervalo de ciclos definido y el intervalo se define por un ciclo inicial y final. Se prefiere que este intervalo comience después de los "códigos de barras" (sexto ciclo) y termine en el ciclo vigesimoséptimo (recuento basado en cero).

Las posiciones del objeto se determinan preferiblemente a partir de una imagen de luz blanca correspondiente antes de la ejecución del método de la presente invención. Es decir, las posiciones del objeto están predeterminadas y son conocidas por el sistema y método conforme a la presente invención

- 5 El al menos un primer objeto es preferiblemente una cuenta viva que ha recibido al menos una cadena de DNA/RNA y el al menos un segundo objeto es preferiblemente una cuenta en blanco que no ha recibido una cadena de DNA/RNA. La cadena de RNA/RNA es preferiblemente un fragmento de DNA/RNA monocatenario. Tales cuentas están disponibles comercialmente por ejemplo de Thermo Fisher Scientific Inc., por ejemplo bajo la marca Dynabeads, véase la referencia [1].
- Preferiblemente, cada cadena de DNA está configurada para recibir moléculas, preferiblemente nucleótidos. Los nucleótidos comprenden preferiblemente un compuesto fluorescente escindible y una cápsula escindible. Adicionalmente, se prefiere que las moléculas sean una pluralidad de cuatro moléculas diferentes, cada una de las cuales tiene un compuesto fluorescente específico para cada una de las cuatro moléculas diferentes.
- 10 Cada uno de la pluralidad de ciclos se caracteriza preferiblemente por que una de las cuatro moléculas está fijada a una de las cadenas de DNA, en donde la cápsula escindible asegura que solamente una molécula está fijada a una cadena de DNA durante un ciclo.
- Se prefiere que para cada ciclo se tomen una pluralidad de imágenes, preferiblemente cuatro, una para cada uno de los cuatro compuestos fluorescentes fijados a las moléculas. Adicionalmente, se prefiere que los cuatro compuestos fluorescentes tengan uno de un espectro de emisión principal en el régimen espectral rojo, azul, amarillo o verde.
- 15 Después de cada ciclo, la cápsula escindible se retira, a fin de que pueda fijarse otra molécula a la cadena de DNA en el ciclo siguiente.
- El sistema óptico de obtención de imágenes comprende preferiblemente un dispositivo de iluminación, por ejemplo un LED, para activar/estimular el o los compuestos fluorescentes a fin de excitar la radiación electromagnética y un sistema de detección, por ejemplo un detector CCD, para detectar los espectros de emisión (radiación electromagnética) del o los compuestos fluorescentes.
- 20 El método comprende además preferiblemente antes del paso a): determinar los valores de intensidad mínimos dentro de una pluralidad de áreas predeterminadas de cada una de las imágenes digitales y corregir la iluminación causada por el sistema óptico de obtención de imágenes para cada objeto en cada una de las imágenes digitales mediante división de los valores de intensidad determinados por los valores de intensidad mínimos correspondientes.
- 25 Preferiblemente, el método comprende además aplicar un filtro de mediana a los valores de intensidad mínimos determinados de las áreas predeterminadas y/o aplicar un filtro Gaussiano a los valores de intensidad mínimos determinados de las áreas predeterminadas, preferiblemente después de aplicar el filtro de mediana.
- 30 Las áreas predeterminadas son preferiblemente de un número menor que el número de píxeles de la imagen original, es decir menor que el tamaño de la imagen original.
- El método comprende además preferiblemente interpolar los valores de intensidad mínimos determinados de las áreas predeterminadas al tamaño de la imagen digital original, preferiblemente por el uso de una interpolación bicúbica. Sin embargo, cualquier método de interpolación adecuado podría aplicarse preferiblemente a fin de determinar los valores de intensidad mínimos de las áreas predeterminadas al tamaño de la imagen digital original.
- 35 El valor de intensidad predeterminado para corregir los efectos de autofluorescencia se determina preferiblemente tomando un valor de intensidad mínimo predeterminado a partir de las intensidades determinadas, preferiblemente tomando el valor de intensidad mínimo 10% a partir de las intensidades determinadas.
- Preferiblemente, la matriz de intermodulación tiene en cuenta una superposición de un espectro fluorescente verde con un espectro fluorescente amarillo de los compuestos fluorescentes, preferiblemente por aplicación de un factor de 0,4. Es decir, la matriz de intermodulación comprende preferiblemente un solo factor, a saber, para corregir la superposición de los espectros verde y amarillo. Entretanto, las otras entradas de la matriz de intermodulación son preferiblemente cero, es decir no se realiza corrección alguna para los otros espectros.
- 40 La normalización de los valores medios comprende preferiblemente subdividir los valores medios determinados en una pluralidad de grupos y determinar los valores medios mínimos y valores medios máximos para cada grupo.
- 45 El método comprende además preferiblemente aplicar una estrategia de crecimiento regional a los valores mínimos y los valores máximos y/o aplicar un filtro de mediana a los valores mínimos y los valores máximos y/o aplicar un filtro Gaussiano a los valores mínimos y los valores máximos.
- Los valores conforme a la presente invención se representan preferiblemente en una matriz correspondiente. Si uno de los elementos de la matriz está vacío, se aplica una estrategia de crecimiento regional, por ejemplo a los valores mínimos (matriz de mínimos) y/o a los valores máximos (matriz de máximos), dado que se prefiere evitar elementos vacíos en las respectivas matrices.
- 50 El método comprende preferiblemente interpolar los valores mínimos y valores máximos determinados al tamaño de la imagen digital original, preferiblemente por el uso de una interpolación bicúbica. Sin embargo, puede utilizarse de modo igualmente satisfactorio cualquier método de interpolación.

