

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 796 346**

51 Int. Cl.:

C12Q 1/6869 (2008.01)

G01N 27/447 (2006.01)

G01N 33/487 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.11.2013 PCT/US2013/068967**

87 Fecha y número de publicación internacional: **15.05.2014 WO14074727**

96 Fecha de presentación y número de la solicitud europea: **07.11.2013 E 13853114 (0)**

97 Fecha y número de publicación de la concesión europea: **08.04.2020 EP 2917372**

54 Título: **Secuenciación de ácidos nucleicos usando marcas**

30 Prioridad:

09.11.2012 US 201261724869 P
14.12.2012 US 201261737621 P
20.09.2013 US 201361880407 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
26.11.2020

73 Titular/es:

ROCHE SEQUENCING SOLUTIONS, INC. (100.0%)
4300 Hacienda Drive
Pleasanton, CA 94588, US

72 Inventor/es:

DAVIS, RANDALL W.;
CHEN, ROGER;
BIBILLO, ARKADIUSZ;
KORENBLUM, DANIEL y
DORWART, MICHAEL

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 796 346 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

Secuenciación de ácidos nucleicos usando marcas

5 REFERENCIA CRUZADA

Esta solicitud reivindica el beneficio de la solicitud de patente provisional de EE. UU n.º de serie 61/724.869, presentada el 9 de noviembre de 2012, la solicitud de patente provisional de EE. UU n.º de serie 61/737.621, presentada el 14 de diciembre de 2012 y la solicitud de patente provisional de EE. UU n.º de serie 61/880.407, presentada el 20 de septiembre de 2013.

ANTECEDENTES

La secuenciación de ácidos nucleicos es un procedimiento que se puede usar para proporcionar información de la secuencia de una muestra de ácido nucleico. Dicha información de la secuencia puede ser útil para diagnosticar y/o tratar a un sujeto. Por ejemplo, la secuencia de ácido nucleico de un sujeto se puede usar para identificar, diagnosticar y potencialmente desarrollar tratamientos para enfermedades genéticas. Como otro ejemplo, la investigación de patógenos puede dar lugar al tratamiento de enfermedades contagiosas.

Hay procedimientos disponibles que se pueden usar para secuenciar un ácido nucleico. Sin embargo, dichos procedimientos son costosos y pueden no proporcionar información de la secuencia en un plazo y con una exactitud que pueden ser necesarios para diagnosticar y/o tratar a un sujeto.

SUMARIO

Los procedimientos de secuenciación de ácidos nucleicos que pasan una molécula de ácido nucleico monocatenario a través de un nanoporo pueden tener una sensibilidad que puede ser insuficiente o, de otro modo inadecuada para proporcionar una fecha con propósitos de diagnóstico y/o tratamiento. Las bases de ácido nucleico que comprenden la molécula de ácido nucleico (por ejemplo, adenina (A), citosina (C), guanina (G), timina (T) y/o uracilo (U)) pueden no proporcionar una señal suficientemente distinta entre sí. En particular, las purinas (es decir, A y G) tienen un tamaño, conformación y carga similares entre sí y proporcionan una señal insuficientemente distinta en algunos casos. También las pirimidinas (es decir, C, T y U) son de un tamaño, conformación y carga similares entre sí y proporcionan una señal insuficientemente distinta en algunos casos. En el presente documento se reconoce la necesidad de procedimientos mejorados para la identificación de moléculas de ácido nucleico y la secuenciación de ácidos nucleicos.

En un aspecto, un procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor comprende: (a) proporcionar nucleótidos marcados en una cámara de reacción que comprende el nanoporo, en el que un nucleótido marcado individual de los nucleótidos marcados contiene una marca acoplada a un nucleótido, dicha marca es detectable con la ayuda del nanoporo; (b) llevar a cabo una reacción de polimerización, con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de los nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de la muestra de ácido nucleico; y (c) detectar, con la ayuda del nanoporo, una marca asociada con el nucleótido marcado individual durante la incorporación del nucleótido marcado individual, en el que se detecta la marca con la ayuda del nanoporo cuando el nucleótido marcado se asocia con la polimerasa.

Los acontecimientos de incorporación de nucleótidos (por ejemplo, incorporación de un nucleótido en una hebra de ácido nucleico que es complementaria a una hebra molde) pueden presentar marcas a un nanoporo y/o liberar marcas de los nucleótidos que detecta un nanoporo. La base incorporada se puede identificar (es decir, A, C, G, T o U) porque se libera y/o presenta una marca única para cada tipo de nucleótido (es decir, A, C, G, T o U).

Una marca se puede atribuir a un nucleótido incorporado con éxito en base al periodo de tiempo en el que se detecta que la marca interactúa con un nanoporo. El periodo de tiempo puede ser mayor que el periodo de tiempo asociado con la circulación libre de la marca de nucleótidos a través del nanoporo. El periodo de tiempo de detección de una marca de nucleótido incorporada con éxito también puede ser mayor que el periodo de tiempo de un nucleótido no incorporado (por ejemplo, un nucleótido que no se empareja correctamente con la hebra molde).

Una polimerasa se puede asociar con el nanoporo (por ejemplo, se une covalentemente al nanoporo) y la polimerasa realiza acontecimientos de incorporación de nucleótidos. El nanoporo puede detectar la marca cuando el nucleótido marcado está asociado con la polimerasa. En algunos casos, los nucleótidos marcados no incorporados pasan a través del nanoporo. El procedimiento puede distinguir entre marcas asociadas con nucleótidos no incorporados y marcas asociadas con nucleótidos incorporados en base al periodo de tiempo en el que el nanoporo detecta el nucleótido marcado. En un modo de realización, el nanoporo detecta un nucleótido no incorporado durante menos de aproximadamente 1 milisegundo y el nanoporo detecta un nucleótido incorporado durante al menos aproximadamente 1 milisegundo.

La polimerasa puede tener una etapa cinética lenta donde la marca es detectable por el nanoporo durante al menos 1 milisegundo, con un tiempo de detección promedio de aproximadamente 100 ms. La polimerasa puede ser una ADN polimerasa del fago Φ 29 mutada.

La polimerasa se puede mutar para reducir la velocidad a la que la polimerasa incorpora un nucleótido en una hebra de ácido nucleico (por ejemplo, una hebra de ácido nucleico en crecimiento). En algunos casos, la velocidad a la que se incorpora un nucleótido en una hebra de ácido nucleico se puede reducir funcionalizando la hebra de nucleótidos y/o molde para proporcionar impedimento estérico, tal como, por ejemplo, a través de metilación de la hebra de ácido nucleico molde. En algunos casos, la velocidad se reduce incorporando nucleótidos metilados.

En algunos modos de realización, la marca se detecta una pluralidad de veces mientras está asociada con la polimerasa.

En algunos modos de realización, un electrodo se recarga entre periodos de detección de marcas.

En algunos modos de realización, el procedimiento distingue entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado en base al periodo de tiempo en el que el nanoporo detecta el nucleótido marcado.

La proporción entre el tiempo en el que el nanoporo detecta un nucleótido marcado incorporado y el tiempo en el que el nanoporo puede detectar un nucleótido marcado no incorporado es al menos aproximadamente 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 1000 o 10.000.

La proporción entre el periodo de tiempo en el que una marca asociada con un nucleótido incorporado interactúa con (y se detecta con la ayuda de) un nanoporo y el periodo de tiempo en el que una marca asociada con un nucleótido no incorporado interactúa con el nanoporo puede ser al menos aproximadamente 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 1000 o 10.000.

En algunos modos de realización, el nucleótido se asocia con la polimerasa durante un periodo de tiempo promedio (o medio) de al menos aproximadamente 1 milisegundo.

En algunos modos de realización, el nucleótido marcado pasa a través del nanoporo en menos de 1 milisegundo (ms) cuando el nucleótido no está asociado con la polimerasa.

La marca puede tener una longitud seleccionada para que sea detectable por el nanoporo.

La incorporación de un primer nucleótido marcado puede interferir o no en la detección por el nanoporo de una marca asociada con un segundo nucleótido marcado.

La detección por el nanoporo de una marca asociada con un primer nucleótido marcado puede interferir o no en la incorporación de un segundo nucleótido marcado.

En algunos modos de realización, el nanoporo puede distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos un 95 %.

En algunos modos de realización, el nanoporo puede distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos un 99 %.

La marca asociada con el nucleótido marcado individual se puede detectar cuando la marca se libera del nucleótido marcado individual.

En un modo de realización de la invención, el procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor se caracteriza por que durante la incorporación del nucleótido marcado individual, se diferencia, con la ayuda del nanoporo, una marca asociada con el nucleótido marcado individual de una o más marcas asociadas con uno o más nucleótidos marcados individuales no incorporados.

La enzima puede ser una polimerasa de ácido nucleico o cualquier enzima que pueda extender una hebra recién sintetizada en base a un polímero molde.

En algunos modos de realización, el nucleótido marcado individual incorporado en (b) se diferencia de los nucleótidos marcados individuales no incorporados en base a los periodos de tiempo y/o las proporciones de tiempo en las que se detecta el nucleótido marcado individual incorporado en (b) y los nucleótidos marcados individuales no incorporados con la ayuda del nanoporo.

- Un procedimiento descrito para secuenciar un ácido nucleico con la ayuda de un nanoporo en una membrana comprende: (a) proporcionar nucleótidos marcados en una cámara de reacción que comprende el nanoporo, en el que un nucleótido marcado individual de los nucleótidos marcados contiene una marca que es detectable por el nanoporo; (b) incorporar los nucleótidos marcados en una cadena de ácido nucleico en crecimiento, en el que la marca asociada con un nucleótido marcado individual de los nucleótidos marcados reside en o próxima a al menos una parte del nanoporo durante la incorporación, en el que la proporción entre el tiempo en el que el nanoporo puede detectar un nucleótido marcado incorporado y el tiempo en el que el nanoporo puede detectar una marca no incorporada es al menos 1,1, 1,2, 1,3, 1,4, 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 1000, 10.000; y (c) detectar la marca con la ayuda del nanoporo.
- La proporción entre el tiempo en el que el nanoporo puede detectar un nucleótido marcado incorporado y el tiempo en el que el nanoporo puede detectar una marca no incorporada puede ser al menos aproximadamente 1,1, 1,2, 1,3, 1,4, 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 1000 o 10.000.
- La marca puede permanecer asociada con un nucleótido individual tras la incorporación del nucleótido.
- La marca asociada con un nucleótido individual se puede liberar tras la incorporación del nucleótido.
- El procedimiento puede comprender además expulsar la marca del nanoporo.
- La marca se puede expulsar en la dirección inversa a la cual la marca entró en el nanoporo.
- La marca puede residir en el nanoporo durante al menos aproximadamente 100 ms.
- La marca puede residir en el nanoporo durante al menos aproximadamente 10 ms.
- La marca puede residir en el nanoporo durante al menos aproximadamente 1 ms.
- Los nucleótidos marcados se pueden incorporar a una velocidad de como máximo aproximadamente 1 nucleótido por segundo.
- El nanoporo puede expulsar la molécula marcadora con un impulso de tensión.
- La molécula marcadora puede tener al menos un 99 % de probabilidades de ser expulsada con el impulso de tensión.
- El nanoporo puede expulsar la molécula marcadora dentro de un periodo de tiempo de modo que dos moléculas marcadoras no estén presentes en el nanoporo al mismo tiempo.
- El nanoporo puede expulsar la molécula marcadora dentro de un periodo de tiempo de modo que la probabilidad de que dos moléculas marcadoras estén presentes en el nanoporo al mismo tiempo sea como máximo un 1 %.
- La marca puede tener un diámetro inferior a aproximadamente 1,4 nm.
- Cada marca asociada con un nucleótido marcado incorporado se puede detectar con la ayuda del nanoporo mientras la marca está unida al nucleótido.
- En algunos modos de realización, la marca asociada con el nucleótido marcado individual se detecta cuando la marca se libera del nucleótido marcado individual.
- También se divulga un chip para secuenciar una muestra de ácido nucleico, que comprende: una pluralidad de nanoporos, un nanoporo de la pluralidad teniendo al menos un nanoporo en una membrana dispuesta contigua o próxima a un electrodo, en el que cada nanoporo detecta una marca asociada con un nucleótido marcado individual durante la incorporación del nucleótido marcado en una cadena de ácido nucleico en crecimiento. En algunos modos de realización, los nanoporos son direccionables individualmente.
- El nanoporo individual puede detectar la marca asociada con el nucleótido durante el paso posterior de la marca a través de o contigua al nanoporo.
- El chip puede comprender al menos 500 electrodos direccionables individualmente por milímetro cuadrado. En algunos modos de realización, el chip comprende al menos 50 electrodos direccionables individualmente por milímetro cuadrado.
- El chip puede distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado en base, al menos en parte, al periodo de tiempo en el que el nanoporo detecta el nucleótido marcado.

La proporción entre el tiempo en el que un nucleótido marcado incorporado es detectable por el nanoporo y el tiempo en el que una marca no incorporada es detectable por el nanoporo puede ser al menos aproximadamente 1,5.

5 La incorporación de un primer nucleótido marcado puede interferir o no en la detección por el nanoporo de una marca asociada con un segundo nucleótido marcado.

La detección por el nanoporo de una marca asociada con un primer nucleótido marcado puede interferir o no en la incorporación de un segundo nucleótido marcado.

10 El nanoporo puede ser capaz de distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos un 95 %.

15 El nanoporo puede ser capaz de distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos un 99 %.

El electrodo puede ser parte de un circuito integrado.

20 El electrodo se puede acoplar a un circuito integrado.

Cada marca asociada con un nucleótido marcado incorporado se puede detectar con la ayuda del nanoporo mientras la marca está unida al nucleótido.

25 La marca asociada con el nucleótido marcado individual se puede detectar cuando la marca se libera del nucleótido marcado individual.

30 También se divulga un chip para secuenciar una muestra de ácido nucleico, que comprende: una pluralidad de nanoporos, en el que un nanoporo de la pluralidad contiene al menos un nanoporo en una membrana dispuesta contigua a un electrodo, en el que cada nanoporo puede detectar una especie marcadora tras o durante la incorporación de una molécula de ácido nucleico que comprende la especie marcadora en una cadena de ácido nucleico en crecimiento, en el que la proporción entre el tiempo en el que un nucleótido marcado incorporado es detectable por el nanoporo y el tiempo en el que una marca no incorporada es detectable por el nanoporo es al menos aproximadamente 1,5. En algunos modos de realización, la pluralidad de nanoporos son direccionables individualmente.

35 La especie marcadora puede pasar o no a través del nanoporo tras la incorporación.

El chip se puede configurar para expulsar la especie marcadora del nanoporo.

40 El nanoporo puede expulsar la especie marcadora con un impulso de tensión.

El electrodo puede ser parte de un circuito integrado.

45 El electrodo se puede acoplar a un circuito integrado.

Cada marca asociada con un nucleótido marcado incorporado se puede detectar con la ayuda del nanoporo mientras la marca está unida al nucleótido.

50 La especie marcadora de la molécula de ácido nucleico se puede detectar sin que la especie marcadora se libere de la molécula de ácido nucleico.

55 También se describe un sistema para secuenciar una muestra de ácido nucleico, que comprende: (a) un chip que comprende uno o más dispositivos de nanoporos, cada uno de los uno o más dispositivos de nanoporos comprendiendo un nanoporo en una membrana que es contigua a un electrodo, en el que el dispositivo de nanoporos detecta una marca asociada con un nucleótido marcado individual durante la incorporación del nucleótido marcado por una polimerasa; y (b) un procesador acoplado al chip, en el que el procesador se programa para ayudar a caracterizar una secuencia de ácido nucleico de la muestra de ácido nucleico en base a señales eléctricas recibidas del dispositivo de nanoporos.

60 El dispositivo de nanoporos puede detectar una marca asociada con un nucleótido marcado individual durante la progresión posterior de la marca a través de o contigua al nanoporo.

El dispositivo de nanoporos puede comprender nanoporos direccionables individualmente.

El chip puede comprender al menos 500 electrodos direccionables individualmente por milímetro cuadrado. En algunos modos de realización, el chip comprende al menos 50 electrodos direccionables individualmente por milímetro cuadrado.

- 5 El chip puede distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado en base, al menos en parte, al periodo de tiempo en el que el nanoporo detecta el nucleótido marcado.

10 La proporción entre el tiempo en el que un nucleótido marcado incorporado es detectable por el nanoporo y el tiempo en el que una marca no incorporada es detectable por el nanoporo puede ser al menos aproximadamente 1,5.

La incorporación de un primer nucleótido marcado puede interferir o no en la detección por el nanoporo de una marca asociada con un segundo nucleótido marcado.

- 15 La detección por el nanoporo de una marca asociada con un primer nucleótido marcado puede interferir o no en la incorporación de un segundo nucleótido marcado.

20 El nanoporo puede ser capaz de distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos un 95 %.

El nanoporo puede ser capaz de distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos un 99 %.

25 El electrodo puede ser parte de un circuito integrado.

El electrodo se puede acoplar a un circuito integrado.

30 Cada marca asociada con un nucleótido marcado incorporado se puede detectar con la ayuda del nanoporo mientras la marca está unida al nucleótido.

La marca asociada con el nucleótido marcado individual se puede detectar cuando la marca se libera del nucleótido marcado individual.

35 Una marca se puede dirigir hacia y a través de al menos una parte del nanoporo usando una fuerza impulsora dada, tal como un potencial eléctrico aplicado al nanoporo o membrana que contiene el nanoporo (V+). La marca se puede dirigir al nanoporo desde una abertura del nanoporo. A continuación, la fuerza impulsora se puede invertir (por ejemplo, potencial eléctrico de polaridad opuesta aplicada, o V-) para expulsar al menos una parte de la marca del nanoporo a través de la abertura. A continuación, la fuerza impulsora se puede aplicar nuevamente (por ejemplo, V+) para impulsar al menos una parte de la marca hacia el nanoporo a través de la abertura. De forma alternativa, se puede invertir una polaridad de la marca y se puede usar una secuencia de potenciales que incluye V-, V+ y V-. Esto puede incrementar el periodo de tiempo en el que la marca se puede detectar con la ayuda del nanoporo.

45 Se puede configurar un nanoporo y/o marca para proporcionar un panorama energético de modo que, en el nanoporo, la marca asociada con un nucleótido tenga más probabilidades de moverse en una dirección (por ejemplo, hacia el poro) que en otra dirección (por ejemplo, fuera del poro).

50 La detección de bases modificadas (por ejemplo, metiladas) en una hebra de muestra molde se puede detectar, por ejemplo, mediante la diferencia en el tiempo en el que un nanoporo detecta una marca de un nucleótido marcado mientras está asociada con una polimerasa durante y/o tras la incorporación de la parte de nucleótido de la marca de nucleótidos en una hebra recién sintetizada. En algunos casos, el tiempo en el que una marca de nucleótido se asocia con una enzima es mayor cuando el nucleótido opuesto de la secuencia de muestra es un nucleótido metilado en comparación con un nucleótido no metilado.

55 Los ejemplos de nucleótidos marcados descritos en el presente documento pueden ser cualquier nucleótido natural modificado con marcas escindibles o análogos de nucleótidos sintéticos no naturales modificados con marcas escindibles. Por ejemplo, las bases universales modificadas con marcas escindibles o no escindibles se pueden usar para contabilizar simplemente el número de bases en una hebra de muestra.

60 Los ejemplos de nucleótidos marcados descritos en el presente documento pueden ser nucleótidos diméricos o análogos de nucleótidos diméricos que se pueden extender como unidades diméricas y las marcas informan de la composición combinada de dímeros del nucleótido dimérico en base al tiempo en asociación con la enzima polimerasa y el nivel de señal detectado por un dispositivo de nanoporos.

65 Aunque el tiempo en el que una marca se asocia con una enzima se puede usar para diferenciar entre nucleótidos incorporados y nucleótidos no incorporados, los niveles de corriente única y/o la respuesta eléctrica de una marca

en un nanoporo a un potencial aplicado o un potencial aplicado variable, permite la diferenciación entre marcas asociadas con diferentes nucleótidos.

5 En un aspecto, un procedimiento para secuenciar un ácido nucleico comprende aplicar una forma de onda de corriente alterna (CA) a un circuito próximo a un nanoporo y un electrodo sensor, en el que una marca asociada con un nucleótido que se incorpora a una hebra de ácido nucleico en crecimiento complementaria a una hebra de ácido nucleico molde se detecta cuando la forma de onda tiene una primera polaridad y el electrodo se recarga cuando la forma de onda tiene una segunda polaridad.

10 También se describe un procedimiento para secuenciar una molécula de ácido nucleico que comprende: (a) proporcionar uno o más nucleótidos marcados a un nanoporo en una membrana contigua a un electrodo; (b) incorporar un nucleótido marcado individual de dicho uno o más nucleótidos marcados en una hebra complementaria a dicha molécula de ácido nucleico; y (c) detectar una marca asociada con dicho nucleótido marcado una o más veces con la ayuda de una forma de onda de corriente alterna (CA) aplicada a dicho electrodo,
15 en el que dicha marca se detecta mientras dicha marca está unida a dicho nucleótido marcado individual incorporado en dicha hebra.

La forma de onda puede ser tal que el electrodo no se agote durante un periodo de tiempo de al menos aproximadamente 1 segundo, 10 segundos, 30 segundos, 1 minuto, 10 minutos, 20 minutos, 30 minutos, 1 hora,
20 2 horas, 3 horas, 4 horas, 5 horas, 6 horas, 12 horas, 24 horas, 1 día, 2 días, 3 días, 4 días, 5 días, 6 días, 1 semana, 2 semanas, 3 semanas o 1 mes.

La identidad de la marca se puede determinar mediante la relación entre la corriente medida y la tensión aplicada por la forma de onda a diversas tensiones.

25 En algunos modos de realización, los nucleótidos comprenden adenina (A), citosina (C), timina (T), guanina (G), uracilo (U) o cualquier derivado de los mismos.

La metilación de una base de la hebra de ácido nucleico molde se puede determinar por la detección de la marca durante un periodo de tiempo mayor cuando la base está metilada que cuando la base no está metilada.

En un aspecto, un procedimiento para determinar la longitud de un ácido nucleico o segmento del mismo con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor comprende (a) proporcionar nucleótidos marcados en una cámara de reacción que comprende dicho nanoporo, en el que los nucleótidos que tienen al
35 menos dos bases diferentes contienen la misma marca acoplada a un nucleótido, dicha marca es detectable con la ayuda de dicho nanoporo; (b) llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de dichos nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de dicha muestra de ácido nucleico; y (c) detectar, con la ayuda de dicho nanoporo, una marca asociada con dicho nucleótido marcado individual durante la incorporación de dicho nucleótido marcado individual.

También se describe un procedimiento para determinar la longitud de un ácido nucleico o segmento del mismo con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor que comprende (a) proporcionar nucleótidos marcados en una cámara de reacción que comprende dicho nanoporo, en el que un nucleótido
45 marcado individual de dichos nucleótidos marcados contiene una marca acoplada a un nucleótido, dicha marca puede reducir la magnitud de la corriente que fluye a través de dicho nanoporo con respecto a la corriente cuando la marca no está presente; (b) llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de dichos nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de dicha muestra de ácido nucleico
50 y reducir la magnitud de la corriente que fluye a través de dicho nanoporo; y (c) detectar, con la ayuda de dicho nanoporo, periodos de tiempo entre la incorporación de dichos nucleótidos marcados individuales. En algunos modos de realización, la magnitud de la corriente que fluye a través de dicho nanoporo vuelve a al menos un 80 % de la corriente máxima durante los periodos de tiempo entre la incorporación de dichos nucleótidos marcados individuales.

55 Todos los nucleótidos pueden tener la misma marca acoplada al nucleótido. En algunos modos de realización, al menos algunos de los nucleótidos tienen marcas que identifican el nucleótido. En algunos modos de realización, como máximo un 20 % de los nucleótidos tienen marcas que identifican el nucleótido. En algunos modos de realización, todo el nucleótido se identifica como adenina (A), citosina (C), guanina (G), timina (T) y/o uracilo (U).
60 En algunos modos de realización, todo el ácido nucleico o un segmento del mismo es una repetición corta en tándem (STR).

También se describe un procedimiento para ensamblar una proteína que tiene una pluralidad de subunidades que comprende (a) proporcionar una pluralidad de primeras subunidades; (b) proporcionar una pluralidad de segundas
65 subunidades, en el que las segundas subunidades se modifican con respecto a las primeras subunidades; (c) poner en contacto las primeras subunidades con las segundas subunidades en una primera proporción para formar

- una pluralidad de proteínas que tienen las primeras subunidades y las segundas subunidades, en el que la pluralidad de proteínas tiene una pluralidad de proporciones entre las primeras subunidades y las segundas subunidades; y (d) fraccionar la pluralidad de proteínas para enriquecer proteínas que tienen una segunda proporción entre las primeras subunidades y las segundas subunidades, en el que la segunda proporción es de una segunda subunidad por (n-1) primeras subunidades, en la que 'n' es el número de subunidades que comprenden la proteína.
- La proteína puede ser un nanoporo.
- El nanoporo puede ser al menos un 80 % homólogo a la hemolisina alfa.
- Las primeras subunidades o las segundas subunidades pueden comprender una marca de purificación.
- La marca de purificación puede ser una marca de polihistidina.
- El fraccionamiento se puede realizar usando cromatografía de intercambio iónico.
- La segunda proporción puede ser de 1 segunda subunidad por 6 primeras subunidades.
- La segunda proporción puede ser de 2 segundas subunidades por 5 primeras subunidades y una única polimerasa se une a cada una de las segundas subunidades.
- Las segundas subunidades pueden comprender un resto químicamente reactivo y el procedimiento puede comprender además (e) realizar una reacción para unir una entidad al resto químicamente reactivo.
- La proteína puede ser un nanoporo y la entidad puede ser una polimerasa.
- Las primeras subunidades pueden ser naturales.
- Las primeras subunidades y/o segundas subunidades pueden ser recombinantes.
- La primera proporción puede ser aproximadamente igual a la segunda proporción.
- La primera proporción puede ser mayor que la segunda proporción.
- El procedimiento puede comprender además insertar las proteínas que tienen las subunidades de la segunda proporción en una bicapa.
- El procedimiento puede comprender además secuenciar una molécula de ácido nucleico con la ayuda de las proteínas que tienen las subunidades de la segunda proporción.
- Un nanoporo puede comprender una pluralidad de subunidades, en el que se une una polimerasa a una de las subunidades y al menos una y menos de todas las subunidades comprenden una primera marca de purificación.
- El nanoporo puede ser al menos un 80 % homólogo a la hemolisina alfa.
- Todas las subunidades pueden comprender una primera marca de purificación o una segunda marca de purificación.
- La primera marca de purificación puede ser una marca de polihistidina.
- La primera marca de purificación puede estar en la subunidad que tiene la polimerasa unida.
- La primera marca de purificación puede estar en las subunidades que no tienen la polimerasa unida.
- En un modo de realización, el procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor comprende: (a) proporcionar nucleótidos marcados en una cámara de reacción que comprende dicho nanoporo, en el que un nucleótido marcado individual de dichos nucleótidos marcados contiene una marca acoplada a un nucleótido, dicha marca es detectable con la ayuda de dicho nanoporo; (b) llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de dichos nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de dicha muestra de ácido nucleico; (c) detectar, con la ayuda de dicho nanoporo, una marca asociada con dicho nucleótido marcado individual durante la incorporación de dicho nucleótido marcado individual, en el que dicha marca se detecta con la ayuda de dicho nanoporo cuando dicho nucleótido está asociado con dicha polimerasa, se caracteriza por que dicha detección

comprende (i) proporcionar una tensión aplicada a través de dicho nanoporo y (ii) medir una corriente con dicho electrodo sensor a dicha tensión aplicada; y (d) calibrar dicha tensión aplicada.

- 5 En algunos modos de realización, dicha calibración comprende (i) medir una pluralidad de tensiones de escape para dicha molécula marcadora, (ii) calcular una diferencia entre las tensiones de escape medidas y un punto de referencia y (iii) desviar la tensión aplicada por la diferencia calculada.

En algunos modos de realización, se estima una distribución de tensiones de escape esperadas frente al tiempo.

- 10 En algunos modos de realización, el punto de referencia es la media o mediana de las tensiones de escape medidas.

El procedimiento elimina las variaciones detectadas en la distribución de tensiones de escape esperadas.

- 15 El procedimiento se realiza en una pluralidad de nanoporos direccionables independientemente, cada uno contiguo a un electrodo sensor.

En algunos modos de realización, la tensión aplicada disminuye con el tiempo.

- 20 La presencia de la marca en el nanoporo reduce la corriente medida con el electrodo sensor a la tensión aplicada.

Los nucleótidos marcados comprenden una pluralidad de marcas diferentes y el procedimiento detecta cada una de la pluralidad de marcas diferentes.

- 25 La etapa (d) puede incrementar la exactitud del procedimiento en comparación con la realización de las etapas (a)-(c).

La etapa (d) puede compensar los cambios en las condiciones electroquímicas con el tiempo.

- 30 La etapa (d) puede compensar que diferentes nanoporos tengan condiciones electroquímicas diferentes en un dispositivo que tiene una pluralidad de nanoporos.

La etapa (d) puede compensar condiciones electroquímicas diferentes para cada realización del procedimiento.

- 35 En algunos modos de realización, el procedimiento comprende además (e) calibrar variaciones en una ganancia de corriente y/o variaciones en una compensación de corriente.

En algunos modos de realización, dicha marca se detecta una pluralidad de veces mientras está asociada con dicha polimerasa.

- 40 Se puede recargar un electrodo entre los periodos de detección de marcas.

El procedimiento distingue entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado en base al periodo de tiempo en el que dicho nanoporo detecta dicho nucleótido marcado.

- 45 En un modo de realización, el procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor comprende además eliminar una secuencia de ácido nucleico repetitiva de la muestra de ácido nucleico para proporcionar una molécula de ácido nucleico monocatenario para la secuenciación.

- 50 La secuencia de ácido nucleico repetitiva puede comprender al menos 20 bases de ácido nucleico consecutivas.

La secuencia de ácido nucleico repetitiva puede comprender al menos 200 bases de ácido nucleico consecutivas.

- 55 La secuencia de ácido nucleico repetitiva puede comprender al menos 20 subunidades repetidas consecutivas de bases de ácido nucleico.

La secuencia de ácido nucleico repetitiva puede comprender al menos 200 subunidades repetidas consecutivas de bases de ácido nucleico.