Preferiblemente se realiza una transformación lineal para cada objeto dividiendo las diferencias entre los valores medios y los valores mínimos correspondientes por las diferencias correspondientes entre los valores máximos y los valores mínimos correspondientes.

5 El método comprende además preferiblemente subdividir los valores medios normalizados en un número predeterminado de subgrupos y determinar los valores medios correspondientes de los objetos para cada subgrupo antes de determinar los umbrales de clasificación para cada subgrupo, y preferiblemente aplicar a continuación una estrategia de crecimiento regional y/o un filtro de mediana y/o un filtro Gaussiano.

El método de selección de umbrales comprende preferiblemente el método de segmentación de Otsu, aplicado preferiblemente a los valores medios determinados de los subgrupos.

10 Además, el método comprende preferiblemente interpolar los umbrales determinados por el método de selección de umbrales al tamaño de la imagen digital original, preferiblemente por el uso de una interpolación bicúbica. La determinación de umbrales de clasificación a partir de los valores medios normalizados utilizando un método de selección de umbrales para obtener una imagen digital binaria se utiliza para clasificar los dos objetos diferentes, es decir, los objetos primero y segundo. A saber, en la imagen binaria, el al menos un primer objeto se representa por un primer valor de intensidad superior al umbral de clasificación correspondiente y el al menos un segundo objeto se representa por un segundo valor de intensidad inferior al umbral de clasificación correspondiente.

15 La invención se refiere también a un sistema para distinguir al menos un primer objeto de al menos un segundo objeto en una pluralidad de imágenes digitales, habiendo recibido el al menos un primer objeto al menos una molécula que comprende información genética, y no habiendo recibido el al menos un segundo objeto una molécula que comprende información genética. La al menos una molécula está configurada para recibir uno de una pluralidad de compuestos fluorescentes en cada uno de una pluralidad de ciclos. Las imágenes digitales se determinan por un sistema óptico de obtención de imágenes durante la emisión de radiación electromagnética por los compuestos fluorescentes, en donde la pluralidad de imágenes digitales comprende una pluralidad de series de imágenes, refiriéndose cada imagen de una serie al espectro de emisión de un compuesto fluorescente respectivo, y en donde las imágenes se toman repetidamente para cada uno de la pluralidad de ciclos. Se prefiere que el sistema esté configurado para llevar a cabo los pasos del método reseñado anteriormente. El sistema comprende una unidad de determinación de intensidad configurada para obtener valores de intensidad a partir de la pluralidad de imágenes digitales para cada objeto, una unidad de corrección configurada para corregir los efectos de autofluorescencia por sustracción de un valor de intensidad predeterminado de los valores de intensidad determinados y para corregir los efectos de intermodulación causados por la superposición de los espectros fluorescentes por aplicación de una matriz de intermodulación predeterminada a los valores de intensidad determinados, y una unidad de computación configurada para computar los valores de desviación estándar de los valores de intensidad corregidos para cada objeto y cada uno de la pluralidad de ciclos y computación de los valores medios de las desviaciones estándar para cada objeto y para computar los valores medios normalizados para cada objeto utilizando una transformación lineal y computar los umbrales de clasificación utilizando un método de selección de umbrales para obtener una imagen digital binaria, en donde el al menos un primer objeto se representa por un primer valor de intensidad superior al umbral de clasificación correspondiente y el al menos un segundo objeto se representa por un segundo valor de intensidad inferior al umbral de clasificación correspondiente.

20 En otro aspecto adicional de la presente invención, un producto de programa de computadora que comprende uno o más medios legibles por computadora que tienen instrucciones ejecutables por computadora para realización de los pasos del método reseñado anteriormente.

25 Así pues, la presente invención proporciona una solución que se origina por la técnica anterior y a los riesgos técnicos que se originan por la utilización del método de selección de umbrales, por ejemplo el método Otsu, véase la referencia [2], que impiden una identificación fiable y robusta de un umbral de clasificación, por ejemplo brillo/saturación de imágenes heterogéneas causado por un efecto de degradación debido al sistema óptico y a los sistemas de iluminación heterogénea e intensidad diferente y características de autofluorescencia entre canales.

A continuación, se describen algunas realizaciones preferidas con referencia a los dibujos. Para propósitos de explicación, se exponen varios detalles específicos, sin desviarse del alcance de la presente invención conforme a las reivindicaciones.

30 La Figura 1 ilustra un sistema conforme a una realización de la presente invención.

La Figura 2a ilustra las imágenes tomadas en ciclos diferentes conforme a una realización de la presente invención.

La Figura 2b ilustra una señal ideal de los diferentes canales.

La Figura 3 ilustra efectos de autofluorescencia.

La Figura 4 ilustra efectos de intermodulación entre los diferentes canales.

La Figura 5 ilustra la computación de los valores de desviación estándar para cada ciclo y un valor medio correspondiente para los ciclos conforme a una realización de la presente invención.

La Figura 6 ilustra un método de selección de umbrales conforme a una realización de la presente invención.

5 La Figura 1 ilustra un sistema conforme a la presente invención. El sistema comprende una unidad de determinación 100, una unidad de corrección 200 y una unidad de computación 300. Todas estas unidades están configuradas para ejecutar uno o más de los pasos de la presente invención. Si bien la presente invención se describe utilizando tres unidades independientes 100, 200, 300, es evidente que las unidades independientes pueden formar parte también de una sola unidad con tal que se ejecuten los pasos de la presente invención.

10 En una realización preferida de la presente invención, la unidad de determinación de intensidad 100 está configurada para obtener valores de intensidad a partir de la pluralidad de imágenes digitales para cada objeto. También es posible que la unidad de determinación de la intensidad 100 esté configurada para determinar todos los valores de intensidad de la pluralidad de imágenes. Sin embargo, esto podría conducir a tiempos de procesamiento más largos.

15 La unidad de determinación de la intensidad 100 podría obtener las posiciones de los objetos a partir de una imagen de luz blanca determinada de antemano. Así, se supone que las posiciones de los objetos dentro de las imágenes son conocidas para extraer únicamente los valores de intensidad para cada objeto.

La unidad de corrección 200 está configurada para corregir los efectos de autofluorescencia por sustracción de un valor de intensidad predeterminado de los valores de intensidad determinados y corregir los efectos de intermodulación causados por la superposición de espectros fluorescentes por aplicación de una matriz de intermodulación predeterminada a los valores de intensidad determinados.

20 La unidad de computación 300 está configurada para computar los valores de desviación estándar de los valores de intensidad corregidos para cada objeto y cada uno de la pluralidad de ciclos y computar los valores medios de las desviaciones estándar para cada objeto y computar los valores medios normalizados para cada objeto utilizando una transformación lineal y computar los umbrales de clasificación utilizando un método de selección de umbrales para obtener una imagen digital binaria, en donde el al menos un primer objeto se representa por un primer valor de
25 intensidad superior al umbral de clasificación correspondiente y el al menos un segundo objeto se representa por un segundo valor de intensidad inferior al umbral de clasificación correspondiente.

30 La funcionalidad de la unidad de determinación de la intensidad 100, la unidad de corrección 200 y la unidad de computación 300 se describirá adicionalmente en términos de pasos del método en las realizaciones preferidas siguientes de la presente invención. Es evidente para una persona experta en la técnica que la descripción que sigue de los pasos del método da lugar a las funciones correspondientes de la unidad de determinación de intensidad 100, la unidad de corrección 200 y la unidad de computación 300.

35 La Figura 2a ilustra las imágenes 11-14, 21-24, 31-34, 41-44, 51-54 tomadas en una pluralidad de ciclos 10-50. Esta ilustración no debería interpretarse como limitante en cuanto a la cantidad de imágenes tomadas en un ciclo correspondiente o las cantidades de ciclos. Como puede verse en la Figura 2a, en cada uno de los ciclos 10-50 en este ejemplo se toman, es decir se adquieren, se capturan, etc. cuatro imágenes 11-14, 21-24, 31-34, 41-44, 51-54. En particular, cada una de las cuatro imágenes 11-14, 21-24, 31-34, 41-44, 51-54 en un ciclo 10-50 corresponde a un canal del sistema óptico de obtención de imágenes, a saber, rojo, verde, amarillo y azul. Los diferentes colores son emitidos por compuestos fluorescentes transportados por moléculas diferentes que son recibidas por cadenas de DNA fijadas a los objetos (cuentas). De modo más particular, cada uno de los diferentes compuestos fluorescentes
40 representa una de una base de DNA específica, a saber, timina (T), adenina (A), citosina (C), y guanina (G). Por ejemplo, los compuestos fluorescentes están asociados a las bases de DNA como sigue: T = verde; A= amarillo; C = azul; y G = rojo.

45 En cada ciclo 10-50 las primeras imágenes 11, 21, 31, 41, 51 corresponden a uno de los cuatro canales T, A, C, G, por ejemplo G. Las segundas imágenes 12, 22, 32, 42, 52 corresponden entonces a un segundo de los tres canales restantes T, A, C, por ejemplo C. Las terceras imágenes 13, 23, 33, 43, 53 corresponden luego a un tercero de los dos canales restantes T, A, por ejemplo, A. Las cuartas imágenes 14, 24, 34, 44, 54 corresponden entonces al cuarto canal restante, por ejemplo T.

50 La Figura 2b muestra la potencia de señal determinada a partir de cuatro ciclos para una sola cuenta. Cada uno de los ciclos comprende cuatro imágenes, en donde la Figura 2b ilustra la potencia de señal para los canales respectivos G, C, A, T, es decir, la llamada base para dicho ciclo específico. Esto es, en un caso ideal cada canal proporcionaría una única señal por canal. Sin embargo, debido a diferentes efectos parásitos, como los efectos de autofluorescencia y los efectos de intermodulación entre diferentes canales, las señales para los diferentes canales serán muy probablemente diferentes del caso ideal.

En lo que sigue, se reseña una realización ilustrativa del método conforme a la presente invención.

En primer lugar, se hace una corrección de las intensidades extraídas para cada ciclo 10-50. Este paso es un paso previo al proceso para la computación próxima de la característica de clasificación. Este paso consiste principalmente en corrección de la iluminación, corrección de autofluorescencia, y corrección de la intermodulación.