- 60 En algunos modos de realización, la secuencia de ácido nucleico repetitiva se elimina por hibridación con una secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva.

- 65 La secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva se puede inmovilizar sobre un soporte sólido.

El soporte sólido puede ser una superficie.

El soporte sólido puede ser una microesfera.

- 5 En algunos modos de realización, la secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva comprende ADN Cot-1.

El ADN Cot-1 se puede enriquecer en secuencias de ácido nucleico repetitivas que tienen una longitud de entre aproximadamente 50 y aproximadamente 100 bases de ácido nucleico.

- 10 En un modo de realización, el procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor se caracteriza por que la polimerasa se une al nanoporo mediante un conector.

- 15 El conector puede ser flexible.

El conector puede tener al menos 5 nanómetros de longitud.

El conector puede ser una unión directa.

- 20 El conector puede comprender aminoácidos.

El nanoporo y la polimerasa pueden comprender un único polipéptido.

- 25 El conector puede comprender ácidos nucleicos o polietilenglicol (PEG).

El conector puede comprender un enlace no covalente.

El conector puede comprender biotina y estreptavidina.

- 30 Se puede seleccionar al menos uno de los siguientes: (a) el extremo C de la polimerasa se une al extremo N del nanoporo; (b) el extremo C de la polimerasa se une al extremo C del nanoporo; (c) el extremo N de la polimerasa se une al extremo N del nanoporo; (d) el extremo N de la polimerasa se une al extremo C del nanoporo; y (e) la polimerasa se une al nanoporo donde al menos uno de la polimerasa y el nanoporo no están unidos a un extremo.

- 35 El conector puede orientar la polimerasa con respecto al nanoporo de modo que la marca se detecte con la ayuda del nanoporo.

La polimerasa se puede unir al nanoporo mediante dos o más conectores.

- 40 El conector puede comprender uno o más de SEQ ID N.º 2-35 o un producto de PCR producido a partir de los mismos.

- 45 El conector puede comprender el péptido codificado por uno o más de SEQ ID N.º 1-35 o un producto de PCR producido a partir de los mismos.

El nanoporo puede ser al menos un 80 % homólogo a la hemolisina alfa.

La polimerasa puede ser al menos un 80 % homóloga al fago Φ29.

- 50 Una molécula marcadora puede comprender (a) una primera cadena polimérica que comprende un primer segmento y un segundo segmento, en la que el segundo segmento es más estrecho que el primer segmento; y (b) una segunda cadena polimérica que comprende dos extremos, en la que un primer extremo se fija a la primera cadena polimérica contigua al segundo segmento y un segundo extremo no se fija a la primera cadena polimérica, en la que la molécula marcadora se puede enhebrar a través de un nanoporo en una primera dirección donde la segunda cadena polimérica se alinea contigua al segundo segmento.

- 55 La molécula marcadora puede no ser capaz de enhebrarse a través del nanoporo en una segunda dirección donde la segunda cadena polimérica no se alinea contigua al segundo segmento.

- 60 La base de la segunda cadena polimérica se puede emparejar con la primera cadena polimérica cuando la segunda cadena polimérica no se alinea contigua al segundo segmento.

La primera cadena polimérica se puede fijar a un nucleótido.

- 65

La primera cadena polimérica se puede liberar del nucleótido cuando el nucleótido se incorpora en una hebra de ácido nucleico en crecimiento.

La primera cadena polimérica se puede fijar a un fosfato terminal del nucleótido.

La primera cadena polimérica puede comprender nucleótidos.

El segundo segmento puede comprender nucleótidos abásicos.

El segundo segmento puede comprender una cadena de carbonos.

En otro modo de realización, el procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor se caracteriza por que la marca comprende (i) una primera cadena polimérica que comprende un primer segmento y un segundo segmento, en el que el segundo segmento es más estrecho que el primer segmento y (ii) una segunda cadena polimérica que comprende dos extremos, en el que un primer extremo se fija a la primera cadena polimérica contigua al segundo segmento y un segundo extremo no se fija a la primera cadena polimérica, en el que la molécula marcadora se puede enhebrarse a través de un nanoporo en una primera dirección donde la segunda cadena polimérica se alinea contigua al segundo segmento.

La molécula marcadora no se puede enhebrar a través del nanoporo en una segunda dirección donde la segunda cadena polimérica no se alinea contigua al segundo segmento.

Dicha marca se puede detectar una pluralidad de veces mientras está asociada con dicha polimerasa.

Se puede recargar un electrodo entre los periodos de detección de marcas.

La marca se enhebra en el nanoporo durante la incorporación del nucleótido marcado individual, y en la que la marca no se desenhebra del nanoporo cuando el electrodo se recarga.

El procedimiento distingue entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado en base al periodo de tiempo en el que dicho nanoporo detecta dicho nucleótido marcado.

La proporción entre el tiempo en el que el nanoporo detecta un nucleótido marcado incorporado y el tiempo en el que el nanoporo detecta un nucleótido marcado no incorporado es al menos 1,5.

También se divulga un procedimiento para la secuenciación de ácidos nucleicos que comprende: (a) proporcionar un ácido nucleico monocatenario que se va a secuenciar; (b) proporcionar una pluralidad de sondas, en el que las sondas comprenden (i) un resto de hibridación que puede hibridar con el ácido nucleico monocatenario, (ii) una estructura de bucle que tiene dos extremos, en el que cada extremo se une al resto de hibridación y (iii) un grupo escindible localizado en el resto de hibridación entre los extremos de la estructura de bucle, en el que la estructura de bucle comprende una puerta que evita que la estructura de bucle se enhebre a través de un nanoporo en una dirección inversa; (c) polimerizar la pluralidad de sondas en un orden determinado mediante hibridación de los restos de hibridación con el ácido nucleico monocatenario que se va a secuenciar; escindir los grupos escindibles para proporcionar una hebra extendida que se va a secuenciar; (d) enhebrar la hebra extendida a través de un nanoporo, en el que las puertas evitan que el hebra extendida se enhebre a través del nanoporo en la dirección inversa; y (e) detectar, con la ayuda del nanoporo, las estructuras de bucle de la hebra extendida en el orden determinado mediante hibridación de los restos de hibridación con el ácido nucleico monocatenario que se va a secuenciar, secuenciando de este modo el ácido nucleico monocatenario que se va a secuenciar.

La estructura de bucle puede comprender un segmento estrecho y la puerta es un polímero que comprende dos extremos, en la que un primer extremo se fija a la estructura de bucle contigua al segmento estrecho y un segundo extremo no se fija a la estructura de bucle, en la que la estructura de bucle se puede enhebrar a través de un nanoporo en una primera dirección donde la puerta se alinea contigua al segmento estrecho.

La estructura de bucle puede no ser capaz de enhebrarse a través del nanoporo en la dirección inversa donde la puerta no se alinea contigua al segmento estrecho.

La base de la puerta se puede emparejar con la estructura de bucle cuando la puerta no se alinea contigua al segmento estrecho.

La puerta puede comprender nucleótidos.

El segmento estrecho puede comprender nucleótidos abásicos.

El segmento estrecho puede comprender una cadena de carbonos.

Un electrodo se puede recargar entre periodos de detección.

La hebra extendida puede no enhebrarse a través del nanoporo en la dirección inversa cuando el electrodo se recarga.

Los aspectos y ventajas adicionales de la presente divulgación resultarán fácilmente evidentes para los expertos en esta técnica a partir de la siguiente descripción detallada, en la que solo se muestran y describen modos de realización ilustrativos de la presente divulgación. Como se comprenderá, la presente divulgación permite otros modos de realización diferentes, y sus diversos detalles se pueden modificar en diversos aspectos obvios, todos sin apartarse de la divulgación. En consecuencia, los dibujos y la descripción se deben considerar de naturaleza ilustrativa y no restrictiva.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Los rasgos característicos novedosos de la invención se exponen con particularidad en las reivindicaciones adjuntas. Se obtendrá una mejor comprensión de los rasgos característicos y ventajas de la presente invención con referencia a la siguiente descripción detallada que presenta modos de realización ilustrativos, en los que se usan los principios de la invención, y los dibujos adjuntos de los que:

la **figura 1** muestra esquemáticamente las etapas del procedimiento;

las **figuras 2A, 2B y 2C** muestran ejemplos de detectores de nanoporos, donde la **figura 2A** tiene el nanoporo dispuesto sobre el electrodo, la **figura 2B** tiene el nanoporo insertado en una membrana sobre un pocillo y la **figura 2C** tiene el nanoporo sobre un electrodo sobresaliente;

la **figura 3** ilustra los componentes del dispositivo y del procedimiento;

la **figura 4** ilustra un procedimiento para la secuenciación de ácidos nucleicos donde un nanoporo detecta las marcas liberadas mientras las marcas están asociadas con una polimerasa;

la **figura 5** ilustra un procedimiento para la secuenciación de ácidos nucleicos donde las marcas no se liberan tras los acontecimientos de incorporación de nucleótidos y se detectan por un nanoporo;

la **figura 6** muestra un ejemplo de una señal generada por marcas que permanecen por poco tiempo en un nanoporo;

la **figura 7** muestra una matriz de detectores de nanoporos;

la **figura 8** muestra un ejemplo de una configuración de chip que comprende un nanoporo y no un pocillo;

la **figura 9A-D** muestra un ejemplo de una configuración de matriz de celdas de chip de prueba;

la **figura 10** muestra un ejemplo de circuitos analógicos de celdas;

la **figura 11** muestra un ejemplo de un circuito de medición ultracompacto;

la **figura 12** muestra un ejemplo de un circuito de medición ultracompacto;

la **figura 13** muestra un ejemplo de una molécula marcadora unida al fosfato de un nucleótido;

la **figura 14** muestra ejemplos de localizaciones de marcas alternativas;

la **figura 15** muestra polifosfato con marca detectable y la marca detectable;

la **figura 16** muestra un sistema informático configurado para controlar un secuenciador;

la **figura 17** muestra el acoplamiento de la polimerasa del fago $\Phi 29$ a un nanoporo de hemolisina;

la **figura 18** muestra la densidad de probabilidad para el tiempo de residencia de una polimerasa que presenta dos etapas cinéticas limitantes;

la **figura 19** muestra una marca que se asocia con un compañero de unión en el lado del nanoporo más cercano a los circuitos de detección;

la **figura 20** muestra una marca de tipo espiga que circula a través del nanoporo más fácilmente de lo que circula fuera del nanoporo;

la **figura 21A-C** muestra ejemplos de formas de onda;

la **figura 22** muestra una curva de la señal extraída frente a la tensión aplicada para las cuatro bases de ácido nucleico adenina (A), citosina (C), guanina (G) y timina (T);

la **figura 23** muestra una curva de la señal extraída frente a la tensión aplicada para múltiples secuenciaciones de las cuatro bases de ácido nucleico adenina (A), citosina (C), guanina (G) y timina (T);

la **figura 24** muestra una curva del porcentaje de diferencia conductiva de referencia (% de RCD) frente a la tensión aplicada para múltiples secuenciaciones de las cuatro bases de ácido nucleico adenina (A), citosina (C), guanina (G) y timina (T);

la **figura 25** muestra el uso de obstáculos de oligonucleótidos para ralentizar la progresión de una polimerasa de ácido nucleico;

la **figura 26** muestra el uso de una segunda enzima o proteína tal como una helicasa o proteína de unión a ácido nucleico además de una polimerasa;

la **figura 27** muestra un ejemplo de un procedimiento para formar proteínas multiméricas que tienen un número definido de subunidades modificadas;

la **figura 28** muestra un ejemplo de fraccionamiento de una pluralidad de nanoporos que tienen una distribución de diferentes números de subunidades modificadas;

la **figura 29** muestra un ejemplo de fraccionamiento de una pluralidad de nanoporos que tienen una distribución de diferentes números de subunidades modificadas;

la **figura 30** muestra un ejemplo de calibración de la tensión aplicada;

la **figura 31** muestra un ejemplo de un nucleótido marcado que tiene una puerta;

la **figura 32** muestra un ejemplo de secuenciación de ácidos nucleicos usando un nucleótido marcado que tiene una puerta;

la **figura 33** muestra un ejemplo de una sonda para su uso en la secuenciación por extensión;

la **figura 34** muestra un ejemplo de sondas de extensión que se van a polimerizar;

la **figura 35** muestra un ejemplo de la escisión de los grupos escindibles para proporcionar una hebra extendida que se va a secuenciar;

la **figura 36** muestra un ejemplo del enhebrado de la hebra extendida a través de un nanoporo;

la **figura 37** muestra un ejemplo de conducción no faradaica;

la **figura 38** muestra un ejemplo de la captura de dos moléculas marcadoras;

la **figura 39** muestra un ejemplo de un complejo ternario formado entre un ácido nucleico que se va a secuenciar, un nucleótido marcado y una fusión entre un nanoporo y una polimerasa;

la **figura 40** muestra un ejemplo de la corriente que fluye a través de un nanoporo sin un nucleótido marcado presente;

la **figura 41** muestra un ejemplo del uso de niveles de corriente para distinguir entre diferentes nucleótidos marcados;

la **figura 42** muestra un ejemplo del uso de niveles de corriente para distinguir entre diferentes nucleótidos marcados;

la **figura 43** muestra un ejemplo del uso de niveles de corriente para distinguir entre diferentes nucleótidos marcados;

la **figura 44** muestra un ejemplo del uso de niveles de corriente para secuenciar una molécula de ácido nucleico usando nucleótidos marcados;

5 la **figura 45** muestra un ejemplo del uso de niveles de corriente para secuenciar una molécula de ácido nucleico usando nucleótidos marcados;

la **figura 46** muestra un ejemplo del uso de niveles de corriente para secuenciar una molécula de ácido nucleico usando nucleótidos marcados; y

10 la **figura 47** muestra un ejemplo del uso de niveles de corriente para secuenciar una molécula de ácido nucleico usando nucleótidos marcados.

DESCRIPCIÓN DETALLADA

15 Si bien en el presente documento se muestran y describen diversos modos de realización de la invención, será obvio para los expertos en la técnica que dichos modos de realización se proporcionan solo a modo de ejemplo. A los expertos en la técnica se les pueden ocurrir numerosas variaciones, cambios y sustituciones sin apartarse de la invención. Se debe entender que se pueden emplear diversas alternativas a los modos de realización de la invención descritos en el presente documento.

20 El término "nanoporo", como se usa en el presente documento, se refiere en general a un poro, canal o paso formado o proporcionado de otro modo en una membrana. Una membrana puede ser una membrana orgánica, tal como una bicapa lipídica, o una membrana sintética, tal como una membrana formada de un material polimérico. La membrana puede ser un material polimérico. El nanoporo se puede disponer contiguo o próximo a un circuito sensor o un electrodo acoplado a un circuito sensor, tal como, por ejemplo, un semiconductor complementario de óxido metálico (CMOS) o circuito de transistor de efecto campo (FET). En algunos ejemplos, un nanoporo tiene una anchura o diámetro característico del orden de 0,1 nanómetros (nm) hasta aproximadamente 1000 nm. Algunos nanoporos son proteínas. La hemolisina alfa es un ejemplo de un nanoporo proteico.

25 La expresión "ácido nucleico", como se usa en el presente documento, se refiere en general a una molécula que comprende una o más subunidades de ácido nucleico. Un ácido nucleico puede incluir una o más subunidades seleccionadas de adenosina (A), citosina (C), guanina (G), timina (T) y uracilo (U), o variantes de las mismas. Un nucleótido puede incluir A, C, G, T o U, o variantes de los mismos. Un nucleótido puede incluir cualquier subunidad que se pueda incorporar en una hebra de ácido nucleico en crecimiento. Dicha subunidad puede ser una A, C, G, T o U, o cualquier otra subunidad que sea específica para una o más A, C, G, T o U complementarias, o complementaria a una purina (es decir, A o G, o variantes de las mismas) o a una pirimidina (es decir, C, T o U, o variantes de las mismas). Una subunidad puede permitir que se resuelvan bases de ácidos nucleicos individuales o grupos de bases (por ejemplo, AA, TA, AT, GC, CG, CT, TC, GT, TG, AC, CA u homólogos de uracilo de los mismos). En algunos ejemplos, un ácido nucleico es ácido desoxirribonucleico (ADN) o ácido ribonucleico (ARN), o derivados de los mismos. Un ácido nucleico puede ser monocatenario o bicatenario.

30 El término "polimerasa", como se usa en el presente documento, se refiere en general a cualquier enzima que puede catalizar una reacción de polimerización. Los ejemplos de polimerasas incluyen, sin limitación, una polimerasa de ácidos nucleicos, una transcriptasa o una ligasa. Una polimerasa puede ser una enzima de polimerización.

Procedimientos y sistemas para secuenciar muestras

35 En el presente documento se describen procedimientos, dispositivos y sistemas para secuenciar ácidos nucleicos usando, o con la ayuda de, uno o más nanoporos. El uno o más nanoporos pueden estar en una membrana (por ejemplo, bicapa lipídica) que se dispone contigua o en proximidad de detección a un electrodo que forma parte de, o está acoplado a, un circuito integrado.

40 En algunos ejemplos, un dispositivo de nanoporos incluye un único nanoporo en una membrana que es contigua o está en proximidad de detección a un electrodo. En otros ejemplos, un dispositivo de nanoporos incluye al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 1000 o 10.000 nanoporos próximos a un circuito sensor o electrodos sensores. El uno o más nanoporos se pueden asociar con un electrodo individual y un circuito integrado sensor o una pluralidad de electrodos y circuitos integrados sensores.

45 Un sistema puede incluir una cámara de reacción que incluye uno o más dispositivos de nanoporos. Un dispositivo de nanoporos puede ser un dispositivo de nanoporos direccionables individualmente (por ejemplo, un dispositivo que puede detectar una señal y proporcionar una salida independiente de otros dispositivos de nanoporos en el sistema). Un nanoporo direccionable individualmente se puede leer individualmente. En algunos casos, un nanoporo direccionable individualmente se puede escribir individualmente. Como alternativa, un nanoporo direccionable individualmente se puede leer y escribir individualmente. El sistema puede incluir uno o más

procesadores informáticos para facilitar la preparación de muestras y diversas operaciones de la divulgación, tales como secuenciación de ácidos nucleicos. El procesador se puede acoplar al dispositivo de nanoporos.

Un dispositivo de nanoporos puede incluir una pluralidad de electrodos sensores direccionables individualmente. Cada electrodo sensor puede incluir una membrana contigua al electrodo y uno o más nanoporos en la membrana.

Los procedimientos, dispositivos y sistemas de la divulgación pueden detectar con exactitud acontecimientos de incorporación de nucleótidos individuales, tales como tras la incorporación de un nucleótido en una hebra en crecimiento que es complementaria a un molde. Una enzima (por ejemplo, ADN polimerasa, ARN polimerasa, ligasa) puede incorporar nucleótidos a una cadena polinucleotídica en crecimiento. Las enzimas (por ejemplo, polimerasas) proporcionadas en el presente documento pueden generar cadenas poliméricas.

El nucleótido añadido puede ser complementario a la hebra de ácido nucleico molde correspondiente que hibrida a la hebra en crecimiento (por ejemplo, reacción en cadena de la polimerasa (PCR)). Un nucleótido puede incluir una marca (o especie marcadora) que se acopla a cualquier localización del nucleótido que incluye, pero sin limitarse a, un fosfato (por ejemplo, fosfato gamma), glúcido o resto de base nitrogenada del nucleótido. En algunos casos, las marcas se detectan mientras están asociadas con una polimerasa durante la incorporación de marcas de nucleótidos. La marca puede continuar detectándose hasta que la marca se transloque a través del nanoporo después de la incorporación de nucleótidos y la posterior escisión y/o liberación de la marca. En algunos casos, los acontecimientos de incorporación de nucleótidos liberan marcas de los nucleótidos que pasan a través de un nanoporo y se detectan. La marca se puede liberar por la polimerasa o escindir/liberar de cualquier manera adecuada, que incluye, sin limitación, escisión por una enzima localizada cerca de la polimerasa. De esta manera, se puede identificar la base incorporada (es decir, A, C, G, T o U) porque se libera una marca única de cada tipo de nucleótido (es decir, adenina, citosina, guanina, timina o uracilo). En algunas situaciones, los acontecimientos de incorporación de nucleótidos no liberan marcas. En dicho caso, una marca acoplada a un nucleótido incorporado se detecta con la ayuda de un nanoporo. En algunos ejemplos, la marca se puede mover a través de o próxima al nanoporo y se puede detectar con la ayuda del nanoporo.

Los procedimientos y sistemas de la divulgación pueden permitir la detección de acontecimientos de incorporación de ácidos nucleicos, tales como a una resolución de al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 500, 1000, 5000, 10.000, 50.000 o 100.000 bases de ácido nucleico ("bases") dentro de un periodo de tiempo dado. En algunos ejemplos se usa un dispositivo de nanoporos para detectar acontecimientos de incorporación de ácidos nucleicos individuales, y cada acontecimiento se asocia con una base de ácido nucleico individual. En otros ejemplos se usa un dispositivo de nanoporos para detectar un acontecimiento que se asocia con una pluralidad de bases. Por ejemplo, una señal detectada por el dispositivo de nanoporos puede ser una señal combinada de al menos 2, 3, 4 o 5 bases.

En algunos casos, las marcas no pasan a través del nanoporo. Las marcas se pueden detectar por el nanoporo y salir de este sin pasar a través del nanoporo (por ejemplo, salir desde la dirección inversa a la cual la marca entró en el nanoporo). El chip se puede configurar para expulsar activamente las marcas del nanoporo.

En algunos casos, las marcas no se liberan tras acontecimientos de incorporación de nucleótidos. En algunos casos, los acontecimientos de incorporación de nucleótidos "presentan" marcas al nanoporo (es decir, sin liberar las marcas). El nanoporo puede detectar las marcas sin que se liberen. Las marcas se pueden unir a los nucleótidos mediante un conector de longitud suficiente para presentar la marca al nanoporo para su detección.

Los acontecimientos de incorporación de nucleótidos se pueden detectar en tiempo real (es decir, mientras se producen) y con la ayuda de un nanoporo. En algunos casos, una enzima (por ejemplo, ADN polimerasa) unida o próxima al nanoporo puede facilitar la circulación de una molécula de ácido nucleico a través de o contigua a un nanoporo. Un acontecimiento de incorporación de nucleótidos, o la incorporación de una pluralidad de nucleótidos, puede liberar o presentar una o más especies marcadoras (también "marcas" en el presente documento), que pueden ser detectadas por un nanoporo. La detección se puede producir mientras las marcas circulan a través de o contiguas al nanoporo, mientras las marcas residen en el nanoporo y/o mientras las marcas se presentan al nanoporo. En algunos casos, una enzima unida o próxima al nanoporo puede ayudar a detectar marcas tras la incorporación de uno o más nucleótidos.

Las marcas de la divulgación pueden ser átomos o moléculas, o una colección de átomos o moléculas. Una marca puede proporcionar una firma óptica, electroquímica, magnética o electrostática (por ejemplo, inductiva, capacitiva), que se puede detectar con la ayuda de un nanoporo.

Los procedimientos descritos en el presente documento pueden ser procedimientos de una única molécula. Es decir, la señal que se detecta se genera por una única molécula (es decir, incorporación de un único nucleótido) y no se genera a partir de una pluralidad de moléculas clonales. El procedimiento puede no requerir amplificación de ADN.

Los acontecimientos de incorporación de nucleótidos se pueden producir a partir de una mezcla que comprende una pluralidad de nucleótidos (por ejemplo, desoxirribonucleótido trifosfato (dNTP donde N es adenosina (A), citidina (C), timidina (T), guanosina (G) o uridina (U)). Los acontecimientos de incorporación de nucleótidos no se producen necesariamente a partir de una solución que comprende un único tipo de nucleótido (por ejemplo, dATP).

Los acontecimientos de incorporación de nucleótidos no se producen necesariamente a partir de soluciones alternativas de una pluralidad de nucleótidos (por ejemplo, dATP, seguida de dCTP, seguida de dGTP, seguida de dTTP, seguida de dATP). En algunos casos, una ligasa incorpora una pluralidad de nucleótidos (por ejemplo, dímeros de AA, AG, AC, AT, GA, GG, GC, GT, CA, etc.).

Procedimientos para identificar y secuenciar ácidos nucleicos

Los procedimientos para secuenciar ácidos nucleicos pueden incluir recuperar una muestra biológica que tiene el ácido nucleico que se va a secuenciar, extraer o aislar de otro modo la muestra de ácido nucleico de la muestra biológica y, en algunos casos, preparar la muestra de ácido nucleico para la secuenciación.

La **figura 1** ilustra esquemáticamente un procedimiento para secuenciar una muestra de ácido nucleico. El procedimiento comprende aislar la molécula de ácido nucleico de una muestra biológica (por ejemplo, muestra de tejido, muestra de líquido) y preparar la muestra de ácido nucleico para la secuenciación. En algunos casos, la muestra de ácido nucleico se extrae de una célula. Los ejemplos de técnicas para extraer ácidos nucleicos son usar lisozima, sonicación, extracción, altas presiones o cualquier combinación de los mismos. En algunos casos, el ácido nucleico es ácido nucleico extracelular y no requiere extracción de una célula.

En algunos casos, una muestra de ácido nucleico se puede preparar para la secuenciación mediante un procedimiento que consiste en eliminar proteínas, restos de la pared celular y otros componentes de la muestra de ácido nucleico. Hay muchos productos comerciales disponibles para lograr esto, tales como, por ejemplo, columnas de centrifugación. También se pueden usar precipitación en etanol y centrifugación.

La muestra de ácido nucleico se puede dividir (o romper) en una pluralidad de fragmentos, lo que puede facilitar la secuenciación de ácidos nucleicos, tal como con la ayuda de un dispositivo que incluye una pluralidad de nanoporos en una matriz. Sin embargo, romper la(s) molécula(s) de ácido nucleico que se va(n) a secuenciar puede no ser necesario.

En algunos casos, se determinan secuencias largas (es decir, los procedimientos de "secuenciación aleatoria" pueden no ser necesarios). Se puede determinar cualquier longitud adecuada de secuencia de ácido nucleico. Por ejemplo, se pueden secuenciar al menos aproximadamente 5, aproximadamente 10, aproximadamente 20, aproximadamente 30, aproximadamente 40, aproximadamente 50, aproximadamente 100, aproximadamente 200, aproximadamente 300, aproximadamente 400, aproximadamente 500, aproximadamente 600, aproximadamente 700, aproximadamente 800, aproximadamente 800, aproximadamente 1000, aproximadamente 1500, aproximadamente 2000, aproximadamente 2500, aproximadamente 3000, aproximadamente 3500, aproximadamente 4000, aproximadamente 4500, aproximadamente 5000, aproximadamente 6000, aproximadamente 7000, aproximadamente 8000, aproximadamente 9000, aproximadamente 10.000, aproximadamente 20.000, aproximadamente 40.000, aproximadamente 60.000, aproximadamente 80.000 o aproximadamente 100.000 bases y similares. En algunos casos se secuencian al menos 5, al menos 10, al menos 20, al menos 30, al menos 40, al menos 50, al menos 100, al menos 200, al menos 300, al menos 400, al menos 500, al menos 600, al menos 700, al menos 800, al menos 800, al menos 1000, al menos 1500, al menos 2000, al menos 2500, al menos 3000, al menos 3500, al menos 4000, al menos 4500, al menos 5000, al menos 6000, al menos 7000, al menos 8000, al menos 9000, al menos 10.000, al menos 20.000, al menos 40.000, al menos 60.000, al menos 80.000, al menos 100.000 bases y similares. En algunos casos, las bases secuenciadas son contiguas. En algunos casos, las bases secuenciadas no son contiguas. Por ejemplo, un número dado de bases se puede secuenciar de forma consecutiva. En otro ejemplo, una o más bases secuenciadas pueden estar separadas por uno o más bloques en los que la información de secuencia no se determina y/o no está disponible. Se puede secuenciar un molde múltiples veces (por ejemplo, usando un molde circular de ácido nucleico), en algunos casos generando información de secuencia redundante. En algunos casos se usa un programa informático para proporcionar la secuencia. En algunos casos, la muestra de ácido nucleico se puede dividir antes de la secuenciación. En algunos casos, la hebra de muestra de ácido nucleico se puede procesar, de modo que un ADN bicatenario dado o región de ARN/ADN dada se haga circular de modo que las partes de sentido y antisentido correspondientes del ADN bicatenario o la región de ARN/ADN se incluyan en el ADN circular o molécula circular de ADN/ARN. En dicho caso, las bases secuenciadas de dicha molécula pueden permitir una recopilación y verificación de datos más fácil de las lecturas de posición de bases.

Secuenciación de nanoporos y detección molecular

En el presente documento se describen sistemas y procedimientos para secuenciar una molécula de ácido nucleico con la ayuda de un nanoporo. El nanoporo se puede formar o incluir de otro modo en una membrana dispuesta contigua a un electrodo sensor de un circuito sensor, tal como un circuito integrado. El circuito integrado puede ser un circuito integrado específico de la aplicación (ASIC). En algunos ejemplos, el circuito integrado es un transistor

de efecto campo o un semiconductor complementario de óxido metálico (CMOS). El circuito sensor puede estar situado en un chip u otro dispositivo que tenga el nanoporo o fuera del chip o dispositivo, tal como en una configuración fuera del chip. El semiconductor puede ser cualquier semiconductor, incluyendo, sin limitación, los semiconductores del grupo IV (por ejemplo, silicio) y del grupo III-V (por ejemplo, arseniuro de galio).

En algunos casos, a medida que un ácido nucleico o marca circula a través de o contigua al nanoporo, el circuito sensor detecta una señal eléctrica asociada con el ácido nucleico o marca. El ácido nucleico puede ser una subunidad de una hebra más grande. La marca puede ser un subproducto de un acontecimiento de incorporación de nucleótidos u otra interacción entre un ácido nucleico marcado y el nanoporo o una especie contigua al nanoporo, tal como una enzima que escinde una marca de un ácido nucleico. La marca puede permanecer unida al nucleótido. Una señal detectada se puede recoger y almacenar en una localización de memoria, y usarse más tarde para construir una secuencia del ácido nucleico. La señal recogida se puede procesar para justificar cualquier anomalía en la señal detectada, tal como errores.