5 La corrección de la iluminación en cada imagen de fluorescencia 11-14, 21-24, 31-34, 41-44, 51-54 tiene por objeto compensar la iluminación irregular (equilibración de la iluminación) y comprende los pasos de subdivisión de las imágenes de fluorescencia correspondientes 11-14, 21-24, 31-34, 41-44, 51-54 en varios sub-bloques y determinación del valor de intensidad mínimo de cada sub-bloque (áreas predeterminadas). La subdivisión de la imagen, típicamente 128² píxeles por sub-bloque, se realiza para hacer frente a los efectos de degradación por un procesamiento más local. A continuación, a fin de eliminar los posibles valores extremos en la matriz de mínimos resultante se aplica una filtración de mediana (3 x 3) (eliminación de valores aberrantes, por ejemplo píxeles muertos, etc.). Adicionalmente, para alisar la superficie de la matriz de mínimos se aplica un filtro Gaussiano a la matriz resultante. El alisado se aplica a fin de prevenir/reducir discontinuidades en la imagen ampliada.

10 El objetivo es obtener un valor mínimo correspondiente para cada cuenta (posición de cuenta). Para ello, la matriz de mínimos, por ejemplo 26x20, se amplía por interpolación bicúbica al tamaño de imagen típico de por ejemplo 3296 x 2472 píxeles.

15 Para cada cuenta (ignorando las denominadas cuentas de borde y cuentas perdidas), la corrección de iluminación se realiza dividiendo la intensidad extraída por el valor mínimo correspondiente del paso previo.

Los pasos anteriores para la corrección de iluminación no son esenciales para la presente invención y pueden omitirse, por ejemplo si la iluminación es esencialmente uniforme en la imagen respectiva.

20 A continuación, se hace una sustracción de la autofluorescencia a partir de la intensidad de iluminación corregida. En la Figura 3 se ilustran niveles de intensidad típicos para cuentas vivas y cuentas en blanco con la presencia de autofluorescencia. Es decir, las intensidades de iluminación corregidas contienen todavía la intensidad de autofluorescencia de las cuentas. Debido a las diferentes características de autofluorescencia entre los canales, la autofluorescencia tiene que eliminarse a fin de permitir la computación de características de clasificación fiables.

25 Con objeto de tener en cuenta la señal de autofluorescencia parásita, el valor de autofluorescencia se determina tomando un valor predeterminado, por ejemplo tomando el 10% del valor mínimo de intensidad de iluminación corregido. La autofluorescencia se extrae típicamente de cuentas en blanco, que no contienen DNA. Dado que es evidente que no se dispone de resultado alguno de clasificación blancos/vivos hasta que ha finalizado el método de clasificación de la presente invención, la intensidad de autofluorescencia se aproxima como se ha descrito anteriormente.

30 Otro efecto parásito es la denominada intermodulación entre los diferentes canales A, C, G, T. La Figura 4 ilustra un efecto de intermodulación ilustrativo. En comparación con el caso ideal ilustrado por la Figura 2b, no sólo la señal principal tiene un valor diferente de cero, sino que también los otros canales exhiben valores no nulos, debido a los espectros fluorescentes superpuestos. Es decir, debido a una superposición en las frecuencias de emisión del colorante (compuesto fluorescente), puede observarse una interdependencia entre pares de canales de color (intermodulación). Como consecuencia, las intensidades de cuenta de un ciclo 10-50 están inherentemente sesgadas. La corrección de la intermodulación en las intensidades de autofluorescencia corregidas se hace utilizando una estimación grosera de la matriz de intermodulación, por ejemplo un factor del canal T en A es 0,4, y todos los demás factores son cero.

35 Después de la corrección de la iluminación y la corrección de los efectos parásitos, se realiza la computación de las desviaciones estándar de las intensidades corregidas a lo largo de los canales para cada cuenta y ciclo 10-50. La suposición es que las cuentas en blanco tienen una desviación estándar baja (en el caso ideal 0) y las cuentas vivas tienen una desviación estándar alta. Las desviaciones estándar se guardan para la clasificación viva subsiguiente, es decir un valor de desviación estándar para cada cuenta y ciclo 10-50 del intervalo de ciclos.

40 A continuación, se computa el valor medio de las desviaciones estándar de un ciclo 10-50 como una característica de clasificación para cada cuenta. El concepto básico de la computación de los valores de desviación estándar y el valor medio correspondiente se ilustra en la Figura 5, en donde se representan los diferentes canales para una posición de cuenta específica y un ciclo específico. La desviación estándar se calcula luego a partir de los diferentes canales para una posición de cuenta específica y un ciclo específico. Los valores de desviación estándar obtenidos para todos los ciclos se utilizan luego para calcular un valor medio para una posición de cuenta específica para la pluralidad de ciclos. De este modo, un valor medio representa una posición de cuenta, preferiblemente posición de cuenta viva, en cada ciclo.

45 Seguidamente se realiza una normalización de las características de clasificación, a fin de permitir la determinación de umbrales de clasificación apropiados. Es decir, se realiza una transformación de las características de clasificación en el intervalo [0,1] (transformación lineal común). Para ello, las características de clasificación se subdividen en varios bloques (grupos) basados en las posiciones de cuenta correspondientes y se determinan los valores de las características mínimas y máximas en cada bloque. Al hacerlo así, este procedimiento da como resultado dos matrices, una matriz de mínimos y una matriz de máximos.

En el caso de que las dos matrices comprendan elementos vacíos, las matrices min/max pueden llenarse por una estrategia de crecimiento regional, para propósitos de estabilidad.

Para eliminar posibles valores extremos en las matrices min/max resultantes, podría aplicarse un filtro de mediana (por ejemplo 3 x 3). Para alisar la superficie de las matrices min/max podría aplicarse un filtro Gaussiano.