La **figura 2** muestra ejemplos de un detector (o sensor) de nanoporo que tiene control de la temperatura, que se puede preparar de acuerdo con los procedimientos descritos en la publicación de solicitud de patente de EE. UU. n.º 2011/0193570. Con referencia a la **figura 2A**, el detector de nanoporo comprende un electrodo superior **201** en contacto con una solución conductora (por ejemplo, solución salina) **207**. Un electrodo conductor inferior **202** está cerca, contiguo o próximo a un nanoporo **206**, que se inserta en una membrana **205**. En algunos casos, el electrodo conductor inferior **202** está incluido en un semiconductor **203** en el que están incluidos circuitos eléctricos en un sustrato semiconductor **204**. Una superficie del semiconductor **203** se puede tratar para que sea hidrófoba. Una muestra que se va a detectar atraviesa el poro del nanoporo **206**. El sensor del chip semiconductor está colocado en el paquete **208** y este, a su vez, está próximo a un elemento de control de temperatura **209**. El elemento de control de temperatura **209** puede ser un dispositivo de calefacción y/o refrigeración termoeléctrico (por ejemplo, un dispositivo Peltier). Múltiples detectores de nanoporo pueden formar una matriz de nanoporos.

Con referencia a la **figura 2B**, donde los números similares representan elementos similares, la membrana **205** se puede disponer sobre un pocillo **210**, donde el sensor **202** forma parte de la superficie del pocillo. La **figura 2C** muestra un ejemplo en el que el electrodo **202** sobresale de la superficie semiconductor tratada **203**.

En algunos ejemplos, la membrana **205** se forma en el electrodo conductor inferior **202** y no en el semiconductor **203**. En dicho caso, la membrana **205** puede formar interacciones de acoplamiento con el electrodo conductor inferior **202**. En algunos casos, sin embargo, la membrana **205** se forma en el electrodo conductor inferior **202** y en el semiconductor **203**. Como alternativa, la membrana **205** se puede formar en el semiconductor **203** y no en el electrodo conductor inferior **202**, pero se puede extender sobre el electrodo conductor inferior **202**.

Los nanoporos se pueden usar para secuenciar moléculas de ácido nucleico indirectamente, en algunos casos, con detección eléctrica. La secuenciación indirecta puede ser cualquier procedimiento donde un nucleótido incorporado en una hebra en crecimiento no pasa a través del nanoporo. La molécula de ácido nucleico puede pasar a cualquier distancia adecuada y/o en proximidad al nanoporo, en algunos casos, a una distancia tal que las marcas liberadas de los acontecimientos de incorporación de nucleótidos se detecten en el nanoporo.

El nanoporo puede detectar los subproductos de los acontecimientos de incorporación de nucleótidos. Los "acontecimientos de incorporación de nucleótidos" son la incorporación de un nucleótido en una cadena de polinucleótidos en crecimiento. Un subproducto se puede correlacionar con la incorporación de un nucleótido de un tipo dado. En general, los acontecimientos de incorporación de nucleótidos se catalizan por una enzima, tal como ADN polimerasa, y usan interacciones de pares de bases con una molécula molde para elegir entre los nucleótidos disponibles para la incorporación en cada localización.

Una muestra de ácido nucleico se puede secuenciar usando nucleótidos o análogos de nucleótidos marcados. En algunos ejemplos, un procedimiento para secuenciar una molécula de ácido nucleico comprende (a) incorporar (por ejemplo, polimerizar) nucleótidos marcados, en el que una marca asociada con un nucleótido individual se libera tras la incorporación y (b) detectar la marca liberada con la ayuda de un nanoporo. En algunos casos, el procedimiento comprende además dirigir la marca unida o liberada de un nucleótido individual a través del nanoporo. La marca liberada o unida se puede dirigir mediante cualquier técnica adecuada, en algunos casos, con la ayuda de una enzima (o motor molecular) y/o una diferencia de tensión a través del poro. De forma alternativa, la marca liberada o unida se puede dirigir a través del nanoporo sin el uso de una enzima. Por ejemplo, la marca se puede dirigir por una diferencia de tensión a través del nanoporo como se describe en el presente documento.

Secuenciación con marcas precargadas

Las marcas que se liberan sin cargarse en el nanoporo se pueden difundir lejos del nanoporo y no ser detectadas por el nanoporo. Esto puede provocar errores en la secuenciación de una molécula de ácido nucleico (por ejemplo, fallar en una posición de ácido nucleico o detectar marcas en el orden incorrecto). En el presente documento se proporcionan procedimientos para secuenciar una molécula de ácido nucleico donde las moléculas marcadoras se

"precargan" en el nanoporo antes de que la marca se libere de un nucleótido. Puede ser mucho más probable que el nanoporo detecte las marcas precargadas que las marcas que no están precargadas (por ejemplo, al menos aproximadamente 100 veces más probable). También, la precarga de marcas proporciona un enfoque para determinar si se ha incorporado o no un nucleótido marcado en una hebra de ácido nucleico en crecimiento. Las marcas que se asocian con un nucleótido incorporado se pueden asociar con el nanoporo durante un periodo de tiempo mayor (por ejemplo, un promedio de al menos aproximadamente 50 milisegundos (ms)) que las marcas que pasan a través del nanoporo (y son detectadas por él) sin ser incorporadas (por ejemplo, un promedio de menos de aproximadamente 1 ms). En algunos ejemplos, la marca asociada con un nucleótido incorporado se puede asociar con el nanoporo o mantenerse o acoplarse de otro modo a una enzima (por ejemplo, polimerasa) contigua al nanoporo durante un periodo de tiempo promedio de al menos aproximadamente 1 milisegundo (ms), 20 ms, 30 ms, 40 ms, 50 ms, 100 ms, 200 ms o más de 250 ms. En algunos ejemplos, una señal de marca asociada con un nucleótido incorporado puede tener una vida útil de detección promedio de al menos aproximadamente 1 milisegundo (ms), 20 ms, 30 ms, 40 ms, 50 ms, 100 ms, 200 ms o más de 250 ms. La marca se puede acoplar a un nucleótido incorporado acoplado. Una señal de marca que tiene una vida útil de detección promedio que es menor que una vida útil de detección promedio atribuida a un nucleótido incorporado (por ejemplo, menos de aproximadamente 1 ms) se puede atribuir a un nucleótido no incorporado acoplado a la marca. En algunos casos, una vida útil de detección promedio de al menos 'x' se puede atribuir a un nucleótido incorporado, y una vida útil de detección promedio menor que 'x' se puede atribuir a un nucleótido no incorporado. En algunos ejemplos, 'x' puede ser 0,1 ms, 1 ms, 20 ms, 30 ms, 40 ms, 50 ms, 100 ms, 1 segundo.

Se puede detectar una marca con la ayuda de un dispositivo de nanoporos que tenga al menos un nanoporo en una membrana. La marca se puede asociar con un nucleótido marcado individual durante la incorporación del nucleótido marcado individual. Los procedimientos proporcionados en el presente documento pueden implicar diferenciar, con la ayuda del nanoporo, una marca asociada con el nucleótido marcado individual de una o más marcas asociadas con uno o más nucleótidos marcados individuales no incorporados. En algunos casos, el dispositivo de nanoporos detecta una marca asociada con un nucleótido marcado individual durante la incorporación. El dispositivo de nanoporos detecta, determina o diferencia los nucleótidos marcados (ya sea incorporados en un soporte de ácido nucleico en crecimiento o no incorporados) durante un periodo de tiempo dado, en algunos casos con la ayuda de un electrodo y/o nanoporo del dispositivo de nanoporos. El periodo de tiempo dentro del cual el dispositivo de nanoporos detecta la marca puede ser más corto, en algunos casos sustancialmente más corto, que el periodo de tiempo en el que la marca y/o nucleótido acoplado a la marca es retenido por una enzima, tal como una enzima que facilita la incorporación del nucleótido en una hebra de ácido nucleico (por ejemplo, una polimerasa). En algunos ejemplos, el electrodo puede detectar la marca una pluralidad de veces dentro del periodo de tiempo en el que el nucleótido marcado incorporado está asociado con la enzima. Por ejemplo, el electrodo puede detectar la marca al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 1000, 10.000, 100.000 o 1.000.000 de veces dentro del periodo de tiempo en el que el nucleótido marcado incorporado está asociado con la enzima.

Cualquier lectura de un tiempo de detección o un tiempo promedio de detección puede permitir que una proporción de los tiempos de detección se encuentre por encima o por debajo del tiempo establecido o del tiempo promedio. En algunos casos, los tiempos de detección al secuenciar una pluralidad de bases de ácido nucleico se distribuyen estadísticamente (por ejemplo, una distribución exponencial o una distribución gaussiana). Una distribución exponencial puede tener un porcentaje relativamente grande de tiempos de detección que se encuentran por debajo del tiempo de detección promedio, como se ve en la **figura 18**, por ejemplo.

En algunos ejemplos, la precarga de una marca comprende dirigir al menos una parte de la marca a través de al menos una parte de un nanoporo mientras la marca está unida a un nucleótido, dicho nucleótido se ha incorporado a una hebra de ácido nucleico (por ejemplo, hebra de ácido nucleico en crecimiento), se está incorporando a la hebra de ácido nucleico o aún no se ha incorporado a la hebra de ácido nucleico pero se puede incorporar a la hebra de ácido nucleico. En algunos ejemplos, la precarga de la marca comprende dirigir al menos una parte de la marca a través de al menos una parte del nanoporo antes de que el nucleótido se haya incorporado a la hebra de ácido nucleico o mientras el nucleótido se está incorporando a la hebra de ácido nucleico. En algunos casos, la precarga de la marca puede incluir dirigir al menos una parte de la marca a través de al menos una parte del nanoporo después de que el nucleótido se haya incorporado a la hebra de ácido nucleico.

La **figura 3** muestra los componentes principales del procedimiento. Aquí, se forma un nanoporo **301** en una membrana **302**. Se asocia una enzima **303** (por ejemplo, una polimerasa, tal como una ADN polimerasa) con el nanoporo. En algunos casos, la enzima se une covalentemente al nanoporo como se describe a continuación. La polimerasa se asocia con una molécula de ácido nucleico monocatenario **304** que se va a secuenciar. La molécula de ácido nucleico monocatenario es circular en algunos casos, pero esto no se requiere. En algunos casos, la molécula de ácido nucleico es lineal. En algunos casos, un cebador de ácido nucleico **305** hibrida a una parte de la molécula de ácido nucleico. En algunos casos, el cebador tiene una horquilla (por ejemplo, para evitar enhebrar la hebra de ácido nucleico recién creada desplazada en el nanoporo después del primer pase alrededor de un molde circular). La polimerasa cataliza la incorporación de nucleótidos en el cebador usando una molécula de ácido nucleico monocatenario como molde. Los nucleótidos **306** comprenden especies marcadoras ("marcas") **307** como se describe en el presente documento.

La **figura 4** ilustra esquemáticamente un procedimiento para la secuenciación de ácidos nucleicos con la ayuda de marcas "precargadas". La **parte A** ilustra los componentes como se describe en la **figura 3**. La **parte C** muestra el marcador cargado en el nanoporo. Una marca "cargada" puede ser una que se sitúa en y/o permanece en o cerca del nanoporo durante una cantidad de tiempo apreciable, tal como, por ejemplo, al menos 0,1 milisegundos (ms), al menos 1 ms, al menos 5 ms, al menos 10 ms, al menos 50 ms, al menos 100 ms o al menos 500 ms o al menos 1000 ms. En algunos casos, una marca que se "precarga" se carga en el nanoporo antes de su liberación del nucleótido. En algunos casos, una marca se precarga si la probabilidad de que la marca pase a través de (y/o sea detectada por) el nanoporo después de ser liberada tras un acontecimiento de incorporación de nucleótidos es adecuadamente alta, tal como, por ejemplo, de al menos un 90 %, al menos un 95 %, al menos un 99 %, al menos un 99,5 %, al menos un 99,9 %, al menos un 99,99 % o al menos un 99,999 %.

En la transición de la **parte A** a la **parte B**, un nucleótido se ha asociado con la polimerasa. Las bases del nucleótido asociado se emparejan con la molécula de ácido nucleico monocatenario (por ejemplo, A con T y G con C). Se reconoce que un número de nucleótidos se puede asociar de forma transitoria con la polimerasa sin que sus bases estén emparejadas con la molécula de ácido nucleico monocatenario. La polimerasa puede rechazar los nucleótidos no emparejados y, en general, la incorporación del nucleótido solo continúa en el caso donde las bases de los nucleótidos están emparejadas. En general, estos nucleótidos no emparejados se rechazan dentro de una escala de tiempo que es más corta que la escala de tiempo durante la cual los nucleótidos emparejados correctamente permanecen asociados con la polimerasa. Los nucleótidos no emparejados se pueden rechazar dentro de un periodo de tiempo (media) de al menos aproximadamente 100 nanosegundos (ns), 1 ms, 10 ms, 100 ms, 1 segundo, mientras que los nucleótidos emparejados correctamente permanecen asociados con la polimerasa durante un periodo de tiempo mayor, tal como un periodo de tiempo medio de al menos aproximadamente 1 milisegundo (ms), 10 ms, 100 ms, 1 segundo o 10 segundos. La corriente que pasa a través del nanoporo durante la **parte A** y **B** de la **figura 4** puede ser entre 3 y 30 picoamperios (pA) en algunos casos.

La **figura 4, parte C**, representa el acoplamiento de la polimerasa al nanoporo. La polimerasa se puede atraer hacia el nanoporo con la ayuda de una tensión (por ejemplo, tensión de CC o CA) aplicada a una membrana en la que reside el nanoporo o al nanoporo. La **figura 17** también representa el acoplamiento de la polimerasa al nanoporo, en este caso, la ADN polimerasa del fago $\Phi 29$ (ADN polimerasa de $\Phi 29$) con un nanoporo con hemolisina alfa. La marca se puede introducir en el nanoporo durante el acoplamiento mediante una fuerza eléctrica, tal como una fuerza generada en presencia de un campo eléctrico generado por una tensión aplicada a la membrana y/o al nanoporo. En algunos casos, la corriente que fluye a través del nanoporo durante la **parte C** de la **figura 4** es de aproximadamente 6 pA, aproximadamente 8 pA, aproximadamente 10 pA, aproximadamente 15 pA o aproximadamente 30 pA. La polimerasa experimenta una isomerización y una reacción de transforilación para incorporar el nucleótido en la molécula de ácido nucleico en crecimiento y liberar la molécula marcadora.

En la **parte D**, la marca se representa pasando a través del nanoporo. Como se describe en el presente documento, el nanoporo detecta la marca. La repetición del ciclo (es decir, de la **parte A** a la **E** o de la **parte A** a la **F**) permite la secuenciación de la molécula de ácido nucleico.

En algunos casos, los nucleótidos marcados que no se incorporan en la molécula de ácido nucleico en crecimiento también pasarán a través del nanoporo, como se ve en la **parte F** de la **figura 4**. En algunos casos, el nanoporo puede detectar el nucleótido no incorporado, pero el procedimiento proporciona un medio para distinguir entre un nucleótido incorporado y un nucleótido no incorporado en base, al menos en parte, al tiempo durante el cual el nucleótido se detecta en el nanoporo. Las marcas unidas a nucleótidos no incorporados pasan a través del nanoporo rápidamente y se detectan durante un corto periodo de tiempo (por ejemplo, menos de 100 ms), mientras que las marcas unidas a nucleótidos incorporados se cargan en el nanoporo y se detectan durante un largo periodo de tiempo (por ejemplo, al menos 100 ms).

En algunos casos, el procedimiento distingue entre un nucleótido marcado incorporado (por ejemplo, polimerizado) y un nucleótido marcado no incorporado en base al periodo de tiempo en el que el nanoporo detecta el nucleótido marcado. La marca puede permanecer próxima al nanoporo durante más tiempo cuando se incorpora que cuando no se incorpora. En algunos casos, la polimerasa muta para incrementar la diferencia de tiempo entre los nucleótidos marcados incorporados y los nucleótidos marcados no incorporados. La proporción entre el tiempo en el que el nanoporo detecta un nucleótido marcado incorporado y el tiempo en el que el nanoporo detecta una marca no incorporada puede ser cualquier valor adecuado. En algunos casos, la proporción entre el tiempo en el que el nanoporo detecta un nucleótido marcado incorporado y el tiempo en el que el nanoporo detecta una marca no incorporada es de aproximadamente 1,5, aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 12, aproximadamente 14, aproximadamente 16, aproximadamente 18, aproximadamente 20, aproximadamente 25, aproximadamente 30, aproximadamente 40, aproximadamente 50, aproximadamente 100, aproximadamente 200, aproximadamente 300, aproximadamente 400, aproximadamente 500 o aproximadamente 1000. En algunos casos, la proporción entre el tiempo en el que el nanoporo detecta un nucleótido marcado incorporado (por ejemplo, polimerizado) y el tiempo en el que el nanoporo detecta una marca

no incorporada es al menos aproximadamente 1,5, al menos aproximadamente 2, al menos aproximadamente 3, al menos aproximadamente 4, al menos aproximadamente 5, al menos aproximadamente 6, al menos aproximadamente 7, al menos aproximadamente 8, al menos aproximadamente 9, al menos aproximadamente 10, al menos aproximadamente 12, al menos aproximadamente 14, al menos aproximadamente 16, al menos aproximadamente 18, al menos aproximadamente 20, al menos aproximadamente 25, al menos aproximadamente 30, al menos aproximadamente 40, al menos aproximadamente 50, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 300, al menos aproximadamente 400, al menos aproximadamente 500 o al menos aproximadamente 1000.

El tiempo en el que la marca se carga en (y/o se detecta por) el nanoporo es cualquier valor adecuado. En algunos casos, el nanoporo detecta la marca durante un promedio de aproximadamente 10 milisegundos (ms), aproximadamente 20 ms, aproximadamente 30 ms, aproximadamente 40 ms, aproximadamente 50 ms, aproximadamente 60 ms, aproximadamente 80 ms, aproximadamente 100 ms, aproximadamente 120 ms, aproximadamente 140 ms, aproximadamente 160 ms, aproximadamente 180 ms, aproximadamente 200 ms, aproximadamente 220 ms, aproximadamente 240 ms, aproximadamente 260 ms, aproximadamente 280 ms, aproximadamente 300 ms, aproximadamente 400 ms, aproximadamente 500 ms, aproximadamente 600 ms, aproximadamente 800 ms o aproximadamente 1000 ms. En algunos casos, el nanoporo detecta la marca durante un promedio de al menos aproximadamente 10 milisegundos (ms), al menos aproximadamente 20 ms, al menos aproximadamente 30 ms, al menos aproximadamente 40 ms, al menos aproximadamente 50 ms, al menos aproximadamente 60 ms, al menos aproximadamente 80 ms, al menos aproximadamente 100 ms, al menos aproximadamente 120 ms, al menos aproximadamente 140 ms, al menos aproximadamente 160 ms, al menos aproximadamente 180 ms, al menos aproximadamente 200 ms, al menos aproximadamente 220 ms, al menos aproximadamente 240 ms, al menos aproximadamente 260 ms, al menos aproximadamente 280 ms, al menos aproximadamente 300 ms, al menos aproximadamente 400 ms, al menos aproximadamente 500 ms, al menos aproximadamente 600 ms, al menos aproximadamente 800 ms o al menos aproximadamente 1000 ms.

En algunos ejemplos, una marca que genera una señal durante un periodo de tiempo de al menos aproximadamente 1 ms, al menos aproximadamente 10 ms, al menos aproximadamente 50 ms, al menos aproximadamente 80 ms, al menos aproximadamente 100 ms, al menos aproximadamente 120 ms, al menos aproximadamente 140 ms, al menos aproximadamente 160 ms, al menos aproximadamente 180 ms, al menos aproximadamente 200 ms, al menos aproximadamente 220 ms, al menos aproximadamente 240 ms, o al menos aproximadamente 260 ms se atribuye a un nucleótido que ha sido incorporado en una hebra en crecimiento que es complementaria a al menos una parte de un molde. En algunos casos, una marca que genera una señal durante un periodo de tiempo menor que aproximadamente 100 ms, menor que aproximadamente 80 ms, menor que aproximadamente 60 ms, menor que aproximadamente 40 ms, menor que aproximadamente 20 ms, menor que aproximadamente 10 ms, menor que aproximadamente 5 ms o menor que aproximadamente 1 ms se atribuye a un nucleótido que no se ha incorporado a la hebra en crecimiento.

La molécula de ácido nucleico puede ser lineal (como se ve en la **figura 5**). En algunos casos, como se ve en la **figura 3**, la molécula de ácido nucleico **304** es circular (por ejemplo, ADN circular, ARN circular). El ácido nucleico circular (por ejemplo, monocatenario) se puede secuenciar una pluralidad de veces (por ejemplo, a medida que la polimerasa **303** avanza completamente alrededor del círculo, comienza a volver a secuenciar partes del molde). El ADN circular puede ser hebras de sentido y antisentido para las mismas posiciones genómicas enlazadas juntas (en algunos casos, permitiendo lecturas más robustas y exactas). El ácido nucleico circular se puede secuenciar hasta que se logre una exactitud adecuada (por ejemplo, al menos un 95 %, al menos un 99 %, al menos un 99,9 % o al menos un 99,99 % de exactitud). En algunos casos, el ácido nucleico se secuenciar al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 12, al menos al menos 15, al menos 20 veces, al menos 40 veces, al menos 50 veces, al menos 100 veces o al menos 1000 veces.

El procedimiento y los dispositivos descritos en el presente documento pueden distinguir entre nucleótidos incorporados y nucleótidos no incorporados en base, en parte, a que los nucleótidos incorporados son detectados y/o detectables por el nanoporo durante un periodo de tiempo mayor que los nucleótidos no incorporados. En algunos casos, el desplazamiento de una segunda hebra de ácido nucleico hibridada a la hebra de ácido nucleico que se secuenciar (ácido nucleico bicatenario) incrementa la diferencia de tiempo entre la detección de nucleótidos incorporados y no incorporados. Con referencia a la **figura 3**, después de la primera secuenciación, la polimerasa puede encontrar ácido nucleico bicatenario (por ejemplo, comenzando a partir de cuando encuentra el cebador **305**) y puede ser necesario desplazar la segunda hebra de ácido nucleico del molde para continuar la secuenciación. Este desplazamiento puede ralentizar la polimerasa y/o la velocidad de acontecimientos de incorporación de nucleótidos en comparación con cuando el molde es monocatenario.

En algunos casos, la molécula de ácido nucleico molde es bicatenaria a partir de hibridaciones de oligonucleótidos al molde monocatenario. La **figura 25** muestra un ejemplo en el que se hibrida una pluralidad de oligonucleótidos **2500**. La progresión de la polimerasa **2501** en la dirección indicada **2502** desplaza los oligonucleótidos del molde. La polimerasa puede avanzar más lentamente que si los oligonucleótidos no estuvieran presentes. En algunos casos, el uso de los oligonucleótidos como se describe en el presente documento mejora la resolución entre los nucleótidos incorporados y los nucleótidos no incorporados en bases, en parte, a que los nucleótidos incorporados

son detectados y/o detectables por el nanoporo durante un periodo de tiempo mayor que los nucleótidos no incorporados. Los oligonucleótidos pueden tener cualquier longitud adecuada (por ejemplo, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 12, aproximadamente 14, aproximadamente 16, aproximadamente 18 o aproximadamente 20 bases de longitud). Los oligonucleótidos pueden comprender bases naturales (por ejemplo, adenina (A), citosina (C), guanina (G), timina (T) y/o uracilo (U)), bases universales (por ejemplo, 5-nitroindol, 3-nitropirrol, 3-metil 7-propinil isocarbostirilo (PIM), 3-metil isocarbostirilo (MICS) y/o 5-metil isocarbostirilo (5MICS)) o cualquier combinación de los mismos en cualquier proporción.

En algunos casos, las polimerasas de ácido nucleico progresan más lentamente con moldes de ácidos nucleicos metilados que con moldes de ácidos nucleicos no metilados. Los procedimientos y/o dispositivos descritos en el presente documento pueden usar ácidos nucleicos metilados y/o metilar la molécula de ácido nucleico. En algunos casos, un grupo metilo está presente y/o se añade a la posición 5 del anillo de pirimidina de la citosina y/o el nitrógeno número 6 del anillo de purina de la adenina. El ácido nucleico que se va a secuenciar se puede aislar de un organismo que metila el ácido nucleico. En algunos casos, el ácido nucleico se puede metilar *in vitro* (por ejemplo, usando una enzima de ADN metiltransferasa). El tiempo se puede usar para diferenciar las bases metiladas de las bases no metiladas. Esto puede permitir estudios de epigenética.

En algunos casos, las bases metiladas se pueden distinguir de las bases no metiladas en base a una corriente característica o una conformación característica de una curva de corriente/tiempo. Por ejemplo, un nucleótido marcado puede dar como resultado una corriente de bloqueo diferente dependiendo de si el molde de ácido nucleico tiene una base metilada en una posición dada (por ejemplo, debido a diferencias conformacionales en la polimerasa). En algunos casos, las bases C y/o A están metiladas y la incorporación de los nucleótidos marcados G y/o T correspondientes desvía la corriente eléctrica.

Enzimas para secuenciación de ácidos nucleicos

El procedimiento puede usar una enzima (por ejemplo, una polimerasa, transcriptasa o una ligasa) para secuenciar una molécula de ácido nucleico con el nanoporo y los nucleótidos marcados como se describe en el presente documento. En algunos casos, el procedimiento implica incorporar (por ejemplo, polimerizar) nucleótidos marcados con la ayuda de una polimerasa (por ejemplo, ADN polimerasa). En algunos casos, la polimerasa se ha mutado para permitir que acepte nucleótidos marcados. La polimerasa también se puede mutar para incrementar el tiempo durante el cual el nanoporo detecta la marca (por ejemplo, el tiempo de la **parte C** de la **figura 4**).

La enzima puede ser cualquier enzima que cree una hebra de ácido nucleico mediante enlace fosfato de los nucleótidos. En algunos casos, la ADN polimerasa es la polimerasa 9^N o una variante de la misma, ADN polimerasa I de *E. Coli*, ADN polimerasa del bacteriófago T4, Sequenase, ADN polimerasa Taq, polimerasa 9^N (exo-)A485L/Y409V, ADN polimerasa del fago Φ29 (ADN polimerasa de Φ29), Bct polimerasa o variantes, mutantes u homólogos de las mismas. Un homólogo puede tener cualquier porcentaje de homología adecuado, que incluye, sin limitación, al menos aproximadamente un 30 %, al menos aproximadamente un 40 %, al menos aproximadamente un 50 %, al menos aproximadamente un 60 %, al menos aproximadamente un 70 %, al menos aproximadamente un 80 %, al menos aproximadamente un 90 % o al menos aproximadamente un 95 % de identidad de secuencia.

Con referencia a la **figura 3**, la enzima **303** se puede unir al nanoporo **301**. Los procedimientos adecuados para unir la enzima al nanoporo incluyen reticulación tal como la formación de enlaces disulfuro intramoleculares. El nanoporo y la enzima también pueden ser una proteína de fusión (es decir, codificada por una única cadena polipeptídica). Los procedimientos para producir proteínas de fusión pueden incluir fusionar la secuencia codificante para la enzima en marco y contigua a la secuencia codificante para el nanoporo (sin un codón finalizador en el medio) y expresar esta secuencia de fusión a partir de un único promotor. En algunos ejemplos, la enzima **303** se puede unir o acoplar de otro modo al nanoporo **301** usando grapas moleculares o huellas de proteínas. En algunos casos, la enzima se une a través de una molécula intermedia, tal como, por ejemplo, biotina conjugada a la enzima y al nanoporo con tetrámeros de estreptavidina unidos a ambas biotinas. La enzima también se puede unir al nanoporo con un anticuerpo. En algunos casos, las proteínas que forman un enlace covalente entre sí (por ejemplo, el sistema SpyTagTM/SpyCatcherTM) se usan para unir la polimerasa al nanoporo. En algunos casos, las enzimas fosfatasa o una enzima que escinde la marca del nucleótido también se unen al nanoporo.

En algunos casos, la ADN polimerasa es la ADN polimerasa del fago Φ29. La polimerasa se puede mutar para facilitar y/o mejorar la eficacia de la polimerasa mutada para la incorporación de nucleótidos marcados en una molécula de ácido nucleico en crecimiento con respecto a la polimerasa no mutada. La polimerasa se puede mutar para mejorar la entrada del análogo de nucleótido (por ejemplo, nucleótido marcado) en la región del sitio activo de la polimerasa y/o mutar para coordinarse con los análogos de nucleótido en la región activa.

La polimerasa puede tener una región activa que tiene una secuencia de aminoácidos que es homóloga (por ejemplo, al menos un 70 %, al menos un 80 % o al menos un 90 % de posiciones de aminoácidos idénticas) a la región activa de una polimerasa que acepta análogos de nucleótidos (por ejemplo, VentA 488L).

Las mutaciones adecuadas de la ADN polimerasa del fago $\Phi 29$ incluyen, pero no se limitan a, una delección de los residuos 505-525, una delección dentro de los residuos 505-525, una mutación K135A, una mutación E375H, una mutación E375S, una mutación E375K, una mutación E375R, una mutación E375A, una mutación E375Q, una mutación E375W, una mutación E375Y, una mutación E375F, una mutación E486A, una mutación E486D, una mutación K512A y combinaciones de las mismas. En algunos casos, la ADN polimerasa comprende además una mutación L384R. Las ADN polimerasas adecuadas se describen en la publicación de patente de EE. UU. n.º 2011/0059505. En algunos casos, la polimerasa es una ADN polimerasa del fago $\Phi 29$ que tiene las mutaciones N62D, L253A, E375Y, A484E y/o K512Y.

Las mutaciones adecuadas para la polimerasa del fago $\Phi 29$ no se limitan a mutaciones que confieren una mejorada incorporación de nucleótidos marcados. Otras mutaciones (por ejemplo, sustituciones, inserciones, delecciones aminoacídicas y/o rasgos característicos exógenos) pueden conferir, sin limitación, coordinación de iones metálicos potenciada, actividad de exonucleasa reducida, velocidades de reacción reducidas en una o más etapas del ciclo cinético de la polimerasa, fracción de ramificación disminuida, selectividad de cofactor alterada, rendimiento incrementado, termoestabilidad incrementada, exactitud incrementada, velocidad incrementada, longitud de lectura incrementada, tolerancia a sales incrementada y similares con respecto a la ADN polimerasa del fago $\Phi 29$ no mutada (natural).

Las mutaciones adecuadas de la ADN polimerasa del fago $\Phi 29$ incluyen, pero no se limitan a, una mutación en la posición E375, una mutación en la posición K512 y una mutación en una o más posiciones seleccionadas del grupo que consiste en L253, A484, V250, E239, Y224, Y148, E508 y T368.

La mutación en la posición E375 puede comprender una sustitución aminoacídica seleccionada del grupo que consiste en E375Y, E375F, E375R, E375Q, E375H, E375L, E375A, E375K, E375S, E375T, E375C, E375G y E375N. En algunos casos, la mutación en la posición K512 comprende una sustitución aminoacídica seleccionada del grupo que consiste en K512Y, K512F, K512I, K512M, K512C, K512E, K512G, K512H, K512N, K512Q, K512R, K512V y K512H. En un caso, la mutación en la posición E375 comprende una sustitución E375Y y la mutación en la posición K512 comprende una sustitución K512Y.

En algunos casos, la polimerasa del fago $\Phi 29$ mutada comprende una o más sustituciones aminoacídicas seleccionadas del grupo que consiste en L253A, L253C, L253S, A484E, A484Q, A484N, A484D, A484K, V250I, V250Q, V250L, V250M, V250C, V250F, V250N, V250R, V250T, V250Y, E239G, Y224K, Y224Q, Y224R, Y148I, Y148A, Y148K, Y148F, Y148C, Y148D, Y148E, Y148G, Y148H, Y148K, Y148L, Y148M, Y148N, Y148P, Y148Q, Y148R, Y148S, Y148T, Y148V, Y148W, E508R y E508K.

En algunos casos, la ADN polimerasa del fago $\Phi 29$ comprende una mutación en una o más posiciones seleccionadas del grupo que consiste en D510, E515 y F526. Las mutaciones pueden comprender una o más sustituciones aminoacídicas seleccionadas del grupo que consiste en D510K, D510Y, D510R, D510H, D510C, E515Q, E515K, E515D, E515H, E515Y, E515C, E515M, E515N, E515P, E515R, E515S, E515T, E515V, E515A, F526L, F526Q, F526V, F526K, F526I, F526A, F526T, F526H, F526M, F526V y F526Y. Los ejemplos de ADN polimerasas que se pueden usar con los procedimientos de la divulgación se describen en la publicación de patente de EE. UU. n.º 2012/0034602.

La polimerasa puede tener un perfil de velocidad cinética que sea adecuado para la detección de las marcas por el nanoporo. El perfil de velocidad se puede referir a la velocidad global de incorporación de nucleótidos y/o a una velocidad de cualquier etapa de incorporación de nucleótidos, tal como adición de nucleótidos, isomerización enzimática tal como hacia o desde un estado cerrado, unión o liberación de cofactores, liberación del producto, incorporación de ácidos nucleicos en el ácido nucleico en crecimiento o translocación.

Los sistemas de la divulgación pueden permitir la detección de uno o más acontecimientos asociados con la secuenciación. Los acontecimientos pueden ser observables cinéticamente y/o observables no cinéticamente (por ejemplo, un nucleótido que migra a través de un nanoporo sin entrar en contacto con una polimerasa).

Una polimerasa se puede adaptar para permitir la detección de acontecimientos de secuenciación. En algunos casos, el perfil de velocidad de una polimerasa puede ser tal que una marca se cargue en (y/o se detecte por) el nanoporo durante un promedio de aproximadamente 0,1 milisegundos (ms), aproximadamente 1 ms, aproximadamente 5 ms, aproximadamente 10 ms, aproximadamente 20 ms, aproximadamente 30 ms, aproximadamente 40 ms, aproximadamente 50 ms, aproximadamente 60 ms, aproximadamente 80 ms, aproximadamente 100 ms, aproximadamente 120 ms, aproximadamente 140 ms, aproximadamente 160 ms, aproximadamente 180 ms, aproximadamente 200 ms, aproximadamente 220 ms, aproximadamente 240 ms, aproximadamente 260 ms, aproximadamente 280 ms, aproximadamente 300 ms, aproximadamente 400 ms, aproximadamente 500 ms, aproximadamente 600 ms, aproximadamente 800 ms o aproximadamente 1000 ms. En algunos casos, el perfil de velocidad de una polimerasa puede ser tal que una marca se cargue en (y/o se detecte por) el nanoporo durante un promedio de al menos aproximadamente 5 milisegundos (ms), al menos aproximadamente 10 ms, al menos aproximadamente 20 ms, al menos aproximadamente 30 ms, al menos

aproximadamente 40 ms, al menos aproximadamente 50 ms, al menos aproximadamente 60 ms, al menos aproximadamente 80 ms, al menos aproximadamente 100 ms, al menos aproximadamente 120 ms, al menos aproximadamente 140 ms, al menos aproximadamente 160 ms, al menos aproximadamente 180 ms, al menos aproximadamente 200 ms, al menos aproximadamente 220 ms, al menos aproximadamente 240 ms, al menos aproximadamente 260 ms, al menos aproximadamente 280 ms, al menos aproximadamente 300 ms, al menos aproximadamente 400 ms, al menos aproximadamente 500 ms, al menos aproximadamente 600 ms, al menos aproximadamente 800 ms o al menos aproximadamente 1000 ms. En algunos casos, el nanoporo detecta la marca durante un promedio de entre aproximadamente 80 ms y 260 ms, entre aproximadamente 100 ms y 200 ms o entre aproximadamente 100 ms y 150 ms.

En algunos casos, la reacción de la polimerasa presenta dos etapas cinéticas que se producen a partir de un producto intermedio en el que un nucleótido o un producto de polifosfato está unido a la enzima polimerasa, y dos etapas cinéticas que se producen a partir de un producto intermedio en el que el nucleótido y el producto de polifosfato no están unidos a la enzima polimerasa. Las dos etapas cinéticas pueden incluir isomerización enzimática, incorporación de nucleótidos y liberación del producto. En algunos casos, las dos etapas cinéticas son translocación de moldes y unión de nucleótidos.

La **figura 18** ilustra que, cuando está presente una etapa cinética, puede haber una probabilidad exponencialmente decreciente de un tiempo de residencia dado de la marca en el nanoporo a medida que el tiempo de residencia se incrementa **1800**, proporcionando una distribución en la que hay una probabilidad relativamente alta de que el tiempo de residencia de la marca en el nanoporo sea corto **1801** (y, por lo tanto, de que potencialmente el nanoporo no la detecte). La **figura 18** también ilustra que, para el caso en el que hay dos o más etapas cinéticas (por ejemplo, etapas observables o "lentas") **1802**, la probabilidad de tiempos de residencia muy rápidos de la marca en el nanoporo es relativamente baja **1803** en comparación con el caso que tiene una etapa lenta **1801**. Dicho de otra manera, la adición de dos funciones exponenciales puede dar como resultado una función o distribución gaussiana **1802**. Además, la distribución de probabilidad para dos etapas lentas presenta un pico en la curva de densidad de probabilidad frente al tiempo de residencia **1802**. Este tipo de distribución del tiempo de residencia puede ser ventajoso para la secuenciación de ácidos nucleicos como se describe en el presente documento (por ejemplo, cuando se desea detectar una alta proporción de marcas incorporadas). Un número relativamente mayor de acontecimientos de incorporación de nucleótidos cargan la marca en el nanoporo durante un periodo de tiempo mayor que un tiempo mínimo ($T_{\min}$, que puede ser mayor de 100 ms en algunos casos).

En algunos casos, la ADN polimerasa del fago $\Phi 29$ se muta con respecto a la enzima natural para proporcionar dos etapas cinéticamente lentas y/o para proporcionar un perfil de velocidad que sea adecuado para que el nanoporo detecte marcas. En algunos casos, la ADN polimerasa del fago $\Phi 29$ tiene al menos una sustitución aminoacídica o una combinación de sustituciones seleccionadas de la posición 484, la posición 198 y la posición 381. En algunos casos, las sustituciones aminoacídicas se seleccionan de E375Y, K512Y, T368F, A484E, A484Y, N387L, T372Q, T372L, K478Y, I370W, F198W, L381A y cualquier combinación de las mismas. Las ADN polimerasas adecuadas se describen en la patente de EE. UU. n.º 8.133.672.

La cinética de la enzima también se puede ver afectada y/o controlada al manipular el contenido de la solución en contacto con la enzima. Por ejemplo, los iones divalentes no catalíticos (por ejemplo, iones que no promueven la función polimerasa tal como el estroncio (Sr^{2+})) se pueden mezclar con iones divalentes catalíticos (por ejemplo, iones que promueven la función polimerasa como el magnesio (Mg^{2+}) y/o manganeso (Mn^{2+})) para ralentizar la polimerasa. La proporción entre iones catalíticos y no catalíticos puede ser cualquier valor adecuado, que incluye aproximadamente 20, aproximadamente 15, aproximadamente 10, aproximadamente 9, aproximadamente 8, aproximadamente 7, aproximadamente 6, aproximadamente 5, aproximadamente 4, aproximadamente 3, aproximadamente 2, aproximadamente 1, aproximadamente 0,5, aproximadamente 0,2 o aproximadamente 0,1. En algunos casos, la proporción depende de la concentración de sal monovalente (por ejemplo, cloruro de potasio (KCl)), temperatura y/o pH. En un ejemplo, la solución comprende Mg^{2+} 1 micromolar y Sr^{2+} 0,25 micromolar. En otro ejemplo, la solución comprende Mg^{2+} 3 micromolar y Sr^{2+} 0,7 micromolar. La concentración de magnesio (Mg^{2+}) y manganeso (Mn^{2+}) puede ser cualquier valor adecuado y puede variar para afectar a la cinética de la enzima. En un ejemplo, la solución comprende Mg^{2+} 1 micromolar y Mn^{2+} 0,25 micromolar. En otro ejemplo, la solución comprende Mg^{2+} 3 micromolar y Mg^{2+} 0,7 micromolar.

Secuenciación de nanoporos de moléculas marcadoras precargadas

Las marcas se pueden detectar sin liberarse de los nucleótidos incorporados durante la síntesis de una hebra de ácido nucleico que es complementaria a una hebra diana. Las marcas se pueden unir a los nucleótidos con un conector de modo que la marca se presente al nanoporo (por ejemplo, la marca cuelga hacia abajo o se extiende de otro modo a través de al menos una parte del nanoporo). La longitud del conector puede ser lo suficientemente larga para permitir que la marca se extienda hacia o a través de al menos una parte del nanoporo. En algunos casos, la marca se presenta a (es decir, se mueve hacia) el nanoporo por una diferencia de tensión. También pueden ser adecuadas otras maneras de presentar la marca en el poro (por ejemplo, uso de enzimas, imanes, campos eléctricos, presión diferencial). En algunos casos, no se aplica una fuerza activa a la marca (es decir, la marca se difunde en el nanoporo).

Un aspecto de la invención proporciona un procedimiento para secuenciar un ácido nucleico. El procedimiento comprende incorporar (por ejemplo, polimerizar) nucleótidos marcados. Un nanoporo puede detectar una marca asociada con un nucleótido individual sin que se libere del nucleótido tras la incorporación.

Un chip para secuenciar una muestra de ácido nucleico puede comprender una pluralidad de nanoporos direccionables individualmente. Un nanoporo direccionable individualmente de la pluralidad puede contener al menos un nanoporo formado en una membrana dispuesta contigua a un circuito integrado. Cada nanoporo direccionable individualmente puede ser capaz de detectar una marca asociada con un nucleótido individual. El nucleótido se puede incorporar (por ejemplo, polimerizar) y la marca puede no liberarse del nucleótido tras la incorporación.

Un ejemplo del procedimiento se representa en la **figura 5**. Aquí, la hebra de ácido nucleico **500** pasa a lo largo de o próxima a (pero no a través como se indica por la flecha en **501**) el nanoporo **502**. Una enzima **503** (por ejemplo, ADN polimerasa) extiende una hebra de ácido nucleico en crecimiento **504** incorporando un nucleótido a la vez usando una primera molécula de ácido nucleico como molde **500** (es decir, la enzima cataliza acontecimientos de incorporación de nucleótidos). El nanoporo **502** detecta las marcas. Las marcas pueden residir en el nanoporo durante un periodo de tiempo.

La enzima **503** se puede unir al nanoporo **502**. Los procedimientos adecuados para unir la enzima al nanoporo incluyen reticulación tal como la formación de enlaces disulfuro intramoleculares y/o la creación de una proteína de fusión como se describe anteriormente. En algunos casos, las enzimas fosfatasa también se unen al nanoporo. Estas enzimas se pueden unir además a los fosfatos restantes en una marca escindida y producir señales más claras al incrementar además el tiempo de permanencia en el nanoporo. Las ADN polimerasas adecuadas incluyen la ADN polimerasa del fago $\Phi 29$ (ADN polimerasa de $\Phi 29$) e incluyen además, pero sin limitarse a, las descritas anteriormente.

Con referencia continua a la **figura 5**, la enzima se extrae de un conjunto de nucleótidos (círculos rellenos en la indicación **505**) unidos a moléculas marcadoras (círculos abiertos en la indicación **505**). Cada tipo de nucleótido se une a una molécula marcadora diferente, de modo que, cuando las marcas residen en el nanoporo **502**, se pueden diferenciar entre sí en base a la señal generada en o asociada con el nanoporo.

En algunos casos, las marcas se presentan al nanoporo tras acontecimientos de incorporación de nucleótidos y se liberan del nucleótido. En algunos casos, las marcas liberadas pasan a través del nanoporo. En algunos casos, las marcas no pasan a través del nanoporo. En algunos casos, una marca que se ha liberado tras un acontecimiento de incorporación de nucleótidos se distingue de una marca que puede circular a través del nanoporo, pero que no se ha liberado tras un acontecimiento de incorporación de nucleótidos, al menos en parte, durante el tiempo de permanencia en el nanoporo. En algunos casos, las marcas que permanecen en el nanoporo durante al menos aproximadamente 100 milisegundos (ms) se liberan tras acontecimientos de incorporación de nucleótidos y las marcas que permanecen en el nanoporo durante menos de 100 ms no se liberan tras acontecimientos de incorporación de nucleótidos. En algunos casos, las marcas se pueden capturar y/o guiar a través del nanoporo por una segunda enzima o proteína (por ejemplo, una proteína de unión a ácido nucleico). La segunda enzima puede escindir una marca tras (por ejemplo, durante o después de) la incorporación de nucleótidos. Se puede escindir un conector entre la marca y el nucleótido.

Como se ve en la **figura 26**, la segunda enzima o proteína **2600** se puede unir a la polimerasa. En algunos casos, la segunda enzima o proteína es una helicasa de ácido nucleico que facilita la disociación de molde bicatenario a molde monocatenario. En algunos casos, la segunda enzima o proteína no se une a la polimerasa. La segunda enzima o proteína puede ser una proteína de unión a ácido nucleico que se une al molde de ácido nucleico monocatenario para ayudar a mantener el molde monocatenario. Las proteínas de unión a ácido nucleico se pueden deslizar a lo largo de la molécula de ácido nucleico monocatenario.

Los nucleótidos incorporados se pueden diferenciar de los nucleótidos no incorporados en base al periodo de tiempo en el que se detecta una marca asociada con un nucleótido con la ayuda del nanoporo. En algunos ejemplos, una marca asociada con un nucleótido que se ha incorporado en una hebra de ácido nucleico ("nucleótido incorporado") se detecta por, o con la ayuda de, el nanoporo durante un periodo de tiempo promedio de al menos aproximadamente 5 milisegundos (ms), 10 ms, 20 ms, 30 ms, 40 ms, 50 ms, 60 ms, 70 ms, 80 ms, 90 ms, 100 ms, 200 ms, 300 ms, 400 ms o 500 ms. El nanoporo detecta una marca asociada con un nucleótido no incorporado (por ejemplo, de circulación libre) durante un periodo de tiempo promedio menor que aproximadamente 500 ms, 400 ms, 300 ms, 200 ms, 100 ms, 90 ms, 80 ms, 70 ms, 60 ms, 50 ms, 40 ms, 30 ms, 20 ms, 10 ms, 5 ms o 1 ms. En algunas situaciones, el nanoporo detecta una marca asociada con un nucleótido incorporado durante un periodo de tiempo promedio de al menos aproximadamente 100 ms, y el nanoporo detecta una marca asociada con un nucleótido no incorporado durante un periodo de tiempo promedio menor que 100 ms.

En algunos ejemplos, una marca que se acopla a un nucleótido incorporado se distingue de una marca asociada con un nucleótido que no se ha incorporado en una hebra complementaria en crecimiento en base al tiempo de

residencia de la marca en el nanoporo o a una señal detectada del nucleótido no incorporado con la ayuda del nanoporo. Un nucleótido no incorporado puede generar una señal (por ejemplo, diferencia de tensión, corriente) que es detectable durante un periodo de tiempo de entre aproximadamente 1 nanosegundo (ns) y 100 ms, o entre aproximadamente 1 ns y 50 ms, mientras que un nucleótido incorporado puede generar una señal con una vida útil de entre aproximadamente 50 ms y 500 ms, o 100 ms y 200 ms. En algunos ejemplos, un nucleótido no incorporado puede generar una señal que es detectable durante un periodo de tiempo de entre aproximadamente 1 ns y 10 ms, o 1 ns y 1 ms. En algunos casos, una marca no incorporada es detectable por un nanoporo durante un periodo de tiempo (promedio) que es mayor que el periodo de tiempo en el que una marca incorporada es detectable por el nanoporo.

En algunos casos, los nucleótidos incorporados se detectan y/o son detectables por el nanoporo durante un periodo de tiempo más corto que un nucleótido no incorporado. La diferencia y/o proporción entre estos tiempos se puede usar para determinar si un nucleótido detectado por el nanoporo está o no incorporado, como se describe en el presente documento.

El periodo de detección se puede basar en la libre circulación del nucleótido a través del nanoporo; un nucleótido no incorporado puede permanecer en o próximo al nanoporo durante un periodo de tiempo de entre aproximadamente 1 nanosegundo (ns) y 100 ms, o entre aproximadamente 1 ns y 50 ms, mientras que un nucleótido incorporado puede permanecer en o próximo al nanoporo durante un tiempo de entre aproximadamente 50 ms y 500 ms, o 100 ms y 200 ms. Los periodos de tiempo pueden variar en base a las condiciones de procesamiento; sin embargo, un nucleótido incorporado puede tener un tiempo de permanencia que es mayor que el de un nucleótido no incorporado.

La polimerización (por ejemplo, la incorporación) y la detección se pueden realizar sin interferir entre sí. En algunos casos, la polimerización de un primer nucleótido marcado no interfiere de forma apreciable en la detección por el nanoporo de una marca asociada con un segundo nucleótido marcado. En algunos casos, la detección por el nanoporo de una marca asociada con un primer nucleótido marcado no interfiere en la polimerización de un segundo nucleótido marcado. En algunos casos, la marca es lo suficientemente larga para ser detectada por el nanoporo y/o para ser detectada sin evitar acontecimientos de incorporación de nucleótidos.

Una marca (o especie marcadora) puede incluir un átomo o molécula detectable, o una pluralidad de átomos o moléculas detectables. En algunos casos, una marca incluye uno o más de adenina, guanina, citosina, timina, uracilo o un derivado de los mismos unido en cualquier posición que incluye un grupo fosfato, glúcido o una base nitrogenada de una molécula de ácido nucleico. En algunos ejemplos, una marca incluye uno o más de adenina, guanina, citosina, timina, uracilo o un derivado de los mismos unido covalentemente a un grupo fosfato de una base de ácido nucleico.

Una marca puede tener una longitud de al menos aproximadamente 0,1 nanómetros (nm), 1 nm, 2 nm, 3 nm, 4 nm, 5 nm, 6 nm, 7 nm, 8 nm, 9 nm, 10 nm, 20 nm, 30 nm, 40 nm, 50 nm, 60 nm, 70 nm, 80 nm, 90 nm, 100 nm, 200 nm, 300 nm, 400 nm, 500 nm o 1000 nm.

Una marca puede incluir una cola de subunidades repetitivas, tal como una pluralidad de adenina, guanina, citosina, timina, uracilo o un derivado de los mismos. Por ejemplo, una marca puede incluir una parte de cola que tiene al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 1000, 10.000 o 100.000 subunidades de adenina, guanina, citosina, timina, uracilo o un derivado de los mismos. Las subunidades se pueden enlazar entre sí, y enlazarse en un extremo terminal a un grupo fosfato del ácido nucleico. Otros ejemplos de partes de marca incluyen cualquier material polimérico, tal como polietilenglicol (PEG), polisulfonatos, aminoácidos o cualquier polímero cargado positiva o negativamente total o parcialmente o sin cargar.

Una especie marcadora puede tener una firma electrónica que es única para el tipo de molécula de ácido nucleico que se incorpora durante la incorporación. Por ejemplo, una base de ácido nucleico que es adenina, guanina, citosina, timina o uracilo puede tener una especie marcadora que tiene una o más especies que son únicas para adenina, guanina, citosina, timina o uracilo, respectivamente.

La **figura 6** muestra un ejemplo de diferentes señales que se generan por diferentes marcas a medida que las detecta el nanoporo. Se detectan cuatro intensidades de señal diferentes (**601**, **602**, **603** y **604**). Estas pueden corresponder a cuatro marcas diferentes. Por ejemplo, la marca presentada al nanoporo y/o liberada por la incorporación de adenosina (A) puede generar una señal con una amplitud **601**. Una marca presentada al nanoporo y/o liberada por la incorporación de citosina (C) puede generar una señal con una amplitud mayor **603**; una marca presentada al nanoporo y/o liberada por la incorporación de guanina (G) puede generar una señal con una amplitud incluso mayor **604**; y una marca presentada al nanoporo y/o liberada por la incorporación de timina (T) puede generar una señal con una amplitud aún mayor **602**. La **figura 6** también muestra un ejemplo de la detección de moléculas marcadoras que se han liberado del nucleótido y/o se presentan al nanoporo tras acontecimientos de incorporación de nucleótidos. Los procedimientos descritos en el presente documento pueden distinguir entre una marca insertada en un nanoporo y escindida posteriormente (véase, por ejemplo, la **figura 4, D**) y una marca no escindida que circula libremente (véase, por ejemplo, la **figura 4, F**).

Los procedimientos proporcionados en el presente documento pueden ser capaces de distinguir entre una marca liberada (o escindida) y una marca no liberada (o no escindida) con una exactitud de al menos aproximadamente un 50 %, al menos aproximadamente un 60 %, al menos aproximadamente un 70 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 99 %, al menos aproximadamente un 99,5 %, al menos aproximadamente un 99,9 %, al menos aproximadamente un 99,95 % o al menos aproximadamente un 99,99 % o al menos aproximadamente un 99,999 % o al menos aproximadamente un 99,9999 %.

Con referencia a la **figura 6**, la magnitud de la corriente se puede reducir en cualquier cantidad adecuada por la marca, incluyendo aproximadamente un 5 %, aproximadamente un 10 %, aproximadamente un 15 %, aproximadamente un 20 %, aproximadamente un 25 %, aproximadamente un 30 %, aproximadamente un 40 %, aproximadamente un 50 %, aproximadamente un 60 %, aproximadamente un 70 %, aproximadamente un 80 %, aproximadamente un 90 %, aproximadamente un 95 % o aproximadamente un 99 %. En algunos casos, la magnitud de la corriente se reduce en al menos un 5 %, al menos un 10 %, al menos un 15 %, al menos un 20 %, al menos un 25 %, al menos un 30 %, al menos un 40 %, al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 %, al menos un 95 % o al menos un 99 %. En algunos casos, la magnitud de la corriente se reduce como máximo un 5 %, como máximo un 10 %, como máximo un 15 %, como máximo un 20 %, como máximo un 25 %, como máximo un 30 %, como máximo un 40 %, como máximo un 50 %, como máximo un 60 %, como máximo un 70 %, como máximo un 80 %, como máximo un 90 %, como máximo un 95 % o como máximo un 99 %.

El procedimiento puede comprender además detectar periodos de tiempo entre la incorporación de los nucleótidos marcados individuales (por ejemplo, periodos **605** en la **figura 6**). Los periodos de tiempo entre la incorporación de los nucleótidos marcados individuales pueden tener una gran magnitud de corriente. En algunos casos, la magnitud de la corriente que fluye a través del nanoporo entre los acontecimientos de incorporación de nucleótidos es (por ejemplo, vuelve a) aproximadamente un 50 %, aproximadamente un 60 %, aproximadamente un 70 %, aproximadamente un 80 %, aproximadamente un 90 %, aproximadamente un 95 % o aproximadamente un 99 % de la corriente máxima (por ejemplo, cuando no hay una marca presente). En algunos casos, la magnitud de la corriente que fluye a través del nanoporo entre los acontecimientos de incorporación de nucleótidos es al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 %, al menos un 95 % o al menos un 99 % de la corriente máxima. La detección y/u observación de la corriente durante periodos de tiempo entre la incorporación de los nucleótidos marcados individuales puede mejorar la exactitud de la secuenciación en algunos casos (por ejemplo, cuando se secuencian tramos repetitivos de un ácido nucleico tal como 3 o más de la misma base de forma consecutiva). Los periodos entre los acontecimientos de incorporación de nucleótidos se pueden usar como una señal de reloj que proporciona la longitud de la molécula de ácido nucleico o segmento de la misma que se secuencia.

Los procedimientos descritos en el presente documento pueden distinguir entre un nucleótido marcado incorporado (por ejemplo, polimerizado) y un nucleótido marcado no polimerizado (por ejemplo, **506** y **505** en la **figura 5**). En algunos ejemplos, un nucleótido marcado incorporado se puede distinguir de un nucleótido marcado no incorporado con una exactitud de al menos aproximadamente un 50 %, al menos aproximadamente un 60 %, al menos aproximadamente un 70 %, al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 99 %, al menos aproximadamente un 99,5 %, al menos aproximadamente un 99,9 %, al menos aproximadamente un 99,95 % o al menos aproximadamente un 99,99 % o al menos aproximadamente un 99,999 % o al menos aproximadamente un 99,9999 %.

Asociaciones entre marcas y nanoporos

El procedimiento y los dispositivos descritos en el presente documento pueden distinguir entre los nucleótidos marcados que se incorporan a una molécula de ácido nucleico y los nucleótidos marcados no incorporados en base, en parte, a la cantidad de tiempo (o proporción de veces) que la marca está asociada con y/o es detectable por el nanoporo. En algunos casos, la interacción entre el nucleótido y la polimerasa incrementa la cantidad de tiempo que la marca está asociada con y/o es detectable por el nanoporo. En algunos casos, la marca interactúa y/o se asocia con el nanoporo.

La marca puede ser relativamente más fácil de insertar en el nanoporo que de eliminarla. En algunos casos, la marca entra en el nanoporo más rápidamente y/o con menos fuerza en comparación con la marca que sale del nanoporo. Una vez asociada con el nanoporo, la marca puede pasar a través del nanoporo más rápidamente y/o con menos fuerza en comparación con la marca que sale del nanoporo desde la dirección en la que entró en el nanoporo.

La asociación entre la marca y el nanoporo puede ser cualquier fuerza o interacción adecuada, tal como un enlace no covalente, un enlace covalente que puede ser reversible, fuerzas electrostáticas o electrodinámicas o cualquier combinación de los mismos. En algunos casos, la marca se diseña para interactuar con el nanoporo, el nanoporo

se muta o diseña para interactuar con la marca, o tanto la marca como el nanoporo se diseñan o seleccionan para formar una asociación entre sí.

La asociación entre la marca y el nanoporo puede ser cualquier fuerza adecuada. En algunos casos, la asociación es lo suficientemente fuerte, de modo que los electrodos se puedan recargar sin expulsar la marca del nanoporo. La polaridad de la tensión a través del nanoporo se puede invertir para recargar el electrodo e invertir de nuevo para detectar la marca sin que la marca abandone el nanoporo en algunos casos.

La **figura 19** muestra un ejemplo en el que una parte de marca de un nucleótido marcado **1902** se une a y/o interactúa con un compañero de afinidad (por ejemplo, molécula de afinidad) o compañero de unión **1903** en el lado del nanoporo **1901** opuesto a la polimerasa. La molécula de afinidad o el compañero de unión **1903** pueden estar separados del nanoporo **1901** pero unido al nanoporo o, como alternativa, puede formar parte del nanoporo **1901**. El compañero de unión se puede unir a cualquier superficie adecuada, tal como el nanoporo o la membrana. En algunos ejemplos se puede usar cualquier combinación adecuada de molécula marcadora y compañero de unión. En algunos casos, la molécula marcadora y el compañero de unión comprenden moléculas de ácido nucleico que hibridan entre sí. En algunos casos, la molécula marcadora y el compañero de unión comprenden estreptavidina y biotina que se unen entre sí. Como alternativa, el compañero de unión **1903** puede ser parte del nanoporo **1901**.

La **figura 20** muestra un ejemplo en el que el nucleótido marcado comprende una parte de nucleótido **2001** y una parte de marca **2002** donde la parte de marca es de tipo espiga. La parte de la marca se conforma (por ejemplo, se le da forma de espiga) de modo que la marca circule a través del nanoporo **2003** más fácilmente (por ejemplo, más rápidamente y/o con menos fuerza) que cuando se retira del nanoporo en la dirección en la que entró en el nanoporo. En un caso, la parte de marca comprende un ácido nucleico monocatenario y las bases (por ejemplo, A, C, T, G) se conectan a la cadena principal de la marca de ácido nucleico en un ángulo que apunta hacia la parte de nucleótido **2001** del nucleótido marcado (es decir, se le da forma de espiga). Como alternativa, el nanoporo puede incluir una aleta u otra obstrucción que permite la circulación de una parte de marca a lo largo de una primera dirección (por ejemplo, fuera del nanoporo) y evita la circulación de la parte de marca a lo largo de una segunda dirección (por ejemplo, una dirección opuesta a la segunda dirección). La aleta puede ser cualquier obstrucción articulada.

La marca se puede diseñar o seleccionar (por ejemplo, usando evolución dirigida) para unirse a y/o asociarse con el nanoporo (por ejemplo, en la parte de poro del nanoporo). En algunos casos, la marca es un péptido que tiene una disposición de residuos de aminoácidos hidrófilos, hidrófobos, cargados positivamente y cargados negativamente que se unen al nanoporo. En algunos casos, la marca es un ácido nucleico que tiene una disposición de bases que se unen al nanoporo.