- 5 A continuación, se aplica el aumento de escala de las dos matrices al tamaño original de las imágenes fuente, de tal modo que cada píxel tiene un valor de característica mínimo y máximo correspondiente. Por último, se realiza una normalización del valor de cada píxel $v(x,y)$ por una transformación lineal común

$$\hat{v}(x,y) = \frac{v(x,y) - \min(x,y)}{\max(x,y) - \min(x,y)},$$

donde $\hat{v}(x,y)$ es el valor de píxeles normalizado.

- 10 A continuación, se hace una determinación de los umbrales de clasificación, por ejemplo utilizando el método de segmentación de Otsu, véase la referencia [1]. En primer lugar, la superficie normalizada se subdivide en varios sub-bloques (subgrupos) y se extraen los valores de característica correspondientes en las posiciones de cuenta para cada sub-bloque. En segundo lugar, basándose en los valores de característica extraídos, se aplica en cada sub-bloque un método de umbralización de imágenes basado en agrupación (por ejemplo [1]) para determinar el umbral que distingue
- 15 cuentas vivas de cuentas en blanco. El fundamento del método de Otsu es que el algoritmo asume que una imagen consiste en dos clases de píxeles (píxeles de primer plano y de fondo), y asume por tanto un histograma bimodal. El umbral se calcula luego hallando el umbral óptimo que separaba las dos clases de píxeles. Un umbral optimizado determinado por el método de Otsu se ilustra en la Figura 6, en donde la línea ilustra un umbral óptimo.

- 20 Para propósitos de estabilidad, los elementos vacíos de la matriz de umbrales obtenida del paso previo podrían llenarse por una estrategia de crecimiento regional, en caso necesario.

A continuación, la matriz de umbrales se aumenta a escala al tamaño de la imagen original por aplicación de un filtro de mediana, un filtro Gaussiano con aumento de escala subsiguiente de la matriz de umbrales al tamaño de imagen original como se ha descrito arriba.

- 25 Una vez que se tiene una imagen de umbrales, esta imagen se utiliza luego para distinguir entre cuentas vivas y cuentas en blanco en la imagen normalizada, donde los píxeles brillantes se corresponden con cuentas vivas. En el caso de las cuentas vivas (marcadas como píxeles brillantes) el valor normalizado correspondiente de la cuenta es mayor que el umbral correspondiente. Así, el mapa binario final se utiliza para determinar y ajustar finalmente el estatus de cada cuenta, si la misma es una cuenta viva o cuenta en blanco.

Referencias:

- 30 [1] - <https://www.thermofisher.com/de/de/home/brands/product-brand/dynal.html>,
- [2] - Otsu,N, "Threshold selection method from graylevel histograms", IEEE Transactions on Systems, Man and Cybernetics, 1997, number 1, volume 9, pages 62-66.

REIVINDICACIONES

1. Un método para distinguir al menos un primer objeto de al menos un segundo objeto en una pluralidad de imágenes digitales, en donde el al menos un primer objeto ha recibido al menos una molécula que comprende información genética, no habiendo recibido el al menos un segundo objeto una molécula que comprende información genética, en donde la al menos una molécula está configurada para recibir uno de una pluralidad de compuestos fluorescentes en cada uno de una pluralidad de ciclos, determinándose las imágenes digitales por un sistema óptico de obtención de imágenes durante la emisión de radiación electromagnética por los compuestos fluorescentes, en donde la pluralidad de imágenes digitales comprende una pluralidad de series de imágenes, haciendo referencia cada imagen de una serie al espectro de emisión de un compuesto fluorescente respectivo y en donde la serie de imágenes se toma repetidamente para cada uno de la pluralidad de ciclos, en donde el método comprende los pasos siguientes, preferiblemente el orden siguiente:
- a) determinación de los valores de intensidad a partir de la pluralidad de imágenes digitales para cada objeto;
 - b) corrección de los efectos de autofluorescencia por sustracción de un valor de intensidad predeterminado a partir de los valores de intensidad determinados;
 - 15 c) corrección de los efectos de intermodulación causados por la superposición de los espectros fluorescentes de los compuestos fluorescentes por aplicación de una matriz de intermodulación predeterminada a los valores de intensidad determinados;
 - d) computación de los valores de desviación estándar a través de los diferentes espectros de emisión de los valores de intensidad corregidos para cada objeto y cada uno de la pluralidad de ciclos y computación de un valor medio de las desviaciones estándar para cada objeto;
 - 20 e) normalización de los valores medios para cada objeto utilizando una transformación lineal;
 - f) determinación de los umbrales de clasificación a partir de los valores medios normalizados utilizando un método de selección de umbrales para obtener una imagen digital binaria, en donde el al menos un primer objeto se representa por un primer valor de intensidad superior al umbral de clasificación correspondiente y el al menos un segundo objeto se representa por un segundo valor de intensidad inferior al umbral de clasificación correspondiente.
2. El método conforme a la reivindicación 1, que comprende adicionalmente los pasos siguientes, preferiblemente antes del paso a): determinación de los valores de intensidad mínimos dentro de una pluralidad de áreas predeterminadas de cada una de las imágenes digitales y corrección de la iluminación causada por el sistema óptico de obtención de imágenes para cada objeto en cada una de las imágenes digitales dividiendo los valores de intensidad determinados por los valores de intensidad mínimos correspondientes.
3. El método conforme a la reivindicación 2, que comprende adicionalmente aplicar un filtro de mediana a los valores de intensidad mínimos determinados de las áreas predeterminadas y/o aplicar un filtro Gaussiano a los valores de intensidad mínimos determinados de las áreas predeterminadas, preferiblemente después de la aplicación del filtro de mediana.
- 35 4. El método conforme a la reivindicación 2 ó 3, que comprende adicionalmente interpolar los valores de intensidad mínimos determinados de las áreas predeterminadas al tamaño de la imagen digital original, preferiblemente por el uso de una interpolación bicúbica.
5. El método conforme a cualquiera de las reivindicaciones 1 a 4, en donde el valor de intensidad predeterminado para corregir los efectos de autofluorescencia se determina tomando un valor de intensidad mínimo predeterminado a partir de las intensidades determinadas, preferiblemente tomando el valor 10% de intensidad mínima a partir de las intensidades determinadas.
- 40 6. El método conforme a cualquiera de las reivindicaciones 1 a 5, en donde la matriz de intermodulación da cuenta de una superposición de un espectro fluorescente verde con un espectro fluorescente amarillo de los compuestos fluorescentes, preferiblemente por aplicación de un factor de 0,4.
- 45 7. El método conforme a cualquiera de las reivindicaciones 1 a 6, en donde la normalización de los valores medios comprende subdividir los valores medios determinados en una pluralidad de grupos y determinar valores medios mínimos y valores medios máximos para cada grupo.
8. El método conforme a la reivindicación 7, que comprende adicionalmente aplicar una estrategia de crecimiento regional a los valores mínimos y los valores máximos y/o aplicar un filtro de mediana a los valores mínimos y los valores máximos y/o aplicar un filtro Gaussiano a los valores mínimos y los valores máximos.
- 50 9. El método conforme a la reivindicación 7 u 8, que comprende adicionalmente interpolar los valores mínimos y los valores máximos determinados al tamaño de la imagen digital original, preferiblemente por el uso de una interpolación bicúbica.

10. El método conforme a cualquiera de las reivindicaciones 7 a 9, en donde la transformación lineal se realiza para cada objeto mediante división de las diferencias entre los valores medios y los valores mínimos correspondientes por las diferencias correspondientes entre los valores máximos y los valores mínimos correspondientes.
- 5 11. El método conforme a cualquiera de las reivindicaciones 1 a 10, que comprende adicionalmente subdividir los valores medios normalizados en un número predeterminado de subgrupos y determinar los valores medios correspondientes para cada subgrupo antes de determinar los umbrales de clasificación, y preferiblemente aplicar subsiguientemente una estrategia de crecimiento regional y/o un filtro de mediana y/o un filtro Gaussiano.
- 10 12. El método conforme a cualquiera de las reivindicaciones 1 a 11, en donde el método de selección de umbrales comprende el método de segmentación de Otsu, aplicado preferiblemente a los valores medios determinados de los subgrupos.
13. El método conforme a la reivindicación 11 ó 12, que comprende adicionalmente interpolar los umbrales determinados por el método de selección de umbrales al tamaño de la imagen digital original, preferiblemente por el uso de una interpolación bicúbica.
- 15 14. Un sistema para distinguir al menos un primer objeto de al menos un segundo objeto en una pluralidad de imágenes digitales, en donde el al menos un primer objeto ha recibido al menos una molécula que comprende información genética, no habiendo recibido el al menos un segundo objeto una molécula que comprende información genética, en donde la al menos una molécula está configurada para recibir uno de una pluralidad de compuestos fluorescentes en cada uno de una pluralidad de ciclos, determinándose las imágenes digitales por un sistema óptico de obtención de imágenes durante la emisión de radiación electromagnética por los compuestos fluorescentes, en donde la pluralidad de imágenes digitales comprende una pluralidad de series de imágenes, haciendo referencia cada imagen de una serie al espectro de emisión de un compuesto fluorescente respectivo y en donde la serie de imágenes se toma repetidamente para cada uno de la pluralidad de ciclos, estando configurada preferiblemente para llevar a cabo los pasos de cualquiera de los métodos conforme a las reivindicaciones 1 a 13, que comprende:
- 20 una unidad de determinación de intensidad configurada para obtener valores intensidad a partir de la pluralidad de imágenes digitales para cada objeto;
- 25 una unidad de corrección configurada para corregir los efectos de autofluorescencia por sustracción de un valor de intensidad predeterminado de los valores de intensidad determinados y para corregir los efectos de intermodulación causados por la superposición de los espectros fluorescentes por aplicación de una matriz de intermodulación predeterminada a los valores de intensidad determinados;
- 30 una unidad de computación configurada para computar los valores de desviación estándar a través de los diferentes espectros de emisión de los valores de intensidad corregidos para cada objeto y cada uno de la pluralidad de ciclos y computación de los valores medios de las desviaciones estándar para cada objeto y computar los valores medios normalizados para cada objeto utilizando una transformación lineal y computar los umbrales de clasificación a partir de los valores medios normalizados utilizando un método de selección de umbrales para obtener una imagen digital binaria, en donde el al menos un primer objeto se representa por un primer valor de intensidad superior al umbral de clasificación correspondiente y el al menos un segundo objeto se representa por un segundo valor de intensidad inferior al umbral de clasificación correspondiente.
- 35 15. Un producto de programa de computadora que comprende uno o más medios legibles por computadora que tienen instrucciones ejecutables por computadora que, cuando son ejecutadas por una computadora, hacen que la computadora lleve a cabo los pasos del método de una de las reivindicaciones 1 a 13.
- 40