El nanoporo se puede mutar para asociarse con la molécula marcadora. Por ejemplo, el nanoporo se puede diseñar o seleccionar (por ejemplo, usando evolución dirigida) para tener una disposición de residuos de aminoácidos hidrófilos, hidrófobos, cargados positivamente y cargados negativamente que se unen a la molécula marcadora. Los residuos de aminoácidos pueden estar en el vestíbulo y/o en el poro del nanoporo.

Expulsión de marcas de un nanoporo

Esta divulgación proporciona procedimientos en los que se expulsa una molécula marcadora de un nanoporo. Por ejemplo, un chip se puede adaptar para expulsar una molécula marcadora en los casos donde la marca reside en el nanoporo o se presenta al nanoporo tras acontecimientos de incorporación de nucleótidos, tal como, por ejemplo, durante la secuenciación. La marca se puede expulsar en la dirección opuesta a la que entró en el nanoporo (por ejemplo, sin que la marca pase a través del nanoporo), por ejemplo, la marca se puede dirigir al nanoporo desde una primera abertura y ser expulsada del nanoporo desde una segunda abertura que es diferente de la primera abertura. De forma alternativa, la marca se puede expulsar desde la abertura por la que entró en el nanoporo, por ejemplo, la marca se puede dirigir al nanoporo desde una primera abertura y ser expulsada del nanoporo desde la primera abertura.

La descripción proporciona un chip para secuenciar una muestra de ácido nucleico, comprendiendo el chip una pluralidad de nanoporos direccionables individualmente, un nanoporo direccionable individualmente de la pluralidad que tiene al menos un nanoporo formado en una membrana dispuesta contigua a un circuito integrado, cada uno de los nanoporos direccionables individualmente se adapta para expulsar una molécula marcadora del nanoporo. En algunos casos, el chip se adapta para expulsar (o el procedimiento expulsa) la marca en la dirección desde la que la marca entró en el nanoporo. En algunos casos, el nanoporo expulsa la molécula marcadora con un impulso de tensión o una serie de impulsos de tensión. Un impulso de tensión puede tener una duración de aproximadamente 1 nanosegundo a 1 minuto, o de 10 nanosegundos a 1 segundo.

El nanoporo se puede adaptar para expulsar (o el procedimiento expulsa) la molécula marcadora dentro de un periodo de tiempo de modo que dos moléculas marcadoras no estén presentes en el nanoporo al mismo tiempo. La probabilidad de que dos moléculas estén presentes en el nanoporo al mismo tiempo es como máximo un 1 %.

como máximo un 0,5 %, como máximo un 0,1 %, como máximo un 0,05 % o como máximo un 0,01 % en algunos casos.

En algunos casos, el nanoporo se adapta para expulsar la molécula marcadora en un plazo de (en un periodo de tiempo inferior a) aproximadamente 0,1 ms, 0,5 ms, 1 ms, 5 ms, 10 ms o 50 ms desde que la marca entró en el nanoporo.

Se puede expulsar una marca de un nanoporo usando un potencial eléctrico (o tensión). En algunos casos, la tensión puede tener una polaridad opuesta a la usada para llevar la marca al nanoporo. La tensión se puede aplicar con la ayuda de una forma de onda de corriente alterna (CA) que tiene un ciclo de al menos aproximadamente 1 nanosegundo, 10 nanosegundos, 100 nanosegundos, 500 nanosegundos, 1 microsegundo, 100 microsegundos, 1 milisegundo (ms), 5 ms, 10 ms, 20 ms, 30 ms, 40 ms, 50 ms, 100 ms, 200 ms, 300 ms, 400 ms, 500 ms, 600 ms, 700 ms, 800 ms, 900 ms, 1 segundo, 2 segundos, 3 segundos, 4 segundos, 5 segundos, 6 segundos, 7 segundos, 8 segundos, 9 segundos, 10 segundos, 100 segundos, 200 segundos, 300 segundos, 400 segundos, 500 segundos o 1000 segundos.

Formas de onda de corriente alterna (CA)

La secuenciación de una molécula de ácido nucleico pasando la hebra de ácido nucleico a través de un nanoporo puede requerir la aplicación de una corriente continua (por ejemplo, de modo que la dirección en la que la molécula se mueve a través del nanoporo no se invierta). Sin embargo, hacer funcionar un sensor de nanoporos durante largos periodos de tiempo usando una corriente continua puede cambiar la composición del electrodo, desequilibrar las concentraciones iónicas en el nanoporo y tener otros efectos indeseables. La aplicación de una forma de onda de corriente alterna (CA) puede evitar estos efectos indeseables y tener determinadas ventajas como se describe a continuación. Los procedimientos de secuenciación de ácidos nucleicos descritos en el presente documento que usan nucleótidos marcados son totalmente compatibles con las tensiones aplicadas de CA y, por lo tanto, se pueden usar para lograr dichas ventajas.

La capacidad de recargar el electrodo durante el ciclo de detección puede ser ventajosa cuando se usan electrodos de protección o electrodos que cambian el carácter molecular en las reacciones de conducción de corriente (por ejemplo, electrodos que comprenden plata) o electrodos que cambian el carácter molecular en reacciones de conducción de corriente. Un electrodo se puede agotar durante un ciclo de detección, aunque en algunos casos el electrodo puede no agotarse durante el ciclo de detección. La recarga puede evitar que el electrodo alcance un límite de agotamiento dado, tal como agotarse por completo, lo que puede ser un problema cuando los electrodos son pequeños (por ejemplo, cuando los electrodos son lo suficientemente pequeños para proporcionar una matriz de electrodos que tiene al menos 500 electrodos por milímetro cuadrado). En algunos casos, la vida útil del electrodo se escala y depende, al menos en parte, de la anchura del electrodo.

En algunos casos, el electrodo es poroso y/o "esponjoso". Un electrodo poroso puede tener una capacidad potenciada de la doble capa para el líquido bruto en comparación con un electrodo no poroso. El electrodo poroso se puede formar galvanizando un metal (por ejemplo, un metal noble) sobre una superficie en presencia de detergente. El metal que se galvaniza puede ser cualquier metal adecuado. El metal puede ser un metal noble (por ejemplo, paladio, plata, osmio, iridio, platino, plata u oro). En algunos casos, la superficie es una superficie metálica (por ejemplo, paladio, plata, osmio, iridio, platino, plata u oro). En algunos casos, la superficie tiene aproximadamente 5 micrómetros de diámetro y es lisa. El detergente puede crear espacios intersticiales a escala nanométrica en la superficie, haciéndola porosa o "esponjosa". Otro procedimiento para producir un electrodo poroso y/o esponjoso es depositar óxido metálico (por ejemplo, óxido de platino) y exponerlo a un agente reductor (por ejemplo, H₂ al 4 %). El agente reductor puede reducir el óxido metálico (por ejemplo, óxido de platino) de nuevo a metal (por ejemplo, platino) y, al hacerlo, proporciona un electrodo esponjoso y/o poroso. La esponja (por ejemplo, de paladio) puede absorber el electrolito y crear una gran superficie eficaz (por ejemplo, 33 pico-faradios por micrómetro cuadrado del área de arriba a abajo del electrodo). Incrementar el área superficial del electrodo haciéndolo poroso, como se describe en el presente documento, puede crear un electrodo que tiene una capacidad que no se agota por completo.

En algunos casos, la necesidad de mantener una diferencia de tensión de polaridad conservada a través del nanoporo durante la detección durante largos periodos de tiempo (por ejemplo, al secuenciar un ácido nucleico pasando el ácido nucleico a través del nanoporo) agota los electrodos y puede limitar la duración de la detección y/o tamaño de los electrodos. Los dispositivos y procedimientos descritos en el presente documento permiten tiempos de detección mayores (por ejemplo, infinitos) y/o electrodos que se pueden reducir a un tamaño arbitrariamente pequeño (por ejemplo, según se limita por consideraciones distintas al agotamiento del electrodo durante la detección). Como se describe en el presente documento, la marca se puede detectar solo durante una parte del tiempo que está asociada con la polimerasa. Invertir la polaridad y/o la magnitud de la tensión a través del nanoporo (por ejemplo, aplicando una forma de onda de CA) entre periodos de detección permite recargar los electrodos. En algunos casos, la marca se detecta una pluralidad de veces (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 1000, 10.000, 100.000, 1.000.000 o más veces en un periodo de 100 milisegundos).

En algunos casos, la polaridad de la tensión a través del nanoporo se invierte periódicamente. La polaridad de la tensión se puede invertir después de periodos de detección que duran cualquier cantidad de tiempo adecuada (por ejemplo, aproximadamente 1 ms, aproximadamente 5 ms, aproximadamente 10 ms, aproximadamente 15 ms, aproximadamente 20 ms, aproximadamente 25 ms, aproximadamente 30 ms, aproximadamente 40 ms, aproximadamente 50 ms, aproximadamente 60 ms, aproximadamente 80 ms, aproximadamente 100 ms, aproximadamente 125 ms, aproximadamente 150 ms, aproximadamente 200 ms y similares). El periodo de tiempo y la intensidad del campo eléctrico durante los periodos de recarga de los electrodos (es decir, cuando la polaridad de la tensión es opuesta a la de la tensión para la detección de marcas) es tal que el electrodo se restablezca a su estado anterior a la detección (por ejemplo, masa del electrodo). La tensión neta a través del nanoporo es cero en algunos casos (por ejemplo, los periodos de tensión positiva anulan los periodos de tensión negativa en una escala de tiempo adecuadamente larga, tal como 1 segundo, 1 minuto o 5 minutos). En algunos casos, la tensión aplicada a un nanoporo se equilibra de modo que un electrodo sensor contiguo o próximo al nanoporo detecte una corriente neta cero.

En algunos ejemplos se aplica una forma de onda de corriente alterna (CA) a un nanoporo en una membrana o un electrodo contiguo a la membrana para llevar una marca a través de o próxima al nanoporo y para liberar la marca. La forma de onda de CA puede tener una frecuencia del orden de al menos 10 microsegundos, 1 milisegundo (ms), 5 ms, 10 ms, 20 ms, 100 ms, 200 ms, 300 ms, 400 ms, 500 ms. La forma de onda puede ayudar a capturar la marca y liberar la marca de forma alternativa y secuencial o, de otro modo, a mover la marca en múltiples direcciones (por ejemplo, direcciones opuestas), lo que puede incrementar el periodo de tiempo global en el que la marca está asociada con el nanoporo. Este equilibrio de carga y descarga puede permitir la generación de una señal más larga desde un electrodo de nanoporos y/o una marca dada.

En algunos ejemplos se aplica una forma de onda de CA para dirigir repetidamente al menos una parte de una marca asociada con un nucleótido marcado (por ejemplo, nucleótido marcado incorporado) hacia un nanoporo y dirigir al menos una parte de la marca fuera del nanoporo. Una enzima (por ejemplo, polimerasa) puede retener la marca o el nucleótido acoplado a la marca. Esta carga y expulsión repetitivas de una única marca retenida por la enzima puede proporcionar de forma ventajosa más oportunidades para detectar la marca. Por ejemplo, si la enzima retiene la marca durante 40 milisegundos (ms) y la forma de onda de CA se aplica alta durante 5 ms (para dirigir la marca hacia el nanoporo) y se aplica baja durante 5 ms (para dirigir la marca fuera del nanoporo), el nanoporo se puede usar para leer la marca aproximadamente 4 veces. Las lecturas múltiples pueden permitir la corrección de errores, tales como errores asociados con el enhebrado en y/o desenhebrado de marcas de un nanoporo.

La forma de onda puede tener cualquier conformación adecuada, que incluye conformaciones regulares (por ejemplo, que se repiten durante un periodo de tiempo) y conformaciones irregulares (por ejemplo, que no se repiten durante ningún periodo de tiempo adecuadamente largo, tal como 1 hora, 1 día o 1 semana). La **figura 21** muestra algunas formas de onda adecuadas (regulares). Los ejemplos de formas de onda incluyen ondas triangulares, (**panel A**) ondas sinusoidales (**panel B**), ondas en diente de sierra, ondas cuadradas y similares.

La inversión de la polaridad (es decir, de positiva a negativa o de negativa a positiva) de la tensión a través del nanoporo, tal como, por ejemplo, tras la aplicación de una forma de onda de corriente alterna (CA), se puede realizar por cualquier motivo, que incluye, pero sin limitarse a, (a) recargar el electrodo (p. ej., cambiando la composición química del electrodo metálico), (b) reequilibrar las concentraciones iónicas en el lado cis y trans de la membrana, (c) restablecer una tensión aplicada distinta de cero a través del nanoporo y/o (d) alterar la capacidad de doble capa (por ejemplo, restableciendo la tensión o la carga que existe en el electrodo metálico y la interfaz del analito a un nivel deseado, por ejemplo, cero).

La **figura 21C** muestra una línea discontinua horizontal con una diferencia de potencial cero a través del nanoporo con una tensión positiva que se extiende hacia arriba en proporción a la magnitud y una tensión negativa que se extiende hacia abajo en proporción a la magnitud. Sin importar la conformación de la forma de onda, el "ciclo de trabajo" compara el área combinada bajo la curva de una curva de tensión frente a tiempo en la dirección positiva **2100** con el área combinada bajo la curva en la dirección negativa **2101**. En algunos casos, el área positiva **2100** es igual al área negativa **2101** (es decir, el ciclo de trabajo neto es cero); sin embargo, la forma de onda de CA puede tener cualquier ciclo de trabajo. En algunos casos, el uso judicial de una forma de onda de CA que tiene un ciclo de trabajo optimizado se puede usar para lograr uno cualquiera o más de (a) el electrodo se equilibre electroquímicamente (por ejemplo, ni cargado ni agotado), (b) la concentración iónica entre el lado cis y trans de la membrana se equilibre, (c) se conoce la tensión aplicada a través del nanoporo (por ejemplo, porque la capa doble capacitiva en el electrodo se reajusta periódicamente y el condensador se descarga en la misma medida con cada cambio de polaridad), (d) la molécula marcadora se identifica en el nanoporo una pluralidad de veces (por ejemplo, al expulsar y volver a capturar la marca con cada cambio de polaridad), (e) se obtiene información adicional de cada lectura de la molécula marcadora (por ejemplo, debido a que la corriente medida puede ser una función diferente de la tensión aplicada para cada molécula marcadora), (f) se logra una alta densidad de sensores de nanoporo (por ejemplo, debido a que la composición del electrodo metálico no está cambiando, no se está limitado por la cantidad de metal que comprende el electrodo), y/o (g) se logra un bajo consumo de energía del

chip. Estos beneficios pueden permitir un funcionamiento continuo y prolongado del dispositivo (por ejemplo, al menos 1 hora, al menos 1 día, al menos 1 semana).

En algunas situaciones, tras la aplicación de un potencial positivo a través de un nanoporo, se mide una primera corriente, y tras aplicar un potencial negativo (por ejemplo, de igual magnitud absoluta al potencial positivo) a través del nanoporo, se mide una segunda corriente. La primera corriente puede ser igual a la segunda corriente, aunque en algunos casos la primera corriente y la segunda corriente pueden ser diferentes. Por ejemplo, la primera corriente puede ser menor que la segunda corriente. En algunos casos, solo se mide una de una corriente positiva y una corriente negativa.

En algunos casos, el nanoporo detecta nucleótidos marcados durante periodos de tiempo relativamente largos a una tensión de magnitud relativamente baja (por ejemplo, **figura 21**, indicación **2100**) y recarga el electrodo durante periodos de tiempo relativamente cortos a una tensión de magnitud relativamente grande (por ejemplo, **figura 21**, indicación **2101**). En algunos casos, el periodo de tiempo para la detección es al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 8, al menos 10, al menos 15, al menos 20 o al menos 50 veces mayor que el periodo de tiempo para la recarga del electrodo.

En algunos casos, la forma de onda se altera en respuesta a una entrada. En algunos casos, la entrada es el nivel de agotamiento del electrodo. En algunos casos, la polaridad y/o magnitud de la tensión varía al menos en parte en base al agotamiento del electrodo o de los iones de conducción de corriente y la forma de onda es irregular.

La capacidad de detectar y recargar repetidamente los electrodos durante periodos cortos de tiempo (por ejemplo, durante periodos de menos de aproximadamente 5 segundos, menos de aproximadamente 1 segundo, menos de aproximadamente 500 ms, menos de aproximadamente 100 ms, menos de aproximadamente 50 ms, menos de aproximadamente 10 ms o menos de aproximadamente 1 ms) permite el uso de electrodos más pequeños con respecto a los electrodos que pueden mantener un potencial de corriente continua (CC) y corriente CC constantes y se usan para secuenciar polinucleótidos que se enhebran a través del nanoporo. Los electrodos más pequeños pueden permitir un gran número de sitios de detección (por ejemplo, que comprenden un electrodo, un circuito sensor, un nanoporo y una polimerasa) en una superficie.

La superficie comprende cualquier densidad adecuada de sitios discretos (por ejemplo, una densidad adecuada para secuenciar una muestra de ácido nucleico en una cantidad tiempo dado o por un coste dado). En un caso, la superficie tiene una densidad de sitios discretos mayor o igual a aproximadamente 500 sitios por 1 mm². En algunos casos, la superficie tiene una densidad de sitios discretos de aproximadamente 100, aproximadamente 200, aproximadamente 300, aproximadamente 400, aproximadamente 500, aproximadamente 600, aproximadamente 700, aproximadamente 800, aproximadamente 900, aproximadamente 1000, aproximadamente 2000, aproximadamente 3000, aproximadamente 4000, aproximadamente 5000, aproximadamente 6000, aproximadamente 7000, aproximadamente 8000, aproximadamente 9000, aproximadamente 10.000, aproximadamente 20.000, aproximadamente 40.000, aproximadamente 60.000, aproximadamente 80.000, aproximadamente 100.000 o aproximadamente 500.000 sitios por 1 mm². En algunos casos, la superficie tiene una densidad de sitios discretos de al menos aproximadamente 200, al menos aproximadamente 300, al menos aproximadamente 400, al menos aproximadamente 500, al menos aproximadamente 600, al menos aproximadamente 700, al menos aproximadamente 800, al menos aproximadamente 900, al menos aproximadamente 1000, al menos aproximadamente 2000, al menos aproximadamente 3000, al menos aproximadamente 4000, al menos aproximadamente 5000, al menos aproximadamente 6000, al menos aproximadamente 7000, al menos aproximadamente 8000, al menos aproximadamente 9000, al menos aproximadamente 10.000, al menos aproximadamente 20.000, al menos aproximadamente 40.000, al menos aproximadamente 60.000, al menos aproximadamente 80.000, al menos aproximadamente 100.000 o al menos aproximadamente 500.000 sitios por 1 mm².

El electrodo se puede recargar antes, entre o durante, o después de los acontecimientos de incorporación de nucleótidos. En algunos casos, el electrodo se recarga en aproximadamente 20 milisegundos (ms), aproximadamente 40 ms, aproximadamente 60 ms, aproximadamente 80 ms, aproximadamente 100 ms, aproximadamente 120 ms, aproximadamente 140 ms, aproximadamente 160 ms, aproximadamente 180 ms o aproximadamente 200 ms. En algunos casos, el electrodo se recarga en menos de aproximadamente 20 milisegundos (ms), menos de aproximadamente 40 ms, menos de aproximadamente 60 ms, menos de aproximadamente 80 ms, menos de aproximadamente 100 ms, menos de aproximadamente 120 ms, menos de aproximadamente 140 ms, menos de aproximadamente 160 ms, menos de aproximadamente 180 ms, aproximadamente 200 ms, menos de aproximadamente 500 ms o menos de aproximadamente 1 segundo.

Chips que pueden distinguir entre marcas escindidas y no escindidas

La descripción proporciona chips para secuenciar una muestra de ácido nucleico. En un ejemplo, un chip comprende una pluralidad de nanoporos direccionables individualmente. Un nanoporo direccionable individualmente de la pluralidad puede tener al menos un nanoporo formado en una membrana dispuesta contigua

a un circuito integrado. Cada nanoporo direccionable individualmente se puede adaptar para determinar si una molécula marcadora está unida o no a un nucleótido o para leer el cambio entre diferentes marcas.

En algunos casos, un chip puede comprender una pluralidad de nanoporos direccionables individualmente. Un nanoporo direccionable individualmente de la pluralidad puede tener al menos un nanoporo formado en una membrana dispuesta contigua a un circuito integrado. Cada nanoporo direccionable individualmente se puede adaptar para determinar si una molécula marcadora está unida a un nucleótido incorporado (por ejemplo, polimerizado) o a un nucleótido no incorporado.

Los chips descritos en el presente documento pueden ser capaces de distinguir entre una marca liberada y una marca no liberada (por ejemplo, **D** frente a **F** en la **figura 4**). En algunos casos, el chip puede distinguir entre una marca liberada y una marca no liberada con una exactitud de al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 97 %, al menos aproximadamente un 99 %, al menos aproximadamente un 99,5 %, al menos aproximadamente un 99,9 %, al menos aproximadamente un 99,95 % o al menos aproximadamente un 99,99 %. El nivel de exactitud se puede lograr cuando se detectan grupos de aproximadamente 5, 4, 3 o 2 nucleótidos consecutivos. En algunos casos, la exactitud se logra para una resolución de base única (es decir, 1 nucleótido consecutivo).

Los chips descritos en el presente documento pueden ser capaces de distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado (por ejemplo, **506** y **505** en la **figura 5**). En algunos casos, el chip puede distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos aproximadamente un 80 %, al menos aproximadamente un 85 %, al menos aproximadamente un 90 %, al menos aproximadamente un 95 %, al menos aproximadamente un 97 %, al menos aproximadamente un 99 %, al menos aproximadamente un 99,5 %, al menos aproximadamente un 99,9 %, al menos aproximadamente un 99,95 % o al menos aproximadamente un 99,99 %. El nivel de exactitud se puede lograr cuando se detectan grupos de aproximadamente 5, 4, 3 o 2 nucleótidos consecutivos. En algunos casos, la exactitud se logra para una resolución de base única (es decir, 1 nucleótido consecutivo).

El nanoporo puede ayudar a determinar si una molécula marcadora está unida o no a un nucleótido en base, al menos en parte, a las diferencias en una señal eléctrica. En algunos casos, el nanoporo puede ayudar a determinar si una molécula marcadora está unida o no a un nucleótido en base, al menos en parte, al tiempo de permanencia en el nanoporo. El nanoporo puede ayudar a determinar si una molécula marcadora está unida o no a un nucleótido en base, al menos en parte, a la tensión de caída, que es la tensión a la que la marca o nucleótido marcado sale del nanoporo.

Chips que pueden capturar una alta proporción de marcas escindidas

La descripción proporciona además chips para secuenciar una muestra de ácido nucleico. En un ejemplo, un chip comprende una pluralidad de nanoporos direccionables individualmente. Un nanoporo direccionable individualmente de la pluralidad puede incluir al menos un nanoporo formado en una membrana dispuesta contigua a un circuito integrado. Cada nanoporo direccionable individualmente se puede adaptar para capturar la mayoría de las moléculas marcadoras liberadas tras la incorporación (por ejemplo, polimerización) de nucleótidos marcados.

El chip se puede configurar para capturar cualquier porcentaje de marcas adecuadamente alto (por ejemplo, para determinar la secuencia de ácido nucleico con una exactitud adecuadamente alta). En algunos casos, el chip captura al menos un 90 %, al menos un 99 %, al menos un 99,9 % o al menos un 99,99 % de las moléculas marcadoras.

El nanoporo puede capturar una pluralidad de moléculas marcadoras diferentes (por ejemplo, cuatro moléculas marcadoras distintas liberadas tras la incorporación de los cuatro nucleótidos) a un único nivel de corriente. El chip se puede adaptar para capturar moléculas marcadoras en la misma secuencia en la que se liberan las moléculas marcadoras.

Configuración del dispositivo

La **figura 8** ilustra esquemáticamente un dispositivo (o sensor) de nanoporos **100** que se puede usar para secuenciar un ácido nucleico y/o detectar una molécula marcadora como se describe en el presente documento. La bicapa lipídica que contiene el nanoporo se puede caracterizar por una resistencia y capacidad. El dispositivo de nanoporos **100** incluye una bicapa lipídica **102** formada sobre una superficie compatible con la bicapa lipídica **104** de un sustrato sólido conductor **106**, donde la superficie compatible con la bicapa lipídica **104** se puede aislar mediante superficies incompatibles con la bicapa lipídica **105** y el sustrato sólido conductor **106** se puede aislar eléctricamente mediante materiales aislantes **107**, y donde la bicapa lipídica **102** puede estar rodeada por un lípido amorfo **103** formado sobre la superficie incompatible con la bicapa lipídica **105**. La bicapa lipídica **102** puede tener incluida con una única estructura de nanoporo **108** que tiene un nanoporo **110** lo suficientemente grande para que

pasen las moléculas marcadoras caracterizadas y/o los iones pequeños (por ejemplo, Na^+ , K^+ , Ca^{2+} , Cl^-) entre los dos lados de la bicapa lipídica **102**. Una capa de moléculas de agua **114** se puede absorber en la superficie compatible con la bicapa lipídica **104** e intercalar entre la bicapa lipídica **102** y la superficie compatible con la bicapa lipídica **104**. La película acuosa **114** adsorbida en la superficie hidrófila compatible con la bicapa lipídica **104** puede promover el ordenamiento de las moléculas lipídicas y facilitar la formación de la bicapa lipídica sobre la superficie compatible con la bicapa lipídica **104**. Se pueden proporcionar una cámara de muestra **116** que contiene una solución de la molécula de ácido nucleico **112** y nucleótidos marcados sobre la bicapa lipídica **102**. La solución puede ser una solución acuosa que contiene electrolitos y se encuentra amortiguada a una concentración iónica óptima y se mantiene a un pH óptimo para mantener abierto el nanoporo **110**. El dispositivo incluye un par de electrodos **118** (que incluye un nodo negativo **118a** y un nodo positivo **118b**) acoplado a una fuente de tensión variable **120** para proporcionar estímulo eléctrico (por ejemplo, polarización de tensión) a través de la bicapa lipídica y para detectar características eléctricas de la bicapa lipídica (por ejemplo, resistencia, capacidad y flujo de corriente iónica). La superficie del electrodo positivo **118b** es la superficie compatible con la bicapa lipídica **104** o forma parte de la misma. El sustrato sólido conductor **106** se puede acoplar a uno de los electrodos **118** o formar parte del mismo. El dispositivo **100** también puede incluir un circuito eléctrico **122** para controlar la estimulación eléctrica y para procesar la señal detectada. En algunos casos, la fuente de tensión variable **120** se incluye como parte del circuito eléctrico **122**. Los circuitos eléctricos **122** pueden incluir amplificador, integrador, filtro de ruido, lógica de control de realimentación y/o otros diversos componentes. Los circuitos eléctricos **122** pueden ser circuitos eléctricos integrados que estén integrados dentro de un sustrato de silicio **128** y se pueden acoplar además a un procesador informático **124** acoplado a una memoria **126**.

La superficie compatible con la bicapa lipídica **104** se puede formar a partir de diversos materiales que son adecuados para la transducción iónica y la formación de gas para facilitar la formación de la bicapa lipídica. En algunos casos, se pueden usar materiales hidrófilos conductores o semiconductores, ya que pueden permitir una mejor detección de un cambio en las características eléctricas de la bicapa lipídica. Los materiales de ejemplo incluyen Ag-AgCl, Au, Pt o silicio dopado u otros materiales semiconductores. En algunos casos, el electrodo no es un electrodo de protección.

La superficie incompatible con la bicapa lipídica **105** se puede formar a partir de diversos materiales que no son adecuados para la formación de la bicapa lipídica y son típicamente hidrófobos. En algunos casos, se prefieren materiales hidrófobos no conductores, ya que aíslan eléctricamente las regiones de la bicapa lipídica además de separar las regiones de la bicapa lipídica entre sí. Los ejemplos de materiales incompatibles con la bicapa lipídica incluyen, por ejemplo, nitruro de silicio (por ejemplo, Si_3N_4) y teflón, óxido de silicio (por ejemplo, SiO_2) silanizado con moléculas hidrófobas.

En un ejemplo, el dispositivo de nanoporos **100** de la **figura 8** es un dispositivo de nanoporos de hemolisina alfa (aHL) que tiene una única proteína de hemolisina alfa (aHL) **108** incluida en una bicapa lipídica de difitanoilfosfatidilcolina (DPhPC) **102** formada sobre una superficie de plata (Ag) compatible con la bicapa lipídica **104** revestida sobre un material de aluminio **106**. La superficie de Ag compatible con la bicapa lipídica **104** se aísla mediante superficies de nitruro de silicio incompatibles con la bicapa lipídica **105**, y el material de aluminio **106** se aísla eléctricamente con materiales de nitruro de silicio **107**. El aluminio **106** se acopla a circuitos eléctricos **122** que están integrados en un sustrato de silicio **128**. Un electrodo de cloruro de plata y plata colocado en un chip o extendido hacia abajo desde una placa de cobertura **128** entra en contacto con una solución acuosa que contiene moléculas de ácido nucleico.

El nanoporo de aHL es un conjunto de siete péptidos individuales. La entrada o el vestíbulo del nanoporo de aHL tienen aproximadamente 26 angstroms de diámetro, que es lo suficientemente ancho como para dar cabida a una parte de una molécula de ADNbc. Desde el vestíbulo, el nanoporo de aHL primero se ensancha y a continuación se estrecha para formar un barril que tiene un diámetro de aproximadamente 15 angstroms, que es lo suficientemente ancho para permitir el paso de una sola molécula de ADNmc (o moléculas marcadoras más pequeñas), pero no lo suficientemente ancho como para permitir el paso de una molécula de ADNbc (o moléculas marcadoras más grandes).

Además de la DPhPC, la bicapa lipídica del dispositivo de nanoporos se puede ensamblar a partir de otros diversos materiales anfipáticos adecuados, seleccionados en base a diversas consideraciones, tales como el tipo de nanoporo utilizado, el tipo de molécula que se va a caracterizar y diversas características físicas, químicas y/o eléctricas de la bicapa lipídica formada, tales como estabilidad y permeabilidad, resistencia y capacidad de la bicapa lipídica formada. Entre los ejemplos de materiales anfipáticos se incluyen diversos fosfolípidos, tales como palmitoiloleoilfosfatidilcolina (POPC) y dioleoilfosfatidilmetiléster (DOPME), difitanoilfosfatidilcolina (DPhPC), 1,2-di-O-fitanil-sn-glicero-3-fosfocolina (DoPhPC), dipalmitoilfosfatidilcolina (DPPC), fosfatidilcolina, fosfatidiletanolamina, fosfatidilserina, ácido fosfatídico, fosfatidilinositol, fosfatidilglicerol y esfingomielina.