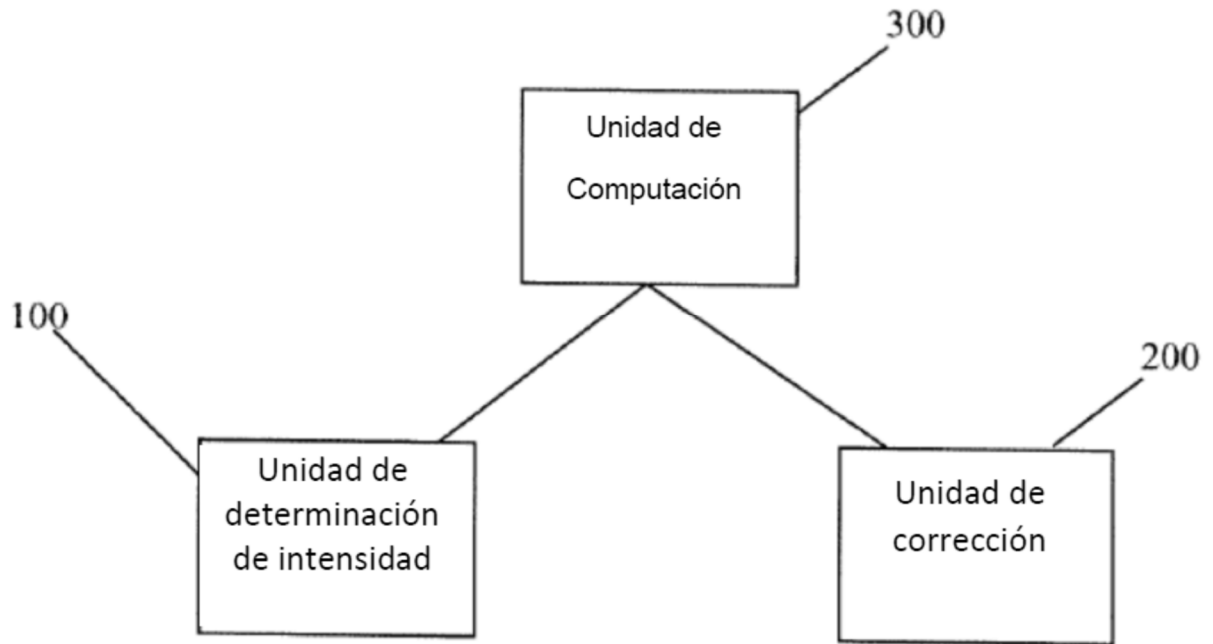


Fig. 1

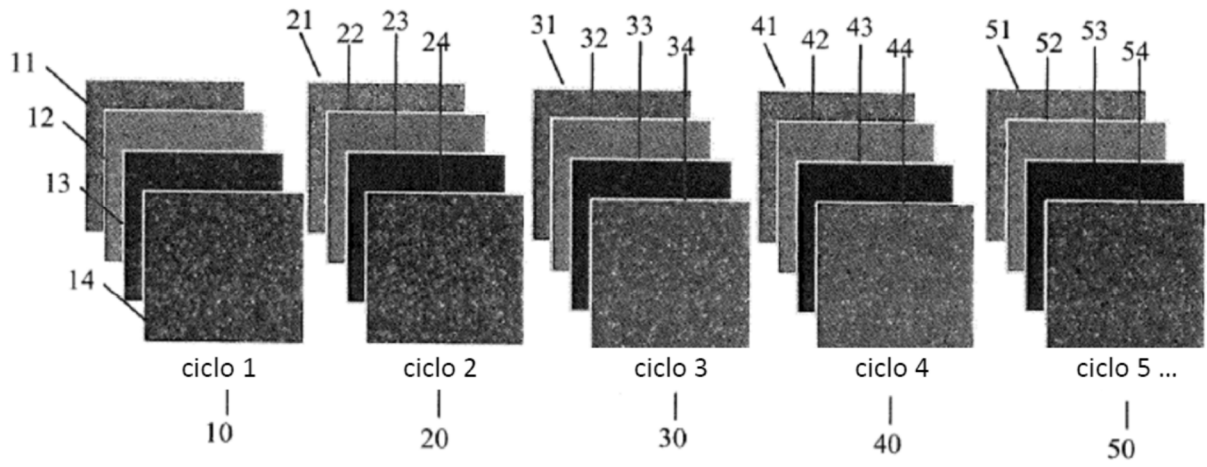


Fig. 2a

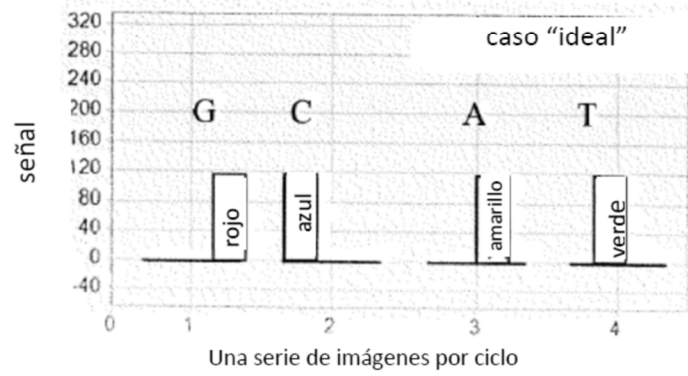


Fig. 2b

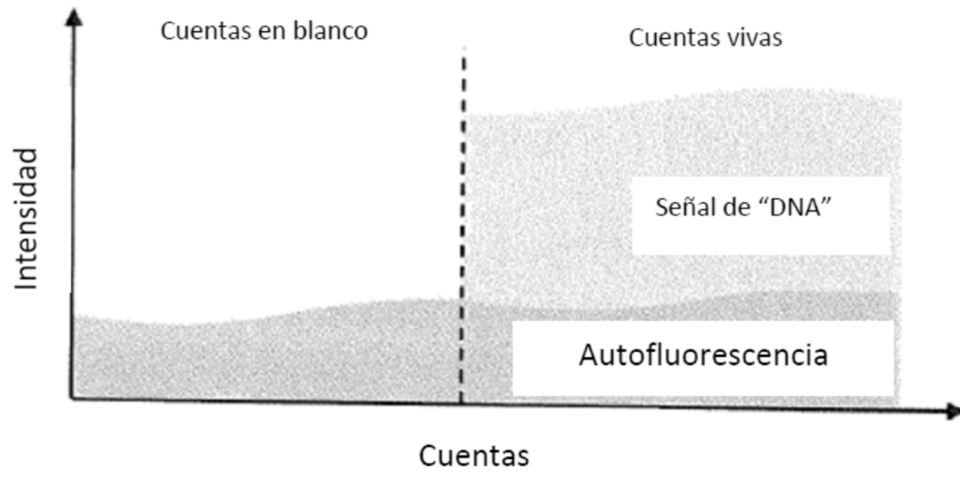


Fig. 3

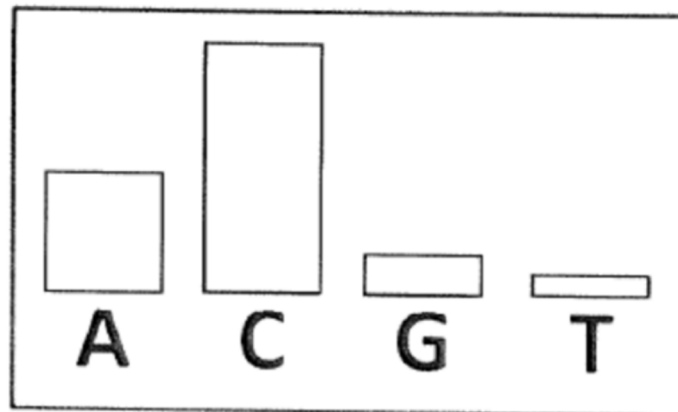


Fig. 4

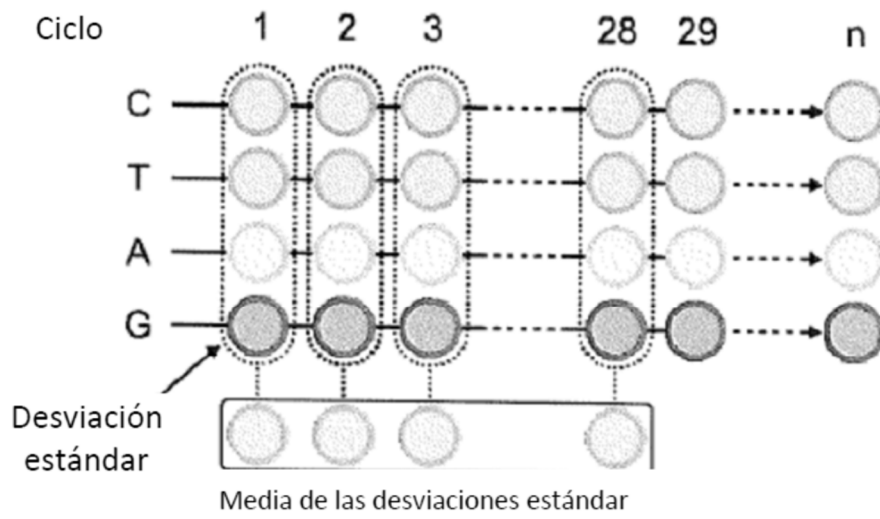
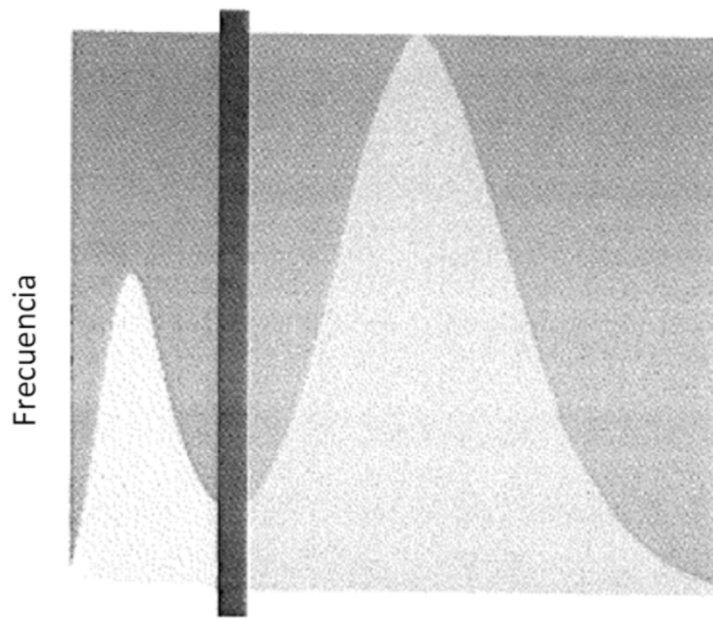


Fig. 5



Valores medios de la desviación estándar

Fig. 6