Además del nanoporo de aHL que se muestra anteriormente, el nanoporo puede ser de otros diversos tipos de nanoporos. Los ejemplos incluyen hemolisina γ , leucocidina, melitina, porina A de *Mycobacterium smegmatis* (MspA) y diversos otros nanoporos naturales, naturales modificados y sintéticos. Se puede seleccionar un nanoporo adecuado en base a diversas características de la molécula del analito, tales como el tamaño de la

molécula del analito con respecto al tamaño de poro del nanoporo. Por ejemplo, el nanoporo de aHL, que tiene un tamaño de poro restrictivo de aproximadamente 15 angstroms.

Medición de la corriente

En algunos casos, la corriente se puede medir a diferentes tensiones aplicadas. Para lograr esto, se puede aplicar un potencial deseado al electrodo, y el potencial aplicado se puede mantener posteriormente a lo largo de la medición. En una implementación, se puede usar una topología de integrador amplificador operacional (op-amp) para este propósito como se describe a continuación. El integrador mantiene el potencial de tensión en el electrodo por medio de retroalimentación capacitiva. El circuito integrador puede proporcionar una linealidad sobresaliente, coincidencia entre celdas y características de compensación. El integrador op-amp típicamente requiere un tamaño grande para lograr el rendimiento requerido. A continuación se describe una topología de integrador más compacta.

En algunos casos, se puede aplicar a la cámara un potencial de tensión "Vliquid" que proporciona un potencial eléctrico común (por ejemplo, 350 mV) para todas las celdas del chip. El circuito integrador puede inicializar el electrodo (que es, eléctricamente, la placa superior del condensador integrador) a un potencial mayor que el potencial del líquido común. Por ejemplo, la polarización a 450 mV puede dar un potencial positivo de 100 mV entre el electrodo y el líquido. Este potencial de tensión positivo puede provocar que una corriente fluya desde el electrodo al contacto de la cámara de líquido. En este caso, los portadores son: (a) iones K^+ que fluyen a través del poro desde el lado del electrodo (trans) de la bicapa al lado del depósito de líquido (cis) de la bicapa y (b) iones de cloro (Cl^-) en el lado trans que reaccionan con el electrodo de plata de acuerdo con la siguiente reacción electroquímica: $Ag + Cl^- \rightarrow AgCl + e^-$.

En algunos casos, K^+ sale de la celda contenida (del lado trans al lado cis de la bicapa) mientras que Cl^- se convierte en cloruro de plata. El lado del electrodo de la bicapa se puede desalinizar como resultado del flujo de corriente. En algunos casos, un material esponjoso líquido o una matriz de plata/cloruro de plata pueden servir como depósito para suministrar iones Cl^- en la reacción inversa que se produce en el contacto eléctrico de la cámara para completar el circuito.

En algunos casos, los electrones fluyen finalmente hacia el lado superior del condensador integrador, lo que crea la corriente eléctrica que se mide. La reacción electroquímica convierte la plata en cloruro de plata y la corriente continuará fluyendo solo mientras haya plata disponible que convertir. En algunos casos, el suministro limitado de plata da lugar a una vida útil del electrodo dependiente de la corriente. En algunos casos se usan materiales de electrodo que no se agotan (por ejemplo, platino).

La marca puede modular una corriente iónica que fluye a través de un nanoporo cuando se aplica un potencial eléctrico constante al detector de nanoporos, permitiendo un registro de la corriente para determinar la identidad de la marca. Sin embargo, un potencial eléctrico constante puede no distinguir adecuadamente entre diferentes marcas (por ejemplo, marcas asociadas con A, C, T o G). La tensión aplicada se puede variar (por ejemplo, realizar un barrido por un intervalo de tensiones) para identificar la marca (por ejemplo, con una confianza de al menos un 90 %, al menos un 95 %, al menos un 99 %, al menos un 99,9 % o al menos un 99,99 %).

La tensión aplicada se puede variar de cualquier manera adecuada, que incluye de acuerdo con cualquiera de las formas de onda mostradas en la **figura 21**. La tensión se puede variar en cualquier intervalo adecuado, que incluye de aproximadamente 120 mV a aproximadamente 150 mV, de aproximadamente 40 mV a aproximadamente 150 mV.

La **figura 22** muestra la señal extraída (por ejemplo, conductancia diferencial logarítmica (DLC)) frente a la tensión aplicada para los nucleótidos adenina (A, verde), citosina (C, azul), guanina (G, negro) y timina (T, rojo). La **figura 23** muestra la misma información para una pluralidad de nucleótidos (muchos ensayos experimentales). Como se ve aquí, la citosina es relativamente fácil de distinguir de la timina a 120 mV, pero son difíciles de distinguir entre sí a 150 mV (por ejemplo, porque la señal extraída es aproximadamente igual para C y T a 150 mV). Además, la timina es difícil de distinguir de la adenina a 120 mV, pero es relativamente más fácil de distinguir a 150 mV. Por lo tanto, en un caso, la tensión aplicada puede variar de 120 mV a 150 mV para distinguir cada uno de los nucleótidos A, C, G y T.

La **figura 24** muestra el porcentaje de diferencia conductiva de referencia (% de RCD) como función de la tensión aplicada para los nucleótidos adenina (A, verde), citosina (C, azul), guanina (G, negro) y timina (T, rojo). La representación gráfica del % de RCD (que es esencialmente la diferencia en conductancia de cada molécula referenciada a una molécula de referencia 30T) puede eliminar la compensación y la variación de ganancia entre los experimentos. La **figura 24** incluye formas de onda de ADN individual del primer bloque de 17/20 ensayos. El % de RCD de todo el ADN de un único nucleótido captura del número 50 al 200 para los 17 ensayos buenos. Se indican las tensiones donde cada uno de los nucleótidos son distinguibles.

Aunque las **figuras 22-24** muestran la respuesta a la tensión aplicada variada para nucleótidos, el concepto de tensión aplicada variada se puede usar para distinguir moléculas marcadoras (por ejemplo, unidas a nucleótidos marcados).

5 **Circuitos de la celda**

En la **figura 12** se muestra un ejemplo de circuitos de la celda. Se aplica una tensión aplicada V_a a un op-amp **1200** por delante de una puerta transportadora de corriente MOSFET **1201**. También se muestran aquí un electrodo **1202** y la resistencia del ácido nucleico y/o marca detectada por el dispositivo **1203**.

Una tensión aplicada V_a puede impulsar la puerta transportadora de corriente **1201**. La tensión resultante en el electrodo es entonces $V_a - V_t$, donde V_t es la tensión de umbral del MOSFET. En algunos casos, esto da como resultado un control limitado de la tensión real aplicada al electrodo, ya que la tensión de umbral de un MOSFET puede variar considerablemente con el procedimiento, la tensión, la temperatura e, incluso, entre dispositivos dentro de un chip. Esta variación de V_t puede ser mayor en niveles de corriente bajos, donde pueden entrar en juego efectos de fuga por debajo del umbral. Por lo tanto, para proporcionar un mejor control de la tensión aplicada, se puede usar un op-amp en una configuración de retroalimentación de seguimiento con el dispositivo transportador de corriente. Esto asegura que la tensión aplicada al electrodo sea V_a , independientemente de la variación de la tensión de umbral del MOSFET.

En la **figura 10** se muestra otro ejemplo de circuitos de la celda e incluye un integrador, un comparador y un circuito lógico digital para desviar los bits de control y, simultáneamente, desviar el estado de la salida del comparador. Los circuitos de la celda se pueden adaptar para su uso con los sistemas y procedimientos proporcionados en el presente documento. Las líneas B0 a B1 pueden salir del registro de desviación. Todas las celdas dentro de un banco comparten las señales analógicas, mientras que las líneas digitales se pueden conectar en batería de una celda a otra.

El circuito lógico digital de la celda comprende el registro de desviación de datos (DSR) de 5 bits, los registros de carga paralela (PLR) de 5 bits, la lógica de control y el circuito integrador analógico. Mediante el uso de la señal LIN, los datos de control desviados hacia el DSR se cargan en paralelo en el PLR. Estos 5 bits controlan la lógica de temporización digital "interrupción previa a la conexión" que controla los interruptores en la celda. Además, el circuito lógico digital tiene un circuito de cierre biestable SR (set-reset) para registrar la conmutación de la salida del comparador.

La arquitectura ofrece una velocidad de muestreo variable que es proporcional a la corriente de la celda individual. Una corriente mayor puede dar como resultado más muestras por segundo que una corriente menor. La resolución de la medición de corriente se relaciona con la corriente que se mide. Se puede medir una corriente pequeña con una resolución mayor que una corriente grande, lo que puede ser un beneficio sobre los sistemas de medición de resolución fija. Hay una entrada analógica que permite al usuario ajustar las frecuencias de muestreo al cambiar la oscilación de tensión del integrador. Puede ser posible incrementar la velocidad de muestreo para analizar procesos biológicamente rápidos o ralentizar la velocidad de muestreo (y, por lo tanto, ganar precisión) para analizar procesos biológicamente lentos.

La salida del integrador se inicializa a la tensión LVB (baja polarización de tensión) y se integra hasta la tensión CMP. Se genera una muestra cada vez que la salida del integrador oscila entre estos dos niveles. Por tanto, cuanto mayor es la corriente, más rápido oscila la salida del integrador y, por lo tanto, más rápida es la velocidad de muestreo. De forma similar, si la tensión CMP se reduce, la oscilación de salida del integrador necesaria para generar una nueva muestra se reduce y, por lo tanto, se incrementa la velocidad de muestreo. Por tanto, simplemente reducir la diferencia de tensión entre LVB y CMP proporciona un mecanismo para incrementar la velocidad de muestreo.

Un chip de secuenciación basado en nanoporos puede incorporar un gran número de celdas de funcionamiento autónomo o direccionables individualmente configuradas como una matriz. Por ejemplo, una matriz de un millón de celdas se puede construir con 1000 filas por 1000 columnas de celdas. Esta matriz permite la secuenciación en paralelo de moléculas de ácido nucleico al medir la diferencia de conductancia cuando, por ejemplo, el nanoporo detecta las marcas liberadas tras acontecimientos de incorporación de nucleótidos. Además, esta implementación de circuitos permite determinar las características de conductancia del complejo poro-molecular que pueden ser valiosas para distinguir entre marcas.

Las estructuras de celdas electrónicas integradas de nanoporos/bicapas pueden aplicar tensiones apropiadas para realizar mediciones de corriente. Por ejemplo, puede ser necesario tanto (a) controlar el potencial de tensión del electrodo como (b) hacer un seguimiento de la corriente del electrodo simultáneamente para que funcione correctamente.

Además, puede ser necesario controlar las celdas independientemente entre sí. Es posible que se requiera el control independiente de una celda para gestionar una gran cantidad de celdas que pueden estar en diferentes

estados físicos. El control preciso del estímulo de forma de onda de tensión lineal por segmentos aplicado al electrodo se puede usar para la transición entre los estados físicos de la celda.

Para reducir el tamaño y la complejidad del circuito, puede ser suficiente proporcionar lógica para aplicar dos tensiones separadas. Esto permite que se apliquen dos agrupaciones independientes de celdas y el estímulo de transición de estado correspondiente. Las transiciones de estado son de naturaleza estocástica, con una probabilidad relativamente baja de aparición. Por tanto, puede ser altamente útil poder mantener la tensión de control apropiada y, posteriormente, realizar una medición para determinar si se ha producido la transición de estado deseada. Por ejemplo, se puede aplicar la tensión apropiada a una celda y, a continuación, medir la corriente para determinar si se ha formado una bicapa. Las celdas se dividen en dos grupos: (a) las que han tenido una forma bicapa y ya no necesitan que se les aplique la tensión. Estas celdas pueden tener una polarización de 0 V aplicada para efectuar la operación nula (NOP), es decir, permanecer en el mismo estado y (b) las que no tienen una bicapa formada. A estas celdas se les aplicará de nuevo la tensión eléctrica de formación de bicapa.

Se puede lograr una simplificación y una reducción del tamaño del circuito sustanciales restringiendo las tensiones aplicadas admisibles a dos y haciendo una transición iterativa de las celdas en lotes entre los estados físicos. Por ejemplo, se puede lograr una reducción de al menos un factor de 1, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50 o 100 restringiendo las tensiones aplicadas admisibles.

En la **figura 11** se muestra un circuito de medición compacto. En algunos casos, el circuito de medición compacto se puede usar para lograr las altas densidades de matriz descritas en el presente documento. Este circuito también se diseña para aplicar una tensión al electrodo mientras mide simultáneamente corrientes de bajo nivel.

La celda funciona como un integrador ultracompacto (UCI) y el funcionamiento básico se describe aquí. La celda se conecta eléctricamente a un electrodo electroquímicamente activo (por ejemplo, AgCl) a través de una conexión de detección de electrodo (ELSNS). El transistor NMOS M11 realiza dos funciones independientes: (1) funciona como un seguidor de fuente para aplicar una tensión al nodo ELSNS dado por ($V_{g1}V_{t1}$) y (2) funciona como un conductor de corriente para mover electrones desde el condensador C1 al nodo ELSNS (y viceversa).

En algunos casos se puede aplicar un potencial de tensión controlado al electrodo ELSNS y esto se puede variar simplemente cambiando la tensión en la puerta del seguidor de fuente de electrodo M11. Además, cualquier corriente del pin de la fuente M11 se propaga directamente y con exactitud al pin de drenaje M11, donde se puede acumular en el condensador C0. Por tanto, M11 y C0 actúan juntos como un integrador ultracompacto. El integrador se puede usar para determinar la corriente suministrada/extraída al/del electrodo, midiendo el cambio de tensión integrada en el condensador de acuerdo con lo siguiente: $I \times t = C \times V$, donde I es corriente, t es tiempo, C es capacidad y V es cambio de tensión.

En algunos casos, el cambio de tensión se mide a un intervalo fijo t (por ejemplo, cada 1 ms).

El transistor M2 se puede configurar como un seguidor de fuente para amortiguar la tensión del condensador y proporcionar una representación de baja impedancia de la tensión integrada. Esto evita que la carga compartida cambie la tensión en el condensador.

El transistor M3 se puede usar como dispositivo de acceso de fila con la salida de tensión analógica AOUT conectada como una columna compartida con muchas otras celdas. Solo se habilita una única fila de la señal AOUT conectada a la columna de modo que se mida una única tensión de la celda.

En una implementación alternativa, el transistor M3 se puede omitir conectando el drenaje del transistor M2 a un "carril conmutado" seleccionable de fila.

Se puede usar un transistor M4 para reajustar la celda a una tensión inicial predeterminada a partir de la cual se integra la tensión. Por ejemplo, la aplicación de una tensión alta (por ejemplo, a $V_{DD} = 1,8$ V) a tanto RST como RV elevará el condensador hasta un valor precargado de ($V_{DD} - V_{t5}$). El valor inicial exacto puede variar de una celda a otra (debido a la variación de V_t de M4 y M2), así como de una medición a otra, debido al ruido térmico del conmutador de reajuste (ruido $\sqrt{kTC}$). Como resultado, se usa una técnica de muestreo doble correlacionado (CDS) para medir la tensión inicial del integrador y la tensión final para determinar el cambio de tensión real durante el periodo de integración.

Téngase en cuenta también que el drenaje del transistor M4 se conecta a una tensión controlada RV (tensión de reajuste). Durante el funcionamiento normal, esto puede llevar a V_{DD} ; sin embargo, también puede llevar a una tensión baja. En efecto, si el "drenaje" de M4 lleva a tierra, entonces el flujo de corriente se puede invertir (es decir, la corriente puede fluir desde el electrodo al circuito a través de M1 y M4 y la noción de fuente y drenaje se puede intercambiar). En algunos casos, cuando se hace funcionar el circuito en este modo, esta tensión RV (asumiendo que V_{g1} y V_{g5} son al menos un umbral mayor que RV) controla la tensión negativa aplicada al electrodo (con respecto a la referencia del líquido). Por tanto, se puede usar una tensión de tierra en RV para aplicar una tensión negativa al electrodo (por ejemplo, para conseguir la electroporación o la formación de la bicapa).

Un convertidor analógico-digital (ADC, no mostrado) mide la tensión AOUT inmediatamente después del reajuste y, de nuevo, después del periodo de integración (realiza la medición del CDS) para determinar la corriente integrada durante un periodo de tiempo fijo. Y el ADC se puede implementar por columna o un transistor separado usado para cada columna como un multiplexor analógico para compartir un único ADC entre múltiples columnas. Este factor de multiplexor de columna se puede variar dependiendo de los requisitos de ruido, exactitud y rendimiento.

En cualquier momento dado, cada celda puede estar en uno de cuatro estados físicos diferentes: (1) cortocircuito a líquido (2) bicapa formada (3) bicapa + poro (4) bicapa + poro + ácido nucleico y/o moléculas marcadoras.

En algunos casos se aplica una tensión para mover las celdas entre estados. La operación NOP se usa para dejar una celda en un estado deseado particular, mientras que otras celdas se estimulan con un potencial aplicado para moverlas de un estado a otro.

Esto se puede lograr teniendo dos (o más) tensiones diferentes que se pueden aplicar a la tensión de puerta del seguidor de fuente M1 que se usa indirectamente para controlar la tensión aplicada al electrodo con respecto al potencial líquido. Por tanto, el transistor M5 se usa para aplicar la tensión A, mientras que el transistor M6 se usa para aplicar la tensión B. Por tanto, M5 y M6 funcionan conjuntamente como un multiplexor analógico con SELA o SELB que se activan alto para seleccionar la tensión.

Ya que cada celda puede estar en un posible estado diferente y que SELA y SELB son complementarios, se puede usar un elemento de memoria en cada celda para seleccionar entre la tensión A o B. Este elemento de memoria puede ser un elemento dinámico (condensador) que se actualizó en cada ciclo o un simple elemento de memoria adaptador-circuito de cierre (inversor de acoplamiento cruzado).

Estructura del chip de prueba de op-amp

En algunos ejemplos, un chip de prueba incluye una matriz de 264 sensores dispuestos en cuatro grupos separados (también llamados bancos) de 66 celdas de sensores cada uno. Cada grupo se divide a su vez en tres "columnas" con 22 "celdas" de sensores en cada columna. La denominación "celda" es apropiada dado que idealmente se forma una celda virtual que consiste en una bicapa lipídica y un nanoporo insertado encima de cada uno de los 264 sensores de la matriz (aunque el dispositivo puede funcionar con éxito con solo una fracción de las celdas del sensor así equipadas).

Existe una única almohadilla analógica de E/S que aplica un potencial de tensión al líquido contenido dentro de un cilindro conductor montado en la superficie de la matriz. Este potencial "líquido" se aplica al lado superior del poro y es común a todas las celdas de una matriz de detectores. El lado inferior del poro tiene un electrodo expuesto y cada celda del sensor puede aplicar a su electrodo un potencial inferior distinto. A continuación, se mide la corriente entre la conexión del líquido superior y la conexión del electrodo de cada celda en el lado inferior del poro. La celda del sensor mide la corriente que viaja a través del poro según la modulación de la molécula marcadora que pasa por el interior del poro.

En algunos casos, cinco bits controlan el modo de cada celda del sensor. Con referencia continua a la **figura 9**, cada una de las 264 celdas en la matriz se puede controlar individualmente. Los valores se aplican por separado a un grupo de 66 celdas. El modo de cada una de las 66 celdas de un grupo se controla desviando en serie los 330 valores digitales (66 x 5 bits/celda) a un registro de desviación de datos (DSR). Estos valores se desvían hacia la matriz usando los pines KIN (reloj) y DIN (entrada dat) con un par de pines separados para cada grupo de 66 celdas.

Por tanto, se usan 330 relojes para desviar 330 bits hacia el registro de desviación DSR. Un segundo registro de carga paralela (PLR) de 330 bits se carga en paralelo desde este registro de desviación cuando el correspondiente LIN<i> (entrada de carga) se mantiene alta. Al mismo tiempo que el PLR se carga en paralelo, el valor de estado de la celda se carga en el DSR.

Una operación completa puede consistir en 330 relojes para desviar 330 bits de datos hacia el DSR, un único ciclo de reloj con la señal LIN mantenida alta, seguido de 330 ciclos de reloj para leer los datos de estado capturados desviados fuera del DSR. La operación se perfecciona de modo que se puedan desviar 330 bits nuevos al DSR simultáneamente mientras se leen los 330 bits de la matriz. Por tanto, a una frecuencia de reloj de 50 MHz, el tiempo de ciclo para una lectura es $331/50 \text{ MHz} = 6,62 \text{ us}$.

Matrices de nanoporos para secuenciación

La divulgación proporciona una matriz de detectores (o sensores) de nanoporo para secuenciar ácidos nucleicos. Con referencia a la **figura 7**, se puede secuenciar una pluralidad de moléculas de ácido nucleico en una matriz de detectores de nanoporo. Aquí, cada localización de nanoporo (por ejemplo, **701**) comprende un nanoporo, en

algunos casos unido a una enzima polimerasa y/o enzimas fosfatasas. En general, también hay un sensor en cada localización de la matriz como se describe en alguna parte del presente documento.

En algunos ejemplos se proporciona una matriz de nanoporos unidos a una polimerasa de ácido nucleico, y se incorporan nucleótidos marcados con la polimerasa. Durante la polimerización, el nanoporo detecta una marca (por ejemplo, liberándola y pasándola al interior o a través del nanoporo, o presentándola al nanoporo). La matriz de nanoporos puede tener cualquier número adecuado de nanoporos. En algunos casos, la matriz comprende aproximadamente 200, aproximadamente 400, aproximadamente 600, aproximadamente 800, aproximadamente 1000, aproximadamente 1500, aproximadamente 2000, aproximadamente 3000, aproximadamente 4000, aproximadamente 5000, aproximadamente 10.000, aproximadamente 15.000, aproximadamente 20.000, aproximadamente 40.000, aproximadamente 60.000, aproximadamente 80.000, aproximadamente 100.000, aproximadamente 200.000, aproximadamente 400.000, aproximadamente 600.000, aproximadamente 800.000, aproximadamente 1.000.000 nanoporos y cantidades similares. En algunos casos, la matriz comprende al menos 200, al menos 400, al menos 600, al menos 800, al menos 1000, al menos 1500, al menos 2000, al menos 3000, al menos 4000, al menos 5000, al menos 10.000, al menos 15.000, al menos 20.000, al menos 40.000, al menos 60.000, al menos 80.000, al menos 100.000, al menos 200.000, al menos 400.000, al menos 600.000, al menos 800.000 o al menos 1.000.000 nanoporos.

En algunos casos, una única marca se libera y/o presenta tras la incorporación de un único nucleótido y se detecta por un nanoporo. En otros casos, se libera y/o presenta una pluralidad de marcas tras la incorporación de una pluralidad de nucleótidos. Un sensor de nanoporos contiguo a un nanoporo puede detectar una marca individual o una pluralidad de marcas. Se pueden detectar y procesar una o más señales asociadas con una pluralidad de marcas para proporcionar una señal promediada.

El sensor puede detectar las marcas como una función del tiempo. Las marcas detectadas con tiempo se pueden usar para determinar la secuencia de ácido nucleico de la muestra de ácido nucleico, tal como con la ayuda de un sistema informático (véase, por ejemplo, la **figura 16**) que se programa para registrar datos del sensor y generar información de secuencia a partir de los datos.

La matriz de detectores de nanoporo puede tener una densidad alta de sitios discretos. Por ejemplo, un número relativamente grande de sitios por unidad de superficie (es decir, densidad) permite la construcción de dispositivos más pequeños, que son portátiles, de bajo coste o tienen otros rasgos característicos ventajosos. Un sitio individual en la matriz puede ser un sitio direccionable individualmente. Un gran número de sitios que comprenden un nanoporo y un circuito sensor pueden permitir la secuenciación simultánea de un número relativamente grande de moléculas de ácido nucleico, tal como, por ejemplo, mediante secuenciación en paralelo. Dicho sistema puede incrementar el rendimiento y/o disminuir el coste de secuenciar una muestra de ácido nucleico.

Una muestra de ácido nucleico se puede secuenciar usando un sensor (o detector) que tiene un sustrato con una superficie que comprende sitios discretos, cada sitio individual teniendo un nanoporo, una polimerasa y, en algunos casos, al menos una enzima fosfatasa unida al nanoporo y un circuito sensor contiguo al nanoporo. El sistema puede comprender además una cubeta de lectura en comunicación fluida con el sustrato, estando la cubeta de lectura adaptada para suministrar uno o más reactivos al sustrato.

La superficie comprende cualquier densidad adecuada de sitios discretos (por ejemplo, una densidad adecuada para secuenciar una muestra de ácido nucleico en una cantidad tiempo dado o por un coste dado). Cada sitio discreto puede incluir un sensor. La superficie puede tener una densidad de sitios discretos mayor o igual a aproximadamente 500 sitios por 1 mm². En algunos casos, la superficie tiene una densidad de sitios discretos de aproximadamente 200, aproximadamente 300, aproximadamente 400, aproximadamente 500, aproximadamente 600, aproximadamente 700, aproximadamente 800, aproximadamente 900, aproximadamente 1000, aproximadamente 2000, aproximadamente 3000, aproximadamente 4000, aproximadamente 5000, aproximadamente 6000, aproximadamente 7000, aproximadamente 8000, aproximadamente 9000, aproximadamente 10.000, aproximadamente 20.000, aproximadamente 40.000, aproximadamente 60.000, aproximadamente 80.000, aproximadamente 100.000 o aproximadamente 500.000 sitios por 1 mm². En algunos casos, la superficie tiene una densidad de sitios discretos de al menos 200, al menos 300, al menos 400, al menos 500, al menos 600, al menos 700, al menos 800, al menos 900, al menos 1000, al menos 2000, al menos 3000, al menos 4000, al menos 5000, al menos 6000, al menos 7000, al menos 8000, al menos 9000, al menos 10.000, al menos 20.000, al menos 40.000, al menos 60.000, al menos 80.000, al menos 100.000 o al menos 500.000 sitios por 1 mm².

Nucleótidos marcados

En algunos casos, un nucleótido marcado comprende una marca que se puede escindir en un acontecimiento de incorporación de nucleótidos y detectar con la ayuda de un nanoporo. La marca se puede unir al fosfato en 5' del nucleótido. En algunos casos, la marca no es un fluoróforo. La marca puede ser detectable por su carga, conformación, tamaño o cualquier combinación de los mismos. Los ejemplos de marcas incluyen diversos polímeros. Cada tipo de nucleótido (es decir, A, C, G, T) comprende en general una marca única.

Las marcas se pueden localizar en cualquier posición adecuada en el nucleótido. La **figura 13** muestra un ejemplo de un nucleótido marcado. Aquí, R₁ es en general OH y R₂ es H (es decir, para ADN) u OH (es decir, para ARN), aunque son aceptables otras modificaciones. En la **figura 13**, X es cualquier conector adecuado. En algunos casos, el conector es escindible. Los ejemplos de conectores incluyen, sin limitación, O, NH, S o CH₂. Los ejemplos de grupos químicos adecuados para la posición de Z incluyen O, S o BH₃. La base es cualquier base adecuada para la incorporación en un ácido nucleico que incluye adenina, guanina, citosina, timina, uracilo o un derivado de las mismas. Las bases universales también son aceptables en algunos casos.

El número de fosfatos (n) es cualquier valor entero adecuado (por ejemplo, un número de fosfatos de modo que el nucleótido se pueda incorporar en una molécula de ácido nucleico). En algunos casos, todos los tipos de nucleótidos marcados tienen el mismo número de fosfatos, pero esto no se requiere. En algunas aplicaciones hay una marca diferente para cada tipo de nucleótido y el número de fosfatos no se usa necesariamente para distinguir las diversas marcas. Sin embargo, en algunos casos, más de un tipo de nucleótido (por ejemplo, A, C, T, G o U) tienen la misma molécula marcadora y la capacidad de distinguir un nucleótido de otro se determina, al menos en parte, por el número de fosfatos (con diversos tipos de nucleótidos teniendo un valor diferente para n). En algunos casos, el valor para n es 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o mayor.

Las marcas adecuadas se describen a continuación. En algunos casos, la marca tiene una carga que es de signo inverso con respecto a la carga del resto del compuesto. Cuando se une la marca, la carga en el compuesto global puede ser neutra. La liberación de la marca puede dar como resultado dos moléculas, una marca cargada y un nucleótido cargado. En algunos casos, la marca cargada pasa a través de un nanoporo y se detecta.

En la **figura 14** se muestran más ejemplos de nucleótidos marcados adecuados. La marca se puede unir a la molécula de glúcido, a la molécula de la base o cualquier combinación de las mismas. Con referencia a la **figura 13**, Y es una marca y X es un conector (en algunos casos, escindible). Además, R₁, si está presente, es en general OH, -OCH₂N₃ u -O-2-nitrobencilo, y R₂, si está presente, es en general H. También, Z es en general O, S o BH₃ y n es cualquier número entero que incluye 1, 2, 3 o 4. En algunos casos, la A es O, S, CH₂, CHF, CFF o NH.

Con referencia continua a la **figura 14**, el tipo de base en cada análogo de dNPP es en general diferente del tipo de base en cada uno de los otros tres análogos de dNPP, y el tipo de marca en cada análogo de dNPP es en general diferente del tipo de marca en cada uno de los otros tres análogos de dNPP. Las bases adecuadas incluyen, pero no se limitan a, adenina, guanina, citosina, uracilo o timina, o un derivado de cada una de las mismas. En algunos casos, la base es una de 7-desazaquanina, 7-desazaadenina o 5-metilcitosina.

En los casos en los que R₁ es -O-CH₂N₃, los procedimientos pueden comprender además tratar el análogo de dNPP incorporado para eliminar el -CH₂N₃ y dar como resultado un grupo OH unido en la posición 3', permitiendo de este modo la incorporación de otro análogo de dNPP.

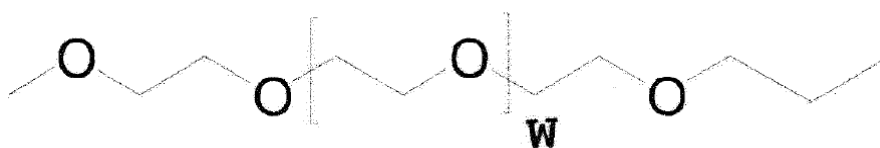
En los casos en los que R₁ es -O-2-nitrobencilo, los procedimientos pueden comprender además tratar el análogo de nucleótido incorporado para eliminar el -2-nitrobencilo y dar como resultado un grupo OH unido en la posición 3', permitiendo de este modo la incorporación de otro análogo de dNPP.

Ejemplos de marcas

Una marca puede ser cualquier grupo químico o molécula que pueda ser detectado en un nanoporo. En algunos casos, una marca comprende uno o más de etilenglicol, un aminoácido, un carbohidrato, un péptido, un tinte, un compuesto quimioluminiscente, un mononucleótido, un dinucleótido, un trinucleótido, un tetranucleótido, un pentanucleótido, un hexanucleótido, un ácido alifático, un ácido aromático, un alcohol, un grupo tiol, un grupo ciano, un grupo nitro, un grupo alquilo, un grupo alqueno, un grupo alquino, un grupo acido o una combinación de los mismos.

También se contempla que la marca comprenda además el número apropiado de lisinas o argininas para equilibrar el número de fosfatos en el compuesto.

En algunos casos, la marca es un polímero. El polietilenglicol (PEG) es un ejemplo de polímero y tiene la estructura como sigue:



Se puede usar cualquier número de unidades de etilenglicol (W). En algunos casos, W es un número entero entre 0 y 100. En algunos casos, el número de unidades de etilenglicol es diferente para cada tipo de nucleótido. En un caso, los cuatro tipos de nucleótidos comprenden marcas que tienen 16, 20, 24 o 36 unidades de etilenglicol. En algunos casos, la marca comprende además un resto identificable adicional, tal como un tinte con base de cumarina. En algunos casos, el polímero está cargado. En algunos casos, el polímero no está cargado y la marca se detecta en una alta concentración de sal (por ejemplo, 3-4 M).

Como se usa en el presente documento, el término "alquilo" incluye grupos de hidrocarburos alifáticos saturados de cadena lineal y ramificada que tienen el número especificado de átomos de carbono y pueden estar sustituidos o no sustituidos. Como se usa en el presente documento, "alqueno" se refiere a un radical hidrocarbonado no aromático, lineal o ramificado, que contiene al menos 1 enlace doble carbono-carbono, y pueden estar presente hasta el número máximo posible de enlaces dobles carbono-carbono no aromáticos, y puede estar sustituido o no sustituido. El término "alquino" se refiere a un radical hidrocarbonado lineal o ramificado, que contiene al menos 1 enlace triple carbono-carbono, y pueden estar presente hasta el número máximo posible de enlaces triples carbono-carbono no aromáticos, y puede estar sustituido o no sustituido. El término "sustituido" se refiere a un grupo funcional como se describe anteriormente, tal como un alquilo o un hidrocarbilo, en el que al menos un enlace a un átomo de hidrógeno contenido en el mismo se reemplaza por un enlace a un átomo distinto de hidrógeno o carbono, siempre que las valencias normales se mantengan y que la(s) sustitución/sustituciones dé/den como resultado un compuesto estable. Los grupos sustituidos también incluyen grupos en los que uno o más enlaces a un átomo de carbono o hidrógeno se reemplazan por uno o más enlaces, que incluyen enlaces dobles o triples, a un heteroátomo.

En algunos casos, la marca solo puede pasar a través del nanoporo en una dirección (por ejemplo, sin invertir la dirección). La marca puede tener una puerta articulada unida a la marca que es lo suficientemente delgada para pasar a través del nanoporo cuando la puerta está alineada con la marca en una dirección, pero no en otra dirección. Con referencia a la **figura 31**, la divulgación proporciona una molécula marcadora que comprende una primera cadena polimérica **3105** que comprende un primer segmento **3110** y un segundo segmento **3115**, donde el segundo segmento es más estrecho que el primer segmento. El segundo segmento puede tener una anchura que es más pequeña que la abertura más estrecha del nanoporo. La molécula marcadora puede incluir una segunda cadena polimérica **3120** que comprende dos extremos, donde un primer extremo se fija a la primera cadena polimérica contigua al segundo segmento y un segundo extremo no se fija a la primera cadena polimérica. La molécula marcadora se puede enhebrar a través de un nanoporo en una primera dirección donde la segunda cadena polimérica se alinea contigua al segundo segmento **3125**. En algunos casos, la molécula marcadora no se puede enhebrar a través del nanoporo en una segunda dirección donde la segunda cadena polimérica no se alinea contigua al segundo segmento **3130**. La segunda dirección puede ser opuesta a la primera dirección.

La primera y/o la segunda cadena polimérica pueden comprender nucleótidos. En algunos casos, la base de la segunda cadena polimérica se empareja con la primera cadena polimérica cuando la segunda cadena polimérica no se alinea contigua al segundo segmento. En algunos casos, la primera cadena polimérica se fija a un nucleótido **3135** (por ejemplo, a un fosfato terminal del nucleótido). La primera cadena polimérica se puede liberar del nucleótido cuando el nucleótido se incorpora en una hebra de ácido nucleico en crecimiento.

El segundo segmento puede comprender cualquier polímero u otra molécula que sea lo suficientemente delgada para pasar a través de un nanoporo cuando se alinea con la puerta (segundo polímero). Por ejemplo, el segundo segmento puede comprender nucleótidos abásicos (es decir, una cadena de ácido nucleico que no tiene ninguna base de ácido nucleico) o una cadena de carbonos.

La divulgación también proporciona un procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor. Con referencia a la **figura 32**, el procedimiento puede comprender proporcionar nucleótidos marcados **3205** en una cámara de reacción que comprende el nanoporo, donde un nucleótido marcado individual de los nucleótidos marcados contiene una marca acoplada a un nucleótido donde la marca es detectable con la ayuda del nanoporo. La marca comprende una primera cadena polimérica que comprende un primer segmento y un segundo segmento, donde el segundo segmento es más estrecho que el primer segmento y una segunda cadena polimérica que comprende dos extremos, donde un primer extremo se fija a la primera cadena polimérica contigua al segundo segmento y un segundo extremo no se fija a la primera cadena polimérica. La molécula marcadora se puede enhebrar a través de un nanoporo en una primera dirección **3210** donde la segunda cadena polimérica se alinea contigua al segundo segmento.

El procedimiento incluye llevar a cabo una reacción de polimerización con la ayuda de una polimerasa **3215**, incorporando de este modo un nucleótido marcado individual de los nucleótidos marcados en una hebra en crecimiento **3220** complementaria a una molécula de ácido nucleico monocatenario **3225** de la muestra de ácido nucleico. El procedimiento puede incluir detectar, con la ayuda del nanoporo **3230**, una marca asociada con el nucleótido marcado individual durante la incorporación del nucleótido marcado individual, en el que la marca se detecta con la ayuda del nanoporo cuando el nucleótido está asociado con la polimerasa.

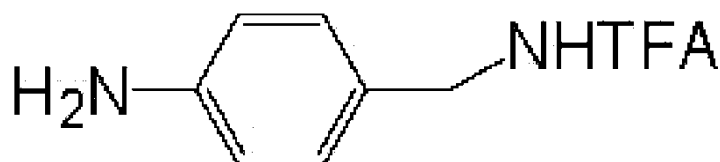
En algunos casos, la molécula marcadora no se puede enhebrar a través del nanoporo en una segunda dirección donde la segunda cadena polimérica no se alinea contigua al segundo segmento.

La marca se puede detectar una pluralidad de veces mientras está asociada con la polimerasa. En algunos casos, un electrodo se recarga entre periodos de detección de marcas. En algunos casos, la marca se enhebra en el nanoporo durante la incorporación del nucleótido marcado individual y la marca no se desenhebra del nanoporo cuando el electrodo se recarga.

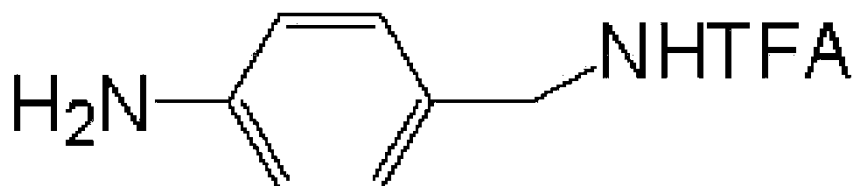
Procedimientos para unir marcas

Se puede usar cualquier procedimiento adecuado para unir las marcas. En un ejemplo, las marcas se pueden unir al fosfato terminal (a) poniendo en contacto un nucleótido trifosfato con diciclohexilcarbodiimida/dimetilformamida en condiciones que permitan la producción de un trimetafosfato cíclico; (b) poniendo en contacto el producto resultante de la etapa a) con un nucleófilo para formar un compuesto funcionalizado con -OH o -NH₂; y (c) haciendo reaccionar el producto de la etapa b) con una marca que tiene un grupo -COR unido a la misma en condiciones que permitan que la marca se una indirectamente a un fosfato terminal, formando de este modo el análogo de nucleótido trifosfato.

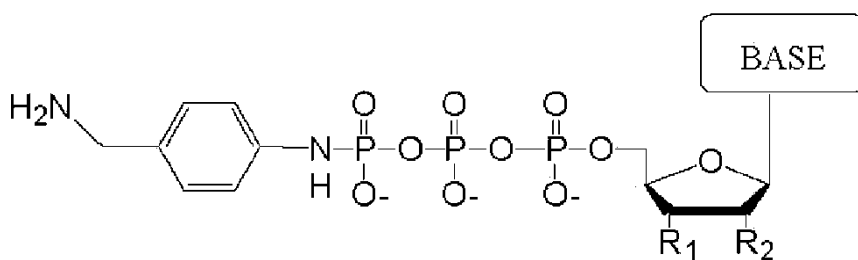
En algunos casos, el nucleófilo es H₂NR-OH, H₂NR-NH₂, R'S-R-OH, R'S-R-NH₂ o



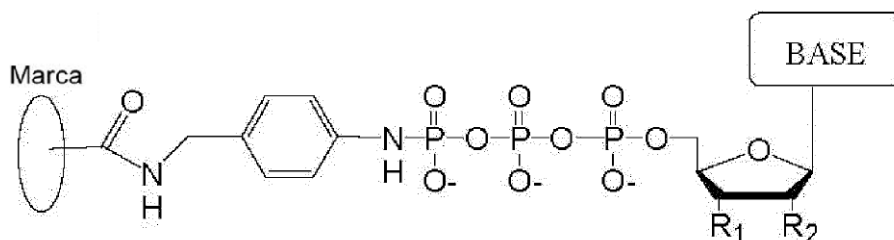
En algunos casos, el procedimiento comprende, en la etapa b), poner en contacto el producto resultante de la etapa a) con un compuesto que tiene la estructura:



y, posteriormente o simultáneamente, poner en contacto el producto con NH₄OH para formar un compuesto que tiene la estructura:



El producto de la etapa b) se puede hacer reaccionar a continuación con una marca que tiene un grupo -COR unido a la misma en condiciones que permitan que la marca se una indirectamente a un fosfato terminal, formando de este modo el análogo de nucleótido trifosfato que tiene la estructura:



en el que R₁ es OH, en el que R₂ es H u OH, en el que la base es adenina, guanina, citosina, timina, uracilo, una 7-desazapurina o una 5-metilpirimidina.

5 **Liberación de marcas**

Se puede liberar una marca de cualquier manera. Se puede liberar una marca durante o después de la incorporación de un nucleótido que tiene la marca en una hebra de ácido nucleico en crecimiento. En algunos casos, la marca se une a polifosfato (por ejemplo, **figura 13**) y la incorporación del nucleótido en una molécula de ácido nucleico da como resultado la liberación de un polifosfato que tiene la marca unida al mismo. La incorporación se puede catalizar con al menos una polimerasa, que se puede unir al nanoporo. En algunos casos, al menos una enzima fosfatasa también se une al poro. La enzima fosfatasa puede escindir la marca del polifosfato para liberar la marca. En algunos casos, las enzimas fosfatasas se sitúan de modo que el pirofosfato producido por la polimerasa en una reacción de la polimerasa interactúe con las enzimas fosfatasas antes de entrar en el poro.

En algunos casos, la marca no se une a polifosfato (véase, por ejemplo, la **figura 14**). En estos casos, la marca se une mediante un conector (X), que puede ser escindible. Los procedimientos para la producción de análogos de nucleótidos protegidos terminalmente de forma escindible y/o unidos de forma escindible se describen en la patente de EE. UU. n.º 6.664.079. El conector no necesita ser escindible.

El conector puede ser cualquier conector adecuado y se puede escindir de cualquier manera adecuada. Los conectores pueden ser fotoescindibles. En un caso, se usa luz UV para escindir fotoquímicamente los conectores y restos fotoquímicamente escindibles. En un caso, el conector fotoescindible es un resto 2-nitrobencilo.

El grupo -CH₂N₃ se puede tratar con TCEP (tris(2-carboxietil)fosfina) para eliminarlo del átomo de O en 3' de un análogo de dNPP, o análogo de rNPP, creando de este modo un grupo OH en 3'.

Detección de marcas

En algunos casos, una polimerasa se extrae de un conjunto de nucleótidos marcados que comprende una pluralidad de bases diferentes (por ejemplo, A, C, G, T y/o U). También es posible poner en contacto de forma iterativa la polimerasa con los diversos tipos de bases marcadas. En este caso, puede no ser necesario que cada tipo de nucleótido tenga una base única, pero el ciclo entre diferentes tipos de bases agrega costes y complejidad al proceso en algunos casos; no obstante, la presente invención abarca este caso.

La **figura 15** muestra que la incorporación del nucleótido marcado en una molécula de ácido nucleico (por ejemplo, usando una polimerasa para extender una base de cebador emparejada a un molde) puede liberar una Marca-polifosfato detectable en algunos casos. En algunos casos, la Marca-polifosfato se detecta a medida que pasa a través del nanoporo. En algunos casos, la Marca-polifosfato se detecta mientras reside en el nanoporo.

En algunos casos, el procedimiento distingue el nucleótido en base al número de fosfatos que comprende el polifosfato (por ejemplo, incluso cuando los marcas son idénticas). Sin embargo, cada tipo de nucleótido tiene en general una marca única.

Con referencia a la **figura 15**, la Marca-polifosfato se puede tratar con fosfatasa (por ejemplo, fosfatasa alcalina) antes de introducir la marca en y/o pasarla a través de un nanoporo y medir la corriente iónica.

Las marcas pueden circular a través de un nanoporo después de ser liberadas del nucleótido. En algunos casos se aplica una tensión para llevar las marcas a través del nanoporo. Al menos aproximadamente un 85 %, al menos un 90 %, al menos un 95 %, al menos un 99 %, al menos un 99,9 % o al menos un 99,99 % de las marcas liberadas se pueden translocar a través del nanoporo.

En algunos casos, las marcas residen en el nanoporo durante un periodo de tiempo donde se detectan. En algunos casos se aplica una tensión para llevar las marcas hacia el nanoporo, detectar las marcas, expulsar las marcas del nanoporo o cualquier combinación de los mismos. Las marcas se pueden liberar o pueden permanecer unidas al nucleótido tras acontecimientos de incorporación de nucleótidos.

La marca se puede detectar en el nanoporo (al menos en parte) debido a su carga. En algunos casos, el compuesto marcador es un compuesto cargado de forma alternativa que tiene una primera carga neta y, después de una reacción química, física o biológica, una segunda carga neta diferente. En algún caso, la magnitud de la carga en la marca es la misma que la magnitud de la carga en el resto del compuesto. En un caso, la marca tiene una carga positiva y la eliminación de la marca cambia la carga del compuesto.

En algunos casos, a medida que la marca pasa a y/o atraviesa el nanoporo, puede generar un cambio electrónico. En algunos casos, el cambio electrónico es un cambio en la amplitud de la corriente, un cambio en la conductancia del nanoporo o cualquier combinación de los mismos.

El nanoporo puede ser biológico o sintético. También se contempla que el poro sea proteínico, por ejemplo, en el que el poro es una proteína hemolisina alfa. Un ejemplo de un nanoporo sintético es un poro en estado sólido o grafeno.

En algunos casos, las enzimas polimerasas y/o las enzimas fosfatasas se unen al nanoporo. Las proteínas de fusión o los enlaces cruzados de disulfuro son ejemplos de procedimientos para unirse a un nanoporo proteínico. En el caso de un nanoporo en estado sólido, la unión a la superficie cerca del nanoporo se puede realizar por medio de enlaces biotina-estreptavidina. En un ejemplo, la ADN polimerasa se une a una superficie sólida por medio de una superficie de oro modificada con una monocapa autoensamblada de alcanotiol funcionalizada con grupos amino, en la que los grupos amino se modifican a ésteres de NHS para la unión a grupos amino en la ADN polimerasa.

El procedimiento se puede realizar a cualquier temperatura adecuada. En algunos casos, la temperatura se encuentra entre 4 °C y 10 °C. En algunos casos, la temperatura es temperatura ambiente.

El procedimiento se puede realizar en cualquier solución y/o tampón adecuados. En algunos casos, el tampón es KCl 300 mM tamponado a pH 7,0 a 8,0 con HEPES 20 mM. En algunos casos, el tampón no comprende cationes divalentes. En algunos casos, el procedimiento no se ve afectado por la presencia de cationes divalentes.

Sistemas informáticos para secuenciar muestras de ácido nucleico

Los sistemas y procedimientos de secuenciación de ácidos nucleicos de la divulgación se pueden regular con la ayuda de sistemas informáticos. La **figura 16** muestra un sistema **1600** que comprende un sistema informático **1601** acoplado a un sistema de secuenciación de ácidos nucleicos **1602**. El sistema informático **1601** puede ser un servidor o una pluralidad de servidores. El sistema informático **1601** se puede programar para regular la preparación y el procesamiento de muestras, y la secuenciación de ácidos nucleicos mediante el sistema de secuenciación **1602**. El sistema de secuenciación **1602** puede ser un secuenciador (o detector) basado en nanoporos, como se describe en otra parte en el presente documento.

El sistema informático se puede programar para implementar los procedimientos de la invención. El sistema informático **1601** incluye una unidad central de procesamiento (CPU, también "procesador" en el presente documento) **1605**, que puede ser un procesador de un único núcleo o de varios núcleos, o una pluralidad de procesadores para procesamiento en paralelo. El sistema informático **1601** también incluye una memoria **1610** (por ejemplo, memoria de acceso aleatorio, memoria de solo lectura, memoria flash), unidad de almacenamiento electrónico **1615** (por ejemplo, disco duro), interfaz de comunicaciones **1620** (por ejemplo, adaptador de red) para que se comunique con uno o más de otros sistemas y dispositivos periféricos **1625**, tales como memoria caché, otra memoria, almacenamiento de datos y/o adaptadores de pantalla electrónica. La memoria **1610**, la unidad de almacenamiento **1615**, la interfaz **1620** y los dispositivos periféricos **1625** están en comunicación con la CPU **1605** a través de un bus de comunicaciones (líneas continuas), tal como una placa base. La unidad de almacenamiento **1615** puede ser una unidad de almacenamiento de datos (o repositorio de datos) para almacenar datos. El sistema informático **1601** puede estar acoplado operativamente a una red informática ("red") con la ayuda de la interfaz de comunicaciones **1620**. La red puede ser Internet, una internet y/o extranet, o una intranet y/o extranet que está en comunicación con Internet. La red puede incluir uno o más servidores informáticos, que pueden permitir el procesamiento distribuido.

Los procedimientos de la invención se pueden implementar por medio de un código ejecutable (o programa informático) por máquina (o procesador informático) almacenado en una localización de almacenamiento electrónico del sistema informático **1601**, tal como, por ejemplo, en la memoria **1610** o unidad de almacenamiento electrónico **1615**. Durante el uso, el procesador **1605** puede ejecutar el código. En algunos casos, el código se puede recuperar de la unidad de almacenamiento **1615** y almacenarse en la memoria **1610** para que el procesador **1605** pueda acceder fácilmente. En algunas situaciones se puede excluir la unidad de almacenamiento electrónico **1615** y las instrucciones ejecutables por máquina se almacenan en la memoria **1610**.

El código puede estar precompilado y configurado para su uso con una máquina que tenga un procesador adaptado para ejecutar el código, o se puede compilar durante el tiempo de ejecución. El código se puede suministrar en un lenguaje de programación que se puede seleccionar para permitir que el código se ejecute de manera precompilada o compilada.

El sistema informático **1601** se puede adaptar para almacenar información de perfiles de usuario, tal como, por ejemplo, un nombre, dirección física, dirección de correo electrónico, número de teléfono, identificador de mensajería instantánea (IM), información académica, información laboral, número de "me gusta" y/o "no me gusta" en redes sociales y otra información de interés potencial para el usuario u otros usuarios. Dicha información de perfiles se puede almacenar en la unidad de almacenamiento **1615** del sistema informático **1601**.

Los aspectos de los sistemas y procedimientos proporcionados en el presente documento, tal como el sistema informático **1601**, se pueden incorporar en la programación. Diversos aspectos de la tecnología se pueden considerar como "productos" o "artículos de elaboración", típicamente en forma de código ejecutable por máquina (o procesador) y/o datos asociados que se incluyen o incorporan en un tipo de medio legible por la máquina. El código ejecutable por máquina se puede almacenar en una unidad de almacenamiento electrónico, tal como una memoria (por ejemplo, ROM, RAM) o un disco duro. Los medios de tipo "almacenamiento" pueden incluir cualquiera o todos de la memoria tangible de los ordenadores, los procesadores o similares o módulos asociados de los mismos, tales como diversas memorias de semiconductores, unidades de cinta, unidades de disco y similares, que pueden proporcionar almacenamiento no transitorio en cualquier momento para la programación del programa informático. En ocasiones, todo el programa informático o partes del mismo se pueden comunicar a través de Internet u otras diversas redes de telecomunicaciones. Dichas comunicaciones, por ejemplo, pueden permitir la carga del programa informático desde un ordenador o procesador a otro, por ejemplo, desde un servidor de administración o equipo host hacia la plataforma informática de un servidor de aplicaciones. Por tanto, otro tipo de medios que pueden soportar los elementos del programa informático incluyen ondas ópticas, eléctricas y electromagnéticas, tales como las que se usan en las interfaces físicas entre dispositivos locales, a través de redes fijas por cable y ópticas y a través de diversos enlaces aéreos. Los elementos físicos que transportan dichas ondas, tales como enlaces por cable o inalámbricos, enlaces ópticos o similares, también se pueden considerar como medios que soportan el programa informático. Como se usa en el presente documento, a menos que esté restringido a medios tangibles de "almacenamiento" no transitorios, las expresiones tales como "medio legible" por ordenador o máquina se refieren a cualquier medio que participe en proporcionar instrucciones a un procesador para su ejecución.

Por tanto, un medio legible por máquina, tal como un código ejecutable por ordenador, puede adoptar muchas formas, que incluyen, pero sin limitarse a, un medio de almacenamiento tangible, un medio de ondas portadoras o un medio de transmisión física. Los medios de almacenamiento no volátiles incluyen, por ejemplo, discos ópticos o magnéticos, tales como cualquiera de los dispositivos de almacenamiento en cualquier ordenador o similar, tales como los que se pueden usar para implantar las bases de datos, etc. que se muestran en los dibujos. Los medios de almacenamiento volátiles incluyen memoria dinámica, tal como la memoria principal de dicha plataforma informática. Los medios de transmisión tangibles incluyen cables coaxiales, cable de cobre y fibra óptica, incluyendo los cables que conforman un bus dentro de un sistema informático. Los medios de transmisión de onda portadora pueden tomar la forma de señales eléctricas o electromagnéticas, u ondas acústicas o lumínicas, tales como las generadas durante las comunicaciones de datos por radiofrecuencia (RF) e infrarrojos (IR). Por lo tanto, las formas frecuentes de medios legibles por ordenador incluyen, por ejemplo, un disquete, un disco flexible, disco duro, cinta magnética, cualquier otro medio magnético, un CD-ROM, DVD o DVD-ROM, cualquier otro medio óptico, tarjetas perforadas, cinta de papel, cualquier otro medio de almacenamiento físico con patrones de orificios, una RAM, una ROM, una PROM y una EPROM, una FLASH-EPROM, cualquier otro chip o cartucho de memoria, una onda portadora que transporte datos o instrucciones, cables o enlaces que transporten dicha onda portadora, o cualquier otro medio desde el que un ordenador pueda leer código y/o datos de programación. Muchas de estas formas de medios legibles por ordenador pueden participar en el transporte de una o más secuencias de una o más instrucciones a un procesador para su ejecución.

Los sistemas y procedimientos de la divulgación se pueden usar para secuenciar diversos tipos de muestras biológicas, tales como ácidos nucleicos (por ejemplo, ADN, ARN) y proteínas. En algunos casos, los procedimientos, dispositivos y sistemas descritos en el presente documento se pueden usar para clasificar muestras biológicas (por ejemplo, proteínas o ácidos nucleicos). Las muestras y/o moléculas clasificadas se pueden dirigir a diversos contenedores para su posterior análisis.

Exactitud de la secuenciación

Los procedimientos proporcionados en el presente documento pueden distinguir con exactitud entre acontecimientos de incorporación de nucleótidos individuales (por ejemplo, acontecimientos de una única molécula). Los procedimientos pueden distinguir con exactitud entre los acontecimientos de incorporación de nucleótidos individuales en un único pase, es decir, sin tener que volver a secuenciar una molécula de ácido nucleico dada. En algunos casos, los procedimientos proporcionados en el presente documento se pueden usar para secuenciar y volver a secuenciar una molécula de ácido nucleico, o para detectar una única vez o múltiples veces una marca asociada con una molécula marcada. Por ejemplo, una marca se puede detectar con la ayuda de un nanoporo al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o 10.000 veces. La marca se puede detectar y volver a detectar con la ayuda de, por ejemplo, una tensión aplicada a una membrana que tiene el nanoporo, que puede llevar la marca hacia el nanoporo o expulsar la marca del nanoporo.

Un procedimiento para la secuenciación de ácidos nucleicos comprende distinguir entre acontecimientos de incorporación de nucleótidos individuales con una exactitud mayor de aproximadamente 4σ . En algunos casos, los acontecimientos de incorporación de nucleótidos se detectan con la ayuda de un nanoporo. Las marcas asociadas con los nucleótidos se pueden liberar tras la incorporación y las marcas pasan a través del nanoporo. En algunos casos, las marcas no se liberan (por ejemplo, se presentan al nanoporo). Aún en más casos, las marcas

se liberan pero residen en el mismo (por ejemplo, no pasan a través del nanoporo). Una marca diferente se puede asociar con y/o liberar de cada tipo de nucleótido (por ejemplo, A, C, T, G) y se detecta por el nanoporo. Los errores incluyen, pero no se limitan a, (a) no detectar una marca, (b) identificar erróneamente una marca, (c) detectar una marca donde no hay marca, (d) detectar marcas en el orden incorrecto (por ejemplo, dos marcas se liberan en un primer orden, pero se detectan en un segundo orden), (e) una marca que no se ha liberado de un nucleótido se detecta como liberada, (f) una marca que no está unida a un nucleótido incorporado se detecta como incorporada a la cadena de nucleótidos en crecimiento, o cualquier combinación de los mismos. En algunos casos, la exactitud de la distinción entre acontecimientos de incorporación de nucleótidos individuales se obtiene restando del 100 % la velocidad a la que se producen los errores (es decir, la tasa de error).

La exactitud de la distinción entre acontecimientos de incorporación de nucleótidos individuales es cualquier porcentaje adecuado. La exactitud de la distinción entre acontecimientos de incorporación de nucleótidos individuales puede ser aproximadamente un 93 %, aproximadamente un 94 %, aproximadamente un 95 %, aproximadamente un 96 %, aproximadamente un 97 %, aproximadamente un 98 %, aproximadamente un 99 %, aproximadamente un 99,5 %, aproximadamente un 99,9 %, aproximadamente un 99,99 %, aproximadamente un 99,999 %, aproximadamente un 99,9999 % y similares. En algunos casos, la exactitud de la distinción entre acontecimientos de incorporación de nucleótidos individuales es al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, al menos un 99,5 %, al menos un 99,9 %, al menos un 99,99 %, al menos un 99,999 %, al menos un 99,9999 % y similares. En algunos casos, la exactitud de la distinción entre acontecimientos de incorporación de nucleótidos individuales se informa en unidades sigma (σ). Sigma es una variable estadística que a veces se usa en la gestión empresarial y la estrategia de fabricación para informar de tasas de error, tales como el porcentaje de productos sin defectos. Aquí, los valores sigma se pueden usar de manera intercambiable con la exactitud de acuerdo con la siguiente relación: 4σ es un 99,38 % de exactitud, 5σ es un 99,977 % de exactitud y 6σ es un 99,99966 % de exactitud.

La distinción entre acontecimientos de incorporación de nucleótidos individuales, de acuerdo con los procedimientos descritos en el presente documento, se puede usar para determinar con exactitud una secuencia de ácido nucleico. En algunos casos, la determinación de la secuencia de ácido nucleico de un ácido nucleico (por ejemplo, ADN y ARN) incluye errores. Los ejemplos de errores incluyen, pero no se limitan a, deleciones (no detectar un ácido nucleico), inserciones (detectar un ácido nucleico donde no hay ninguno presente realmente) y sustituciones (detectar el ácido nucleico incorrecto). La exactitud de la secuenciación de ácidos nucleicos se puede determinar alineando la secuencia de ácido nucleico medida con la secuencia de ácido nucleico verdadera (por ejemplo, de acuerdo con técnicas bioinformáticas) y determinando el porcentaje de posiciones de ácido nucleico que son deleciones, inserciones y/o sustituciones. Los errores son cualquier combinación de deleciones, inserciones y sustituciones. La exactitud varía desde un 0 % a un 100 %, siendo el 100 % una determinación completamente correcta de la secuencia del ácido nucleico. De forma similar, la tasa de error es de un 100 % menos la exactitud y varía desde un 0 % a un 100 %, siendo una tasa de error de un 0 % una determinación completamente correcta de la secuencia del ácido nucleico.

La exactitud de la secuenciación de ácidos nucleicos, según se realiza de acuerdo con los procedimientos y/o el uso de los dispositivos descritos en la presente memoria, es alta. La exactitud es cualquier valor adecuadamente alto. En algunos casos, la exactitud es de aproximadamente un 95 %, aproximadamente un 95,5 %, aproximadamente un 96 %, aproximadamente un 96,5 %, aproximadamente un 97 %, aproximadamente un 97,5 %, aproximadamente un 98 %, aproximadamente un 98,5 %, aproximadamente un 99 %, aproximadamente un 99,5 %, aproximadamente un 99,9 %, aproximadamente un 99,99 %, aproximadamente un 99,999 %, aproximadamente un 99,9999 % y similares. En algunos casos, la exactitud es al menos un 95 %, al menos un 95,5 %, al menos un 96 %, al menos un 96,5 %, al menos un 97 %, al menos un 97,5 %, al menos un 98 %, al menos un 98,5 %, al menos un 99 %, al menos un 99,5 %, al menos un 99,9 %, al menos un 99,99 %, al menos un 99,999 %, al menos un 99,9999 % y similares. En algunos casos, la exactitud se encuentra entre aproximadamente un 95 % y un 99,9999 %, entre aproximadamente un 97 % y un 99,9999 %, entre aproximadamente un 99 % y un 99,9999 %, entre aproximadamente un 99,5 % y un 99,9999 %, entre aproximadamente un 99,9 % y un 99,9999 % y similares.

Se puede lograr una alta exactitud realizando múltiples pases (es decir, secuenciando una molécula de ácido nucleico una pluralidad de veces, por ejemplo, pasando el ácido nucleico a través de o próximo a un nanoporo y secuenciando bases de ácido nucleico de la molécula de ácido nucleico). Se pueden combinar los datos de múltiples pases (por ejemplo, las deleciones, inserciones y/o sustituciones en un primer pase se corrigen usando datos de otros pases repetidos). En algunos casos, la exactitud de la detección de una marca se puede incrementar pasando una marca a través de o contigua a un nanoporo múltiples veces, tal como, por ejemplo, invirtiendo la tensión (por ejemplo, tensión de CC o CA) aplicada al nanoporo o membrana. El procedimiento proporciona una alta exactitud con pocos pases (también denominados lecturas, multiplicidad de cobertura de secuenciación). El número de pases es cualquier número adecuado y no necesita ser un número entero. En algunos casos, la molécula de ácido nucleico se secuencía 1 vez, 2 veces, 3 veces, 4 veces, 5 veces, 6 veces, 7 veces, 8 veces, 9 veces, 10 veces, 12 veces, 14 veces, 16 veces, 18 veces, 20 veces, 25 veces, 30 veces, 35 veces, 40 veces, 45 veces, 50 veces y similares. En algunos casos, la molécula de ácido nucleico se secuencía como máximo 1 vez, como máximo 2 veces, como máximo 3 veces, como máximo 4 veces, como máximo 5 veces, como máximo 6

veces, como máximo 7 veces, como máximo 8 veces, como máximo 9 veces, como máximo 10 veces, como máximo 12 veces, como máximo 14 veces, como máximo 16 veces, como máximo 18 veces, como máximo 20 veces, como máximo 25 veces, como máximo 30 veces, como máximo 35 veces, como máximo 40 veces, como máximo 45 veces, como máximo 50 veces y similares. En algunos casos, la molécula de ácido nucleico se secuencia entre aproximadamente 1 vez y 10 veces, entre aproximadamente 1 vez y 5 veces, entre aproximadamente 1 vez y 3 veces y similares. El nivel de exactitud se puede lograr combinando datos recogidos de un máximo de 20 pases. En algunos casos, el nivel de exactitud se logra combinando datos recogidos de un máximo de 10 pases. En algunos casos, el nivel de exactitud se logra combinando datos recogidos de un máximo de 5 pases. En algunos casos, el nivel de exactitud se logra en un único pase.

La tasa de error es cualquier tasa adecuadamente baja. En algunos casos, la tasa de error es de aproximadamente un 10 %, aproximadamente un 5 %, aproximadamente un 4 %, aproximadamente un 3 %, aproximadamente un 2 %, aproximadamente un 1 %, aproximadamente un 0,5 %, aproximadamente un 0,1 %, aproximadamente un 0,01 %, aproximadamente un 0,001 %, aproximadamente un 0,0001 % y similares. En algunos casos, la tasa de error es como máximo un 10 %, como máximo un 5 %, como máximo un 4 %, como máximo un 3 %, como máximo un 2 %, como máximo un 1 %, como máximo un 0,5 %, como máximo un 0,1 %, como máximo un 0,01 %, como máximo un 0,001 %, como máximo un 0,0001 % y similares. En algunos casos, la tasa de error se encuentra entre un 10 % y un 0,0001 %, entre un 3 % y un 0,0001 %, entre un 1 % y un 0,0001 %, entre un 0,01 % y un 0,0001 % y similares.

Eliminación de secuencias repetitivas

En algunos casos, el ADN genómico puede contener secuencias repetitivas que no son de interés al realizar reacciones de secuenciación de ácidos nucleicos. En el presente documento se proporcionan procedimientos para eliminar estas secuencias repetitivas (por ejemplo, mediante hibridación con secuencias complementarias a las secuencias repetitivas, por ejemplo, ADN Cot-1).

En un modo de realización, el procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor comprende eliminar una secuencia de ácido nucleico repetitiva de la muestra de ácido nucleico para proporcionar una molécula de ácido nucleico monocatenario para la secuenciación. El procedimiento puede incluir además proporcionar nucleótidos marcados en una cámara de reacción que comprende el nanoporo, donde un nucleótido marcado individual de los nucleótidos marcados contiene una marca acoplada a un nucleótido que es detectable con la ayuda del nanoporo. En algunos casos, el procedimiento incluye llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de los nucleótidos marcados en una hebra en crecimiento complementaria a la molécula de ácido nucleico monocatenario. El procedimiento puede incluir detectar, con la ayuda del nanoporo, una marca asociada con el nucleótido marcado individual durante la incorporación del nucleótido marcado individual, en el que la marca se detecta con la ayuda del nanoporo cuando el nucleótido está asociado con la polimerasa.

En algunos casos, las secuencias repetitivas no se eliminan físicamente de la reacción, sino que se hace que no se puedan secuenciar y se dejan en la mezcla de reacción (por ejemplo, mediante hibridación con ADN Cot-1, haciendo que las secuencias repetitivas sean bicatenarias y se "eliminen" eficazmente de la reacción de secuenciación). En algunos casos, las secuencias repetitivas están hechas para ser bicatenarias.

La secuencia repetitiva puede tener cualquier longitud adecuada. En algunos casos, la secuencia de ácido nucleico repetitiva comprende aproximadamente 20, aproximadamente 40, aproximadamente 60, aproximadamente 80, aproximadamente 100, aproximadamente 200, aproximadamente 400, aproximadamente 600, aproximadamente 800, aproximadamente 1000, aproximadamente 5000, aproximadamente 10.000 o aproximadamente 50.000 bases de ácido nucleico. En algunos casos, la secuencia de ácido nucleico repetitiva comprende al menos aproximadamente 20, al menos aproximadamente 40, al menos aproximadamente 60, al menos aproximadamente 80, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 400, al menos aproximadamente 600, al menos aproximadamente 800, al menos aproximadamente 1000, al menos aproximadamente 5000, al menos aproximadamente 10.000 o al menos aproximadamente 50.000 bases de ácido nucleico. En algunos casos, las bases son consecutivas.

La secuencia de ácido nucleico repetitiva puede tener cualquier número de subunidades repetidas. En algunos casos, las subunidades repetidas son consecutivas. En algunos casos, la secuencia de ácido nucleico repetitiva comprende aproximadamente 20, aproximadamente 40, aproximadamente 60, aproximadamente 80, aproximadamente 100, aproximadamente 200, aproximadamente 400, aproximadamente 600, aproximadamente 800, o aproximadamente 1000 subunidades repetidas de bases de ácido nucleico. En algunos casos, la secuencia de ácido nucleico repetitiva comprende al menos aproximadamente 20, al menos aproximadamente 40, al menos aproximadamente 60, al menos aproximadamente 80, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 400, al menos aproximadamente 600, al menos aproximadamente 800, o al menos aproximadamente 1000 subunidades repetidas de bases de ácido nucleico.

En algunos casos, la secuencia de ácido nucleico repetitiva se elimina mediante hibridación con una secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva. La secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva se puede inmovilizar sobre un soporte sólido tal como una superficie o una microesfera. En algunos casos, la secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva comprende ADN Cot-1 (que es un ejemplo de secuencias de ácido nucleico repetitivas que tienen una longitud de entre aproximadamente 50 y aproximadamente 100 bases de ácido nucleico).

Ensamblaje e inserción de nanoporos

Los procedimientos descritos en el presente documento pueden usar un nanoporo que tiene una polimerasa unida al nanoporo. En algunos casos, es deseable tener una y solo una polimerasa por nanoporo (por ejemplo, de modo que solo se secuencie una molécula de ácido nucleico en cada nanoporo). Sin embargo, muchos nanoporos, incluyendo hemolisina alfa (α HL), pueden ser proteínas multiméricas que tienen una pluralidad de subunidades (por ejemplo, 7 subunidades para α HL). Las subunidades pueden ser copias idénticas del mismo polipéptido. En el presente documento se proporcionan proteínas multiméricas (por ejemplo, nanoporos) que tienen una proporción definida entre subunidades modificadas y subunidades no modificadas. En el presente documento también se proporcionan procedimientos para producir proteínas multiméricas (por ejemplo, nanoporos) que tienen una proporción definida entre subunidades modificadas y subunidades no modificadas.

Con referencia a la **figura 27**, un procedimiento para ensamblar una proteína que tiene una pluralidad de subunidades comprende proporcionar una pluralidad de primeras subunidades **2705** y proporcionar una pluralidad de segundas subunidades **2710**, donde las segundas subunidades están modificadas en comparación con las primeras subunidades. En algunos casos, las primeras subunidades son naturales (por ejemplo, purificadas a partir de fuentes naturales o producidas de forma recombinante). Las segundas subunidades se pueden modificar de cualquier forma adecuada. En algunos casos, las segundas subunidades tienen una proteína (por ejemplo, una polimerasa) unida (por ejemplo, como una proteína de fusión). Las subunidades modificadas pueden comprender un resto químicamente reactivo (por ejemplo, una acida o un grupo alquino adecuado para formar un enlace). En algunos casos, el procedimiento comprende además realizar una reacción (por ejemplo, una cicloadición química *Click*) para unir una entidad (por ejemplo, una polimerasa) al resto químicamente reactivo.

El procedimiento puede comprender además poner en contacto las primeras subunidades con las segundas subunidades **2715** en una primera proporción para formar una pluralidad de proteínas **2720** que tienen las primeras subunidades y las segundas subunidades. Por ejemplo, una parte de subunidades de α HL modificadas que tienen un grupo reactivo adecuado para unirse a una polimerasa se puede mezclar con seis partes de subunidades de α HL natural (es decir, siendo la primera proporción 1:6). La pluralidad de proteínas puede tener una pluralidad de proporciones entre las primeras subunidades y las segundas subunidades. Por ejemplo, las subunidades mixtas pueden formar varios nanoporos que tienen una distribución de estequiometrías entre subunidades modificadas y no modificadas (por ejemplo, 1:6, 2:5, 3:4).

En algunos casos, las proteínas se forman simplemente mezclando las subunidades. En el caso de nanoporos de α HL, por ejemplo, un detergente (por ejemplo, ácido desoxicólico) puede provocar que el monómero de α HL adopte la conformación de poro. Los nanoporos también se pueden formar usando un lípido (por ejemplo, 1,2-difitanil-sn-glicero-3-fosfolina (DPhPC) o 1,2-di-O-fitanil-sn-glicero-3-fosfolina (DoPhPC)) y una temperatura moderada (por ejemplo, menos de aproximadamente 100 °C). En algunos casos, mezclar DPhPC con una solución tampón crea grandes vesículas multilaminares (LMV), y añadir subunidades de α HL a esta solución e incubar la mezcla a 40 °C durante 30 minutos da como resultado la formación de poros.

Si se usan dos tipos diferentes de subunidades (por ejemplo, la proteína natural y un segundo monómero de α HL que puede contener una única mutación puntual), las proteínas resultantes pueden tener una estequiometría mixta (por ejemplo, de las proteínas naturales y mutantes). La estequiometría de estas proteínas puede seguir una fórmula que es dependiente de la proporción de las concentraciones de las dos proteínas usadas en la reacción de formación de poros. Esta fórmula es como sigue:

$$100 P_m = 100[n!/m!(n-m)!] \cdot f_{mut}^m \cdot f_{wt}^{n-m}, \text{ donde}$$

P_m = probabilidad de que un poro tenga un número m de subunidades mutantes

n = número total de subunidades (por ejemplo, 7 para α HL)

m = número de subunidades "mutantes"

f_{mut} = fracción o proporción de subunidades mutantes mezcladas juntas

f_{wt} = fracción o proporción de subunidades naturales mezcladas juntas

El procedimiento puede comprender además fraccionar la pluralidad de proteínas para enriquecer proteínas que tienen una segunda proporción entre las primeras subunidades y las segundas subunidades **2725**. Por ejemplo, se pueden aislar proteínas de nanoporo que tienen una y solo una subunidad modificada (por ejemplo, una segunda proporción de 1:6). Sin embargo, cualquier segunda proporción es adecuada. También se puede fraccionar una distribución de segundas proporciones, tal como enriqueciendo proteínas que tienen una o bien dos subunidades modificadas. El número total de subunidades que forman la proteína no siempre es 7 (por ejemplo, se puede usar un nanoporo diferente o se puede formar un nanoporo de hemolisina alfa que tiene seis subunidades) como se representa en la **figura 27**. En algunos casos, se enriquecen las proteínas que tienen solo una subunidad modificada. En dichos casos, la segunda proporción es de 1 segunda subunidad por (n-1) primeras subunidades, donde n es el número de subunidades que comprenden la proteína.

La primera proporción puede ser la misma que la segunda proporción, sin embargo, esto no se requiere. En algunos casos, las proteínas que tienen monómeros mutados se pueden formar menos eficazmente que las que no tienen subunidades mutadas. Si este es el caso, la primera proporción puede ser mayor que la segunda proporción (por ejemplo, si se desea una segunda proporción de 1 subunidad mutada por 6 no mutadas en un nanoporo, formar un número adecuado de proteínas 1:6 puede requerir mezclar las subunidades en una proporción mayor que 1:6).

Las proteínas que tienen diferentes segundas proporciones de subunidades se pueden comportar de forma diferente (por ejemplo, tienen diferentes tiempos de retención) en una separación. En algunos casos, las proteínas se fraccionan usando cromatografía, tal como cromatografía de intercambio iónico o cromatografía de afinidad. Puesto que las primeras y segundas subunidades pueden ser idénticas excepto por la modificación, el número de modificaciones en la proteína puede servir como base para la separación. En algunos casos, las primeras o bien segundas subunidades tienen una marca de purificación (por ejemplo, además de la modificación) para permitir o mejorar la eficacia del fraccionamiento. En algunos casos se usa una marca de polihistidina (marca His), una marca de estreptavidina (marca Strep) u otra marca peptídica. En algunos casos, las primeras y segundas subunidades comprenden cada una diferentes marcas y la etapa de fraccionamiento fracciona en base a cada marca. En el caso de una marca His, se crea una carga en la marca a pH bajo (los residuos de histidina se cargan positivamente por debajo del pKa de la cadena lateral). Con una diferencia significativa de carga en una de las moléculas de α HL en comparación con las otras, se puede usar cromatografía de intercambio iónico para separar los oligómeros que tienen 0, 1, 2, 3, 4, 5, 6 o 7 de las subunidades de α HL "con marca cargada". En principio, esta marca cargada puede ser una cadena de cualquier aminoácido que porte una carga uniforme. La **figura 28** y la **figura 29** muestran ejemplos de fraccionamiento de nanoporos en base a una marca His. La **figura 28** muestra una curva de absorbancia ultravioleta a 280 nanómetros, absorbancia ultravioleta a 260 nanómetros y conductividad. Los picos corresponden a nanoporos con diversas proporciones de subunidades modificadas y no modificadas. La **figura 29** muestra el fraccionamiento de nanoporos de α HL y mutantes de los mismos usando tanto marca His como marcas Strep.

En algunos casos, una entidad (por ejemplo, una polimerasa) se une a la proteína después del fraccionamiento **2730**. La proteína puede ser un nanoporo y la entidad puede ser una polimerasa. En algunos casos, el procedimiento comprende además insertar las proteínas que tienen las subunidades de la segunda proporción en una bicapa.

En algunas situaciones, un nanoporo puede comprender una pluralidad de subunidades. Se puede unir una polimerasa a una de las subunidades y al menos una y menos que todas las subunidades comprenden una primera marca de purificación. En algunos ejemplos, el nanoporo es hemolisina alfa o una variante de la misma. En algunos casos, todas las subunidades comprenden una primera marca de purificación o una segunda marca de purificación. La primera marca de purificación puede ser una marca de polihistidina (por ejemplo, en la subunidad que tiene la polimerasa unida).

Conectores

Los procedimientos descritos en el presente documento pueden usar una enzima (por ejemplo, polimerasa) unida a un nanoporo para la detección por el nanoporo, que incluye la secuenciación de ácidos nucleicos. En algunos casos, el enlace entre la enzima y el nanoporo puede afectar al rendimiento del sistema. Por ejemplo, diseñar una unión de la ADN polimerasa al poro (hemolisina alfa) puede incrementar la concentración eficaz de nucleótidos marcados, disminuyendo de este modo la barrera entrópica. En algunos casos, la polimerasa se une directamente al nanoporo. En otros casos se usa un conector entre la polimerasa y el nanoporo.

La secuenciación de marcas descrita en el presente documento se puede beneficiar de una captura eficaz de un nucleótido con marca específico en el poro de α HL inducida por un potencial eléctrico. La captura puede tener lugar durante o después de la extensión del cebador de polimerasa en base al molde de ADN. Un procedimiento para mejorar la eficacia de la captura es optimizar la conexión entre la polimerasa y el poro de α HL. Sin limitación, tres rasgos característicos de la conexión que hay que optimizar son: (a) la longitud de la conexión (que puede incrementar la concentración eficaz de nucleótidos marcados, afectar a la cinética de captura y/o cambiar la barrera entrópica); (b) la flexibilidad de conexión (que puede influir en la cinética de los cambios conformacionales del conector); y (c) el número y la localización de las conexiones entre la polimerasa y el nanoporo (que pueden reducir

el número de estados conformacionales disponibles, incrementando de este modo la probabilidad de una orientación adecuada poro-polimerasa, incrementar la concentración eficaz de nucleótidos marcados y reducir la barrera entrópica).

- 5 La enzima y la polimerasa se pueden conectar de cualquier manera adecuada. En algunos casos, los marcos abiertos de lectura (ORF) se fusionan, directamente o con un conector de aminoácidos. La fusión puede ser en cualquier orden. En algunos casos, se forma un enlace químico (por ejemplo, mediante química *Click*). En algunos casos, la conexión no es covalente (por ejemplo, grapas moleculares, a través de interacciones biotina-estreptavidina o a través de marcas proteína-proteína tales como PDZ, GBD, SpyTag, HaloTag o ligandos SH3).
- 10 En algunos casos, el conector es un polímero tal como un péptido, ácido nucleico, polietilenglicol (PEG). El conector puede tener cualquier longitud adecuada. Por ejemplo, el conector puede tener aproximadamente 5 nanómetros (nm), aproximadamente 10 nm, aproximadamente 15 nm, aproximadamente 20 nm, aproximadamente 40 nm, aproximadamente 50 nm o aproximadamente 100 nm de longitud. En algunos casos, el conector tiene al menos
- 15 aproximadamente 5 nanómetros (nm), al menos aproximadamente 10 nm, al menos aproximadamente 15 nm, al menos aproximadamente 20 nm, al menos aproximadamente 40 nm, al menos aproximadamente 50 nm o al menos aproximadamente 100 nm de longitud. En algunos casos, el conector tiene como máximo aproximadamente 5 nanómetros (nm), como máximo aproximadamente 10 nm, como máximo aproximadamente 15 nm, como máximo aproximadamente 20 nm, como máximo aproximadamente 40 nm, como máximo aproximadamente 50 nm
- 20 o como máximo aproximadamente 100 nm de longitud. El conector puede ser rígido, flexible o cualquier combinación de los mismos. En algunos casos no se usa un conector (por ejemplo, la polimerasa se une directamente al nanoporo).
- 25 En algunos casos, más de un conector conecta la enzima con el nanoporo. El número y la localización de la conexión entre la polimerasa y el nanoporo pueden variar. Los ejemplos incluyen: extremo C de α HL a extremo N de polimerasa; extremo N de α HL a extremo C de polimerasa y conexiones entre aminoácidos que no están en los extremos.
- 30 En un aspecto, un procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor comprende proporcionar nucleótidos marcados en una cámara de reacción que comprende el nanoporo, donde un nucleótido marcado individual de los nucleótidos marcados contiene una marca acoplada a un nucleótido que es detectable con la ayuda del nanoporo. El procedimiento incluye llevar a cabo una reacción de polimerización con la ayuda de una polimerasa unida mediante un conector al nanoporo, incorporando de este modo un nucleótido marcado individual de los nucleótidos marcados en una
- 35 hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de la muestra de ácido nucleico. El procedimiento incluye detectar, con la ayuda del nanoporo, una marca asociada con el nucleótido marcado individual durante la incorporación del nucleótido marcado individual, donde la marca se detecta con la ayuda del nanoporo cuando el nucleótido está asociado con la polimerasa.
- 40 En algunos casos, el conector orienta la polimerasa con respecto al nanoporo de modo que la marca se detecte con la ayuda del nanoporo. En algunos casos, la polimerasa se une al nanoporo mediante dos o más conectores.
- 45 En algunos casos, el conector comprende uno o más de los SEQ ID N.º 2-35 o un producto de PCR producido a partir de los mismos. En algunos casos, el conector comprende el péptido codificado por uno o más de los SEQ ID N.º 2-35 o un producto de PCR producido a partir de los mismos.

SEQ. ID. N.º	Denominación	Secuencia (5'-3')
2	1-GGGGSG_For	GATCG GGAGGAGGTGGGAGCGGA G
3	1-GGGGSG_Rev	GATCC TCCGCTCCACCTCCTCC C
4	2-GGGGS_For	GATCG GGAGGAGGTGGGAGCGGAGGAGGTGGGAGCGGA G
5	2-GGGGS_Rev	GATCC TCCGCTCCACCTCCTCCGCTCCACCTCCTCC C
6	3-GGGGS_For	GGAGGAGGTGGGAGCGGAGGAGGTGGGAGCGGAGGAGGTGGGA GCGGA G
7	3-GGGGS_Rev	GATCC TCCGCTCCACCTCCTCCGCTCCACCTCCTCCGCTCCACCTCCT CC C
8	4-GGGGS_For	GGAGGAGGTGGGAGCGGAGGAGGTGGGAGCGGAGGAGGTGGGA GCGGAGGAGGTGGGAGCGGA G
9	4-GGGGS_Rev	TCCGCTCCACCTCCTCCGCTCCACCTCCTCCGCTCCACCTCCT CCGCTCCACCTCCTCC C

SEQ. ID. N.º	Denominación	Secuencia (5'-3')
10	5-GGGGS_For	GATCG GGAGGAGGTGGGAGCGGAGGAGGTGGGAGCGGAGGAGGTGGGA GCGGAGGAGGTGGGAGCGGAGGAGGTGGGAGCGGA G
11	5-GGGGS_Rev	GATCC TCCGCTCCCACCTCCTCCGCTCCCACCTCCTCCGCTCCCACCTCCT CCGCTCCCACCTCCTCCGCTCCCACCTCCTCC C
12	6-GGGGS_For	GATCG GGAGGAGGTGGGAGCGGAGGAGGTGGGAGCGGAGGAGGTGGGA GCGGAGGAGGTGGGAGCGGAGGAGGTGGGAGCGGAGGAGGAGG TGGGAGCGGA G
13	6-GGGGS_Rev	GATCC TCCGCTCCCACCTCCTCCTCCGCTCCCACCTCCTCCGCTCCCACCT CCTCCGCTCCCACCTCCTCCGCTCCCACCTCCTCCGCTCCCACCT CCTCC C
14	2-EAAAK_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGA G
15	2-EAAAK_Rev	GATCC TCCTTTGCTGCCGCTTCTTTGCTGCCGCTTCTCC C
16	4-EAAAK_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGC GAAAGAAGCGGCAGCGAAAGGA G
17	4-EAAAK_Rev	GATCC TCCTTTGCTGCCGCTTCTTTGCTGCCGCTTCTTTGCTGCCGCT TCTTTGCTGCCGCTTCTCC C
18	4-EAAAK_Flex_For	GATCGGGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGAGGA GAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGAG
19	4-EAAAK_Flex_Rev	GATCCTCCTTTGCTGCCGCTTCTTTGCTGCCGCTTCTCCTCCTT TCGCTGCCGCTTCTTTGCTGCCGCTTCTCC
20	6-EAAAK_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGC GAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAG CGAAAGGA G
21	6-EAAAK_Rev	GATCC TCCTTTGCTGCCGCTTCTTTGCTGCCGCTTCTTTGCTGCCGCT TCTTTGCTGCCGCTTCTTTGCTGCCGCTTCTTTGCTGCCGCTT CTCC C
22	6-EAAAK_Flex_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGAGGAGAAGC GGCAGCGAAAGAAGCGGCAGCGAAAGGAGGAGAAGCGGCAGCGA AAGAAGCGGCAGCGAAAGGA G
23	6-EAAAK_Flex_Rev	GATCC TCCTTTGCTGCCGCTTCTTTGCTGCCGCTTCTCCTCCTTTGCT GCCGCTTCTTTGCTGCCGCTTCTCCTCCTTTGCTGCCGCTTCTT TCGCTGCCGCTTCTCC C
24	8-EAAAK_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGC GAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAG CGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGA G
25	8-EAAAK_Rev	GATCC TCCTTTGCTGCCGCTTCTTTGCTGCCGCTTCTTTGCTGCCGCT TCTTTGCTGCCGCTTCTTTGCTGCCGCTTCTTTGCTGCCGCTT CTTTGCTGCCGCTTCTTTGCTGCCGCTTCTCC C

SEQ. ID. N.º	Denominación	Secuencia (5'-3')
26	8-EAAAK_Flex_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGC GAAAGAAGCGGCAGCGAAAGGAGGAGAAGCGGCAGCGAAAGAAG CGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGA G
27	8-EAAAK_Flex_Rev	GATCC TCCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCT TCTTTTCGCTGCCGCTTCTCCTCCTTTTCGCTGCCGCTTCTTTTCGCTG CCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTCC C
28	10-EAAAK_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGC GAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAG CGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCA GCGAAAGAAGCGGCAGCGAAAGGA G
29	10-EAAAK_Rev	GATCC TCCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCT TCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTT CTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTT TTTCGCTGCCGCTTCTCC C
30	10-EAAAK_Flex_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGAGGAGAAGC GGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAG CGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGA GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGA G
31	10-EAAAK_Flex_Rev	GATCC TCCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTCCTCCTTTTCGCT GCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTG CCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTCCTCCTTT CGCTGCCGCTTCTTTTCGCTGCCGCTTCTCC C
32	12-EAAAK_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGC GAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAG CGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCA GCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGC AGCGAAAGGA G
33	12-EAAAK_Rev	GATCC TCCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCT TCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTT CTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTT TTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCT CC C
34	12-EAAAK_Flex_For	GATCG GGAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGC GAAAGAAGCGGCAGCGAAAGGAGAAGCGGCAGCGAAAGAAGCGG CAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGGAGAA GCGGCAGCGAAAGAAGCGGCAGCGAAAGAAGCGGCAGCGAAAGA AGCGGCAGCGAAAGGA G
35	12-EAAAK_Flex_Rev	GATCC TCCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCT TCTTTTCGCTGCCGCTTCTCCTTTTCGCTGCCGCTTCTTTTCGCTGCCG CTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTCCTTTTCGCTGC CGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCCGCTTCTTTTCGCTGCC GCTTCTCC C

Calibración de la tensión aplicada

Las señales de salida específicas de la molécula de los dispositivos sensores de nanoporos de una única molécula se pueden originar por la presencia de una diferencia de potencial electroquímico a través de una membrana iónicamente impermeable rodeada por una solución electrolítica. Esta diferencia de potencial transmembrana puede determinar la fuerza de la corriente electroquímica específica de nanoporos que puede detectar la electrónica dentro del dispositivo por medio de reacciones de protección (es decir, faradaicas) o de no protección (es decir, capacitivas) que se producen en las superficies del electrodo.

Para cualquier estado dado del nanoporo (es decir, canal abierto, estado capturado, etc.), el potencial transmembrana dependiente del tiempo puede actuar como una señal de entrada que puede determinar la corriente resultante que fluye a través del complejo de nanoporos como una función del tiempo. Esta corriente de nanoporo puede proporcionar la salida de señal molecular específica del dispositivo sensor de nanoporos. La corriente de nanoporo de canal abierto se puede modular a grados variables por las interacciones entre el nanoporo y las moléculas capturadas que bloquean parcialmente el flujo de iones a través del canal.

Estas modulaciones pueden presentar especificidad para el tipo de molécula que se ha capturado, permitiendo que algunas moléculas se identifiquen directamente a partir de sus modulaciones de corriente de nanoporo. Para un tipo de molécula dado y un conjunto fijo de condiciones del dispositivo, el grado de modulación de la corriente de nanoporo de canal abierto por una molécula capturada de este tipo puede variar dependiendo del potencial transmembrana aplicado, mapeando cada tipo de molécula a una curva particular de corriente frente a tensión (I-V).

Las compensaciones sistemáticamente variables entre los ajustes de tensión aplicada y el potencial transmembrana pueden introducir desviaciones horizontales de esta curva I-V a lo largo del eje de tensión horizontal, reduciendo potencialmente la exactitud de la identificación molecular en base a las señales de corriente medidas informadas por el dispositivo sensor de nanoporos como una señal de salida. Por lo tanto, la compensación no controlada entre el potencial aplicado y el potencial transmembrana puede ser problemática para comparar con exactitud las mediciones de la misma molécula en las mismas condiciones.

Esta llamada "compensación de potencial" entre el potencial aplicado externamente y el potencial transmembrana real puede variar tanto dentro de los experimentos como entre ellos. Las variaciones en la compensación de potencial se pueden producir tanto por variaciones en las condiciones iniciales como por variaciones dependientes del tiempo (deriva) en las condiciones electroquímicas dentro del dispositivo sensor de nanoporos.

La eliminación de estos errores de medición se puede realizar como se describe aquí calibrando la compensación dependiente del tiempo entre la tensión aplicada y el potencial transmembrana para cada experimento. Físicamente, la probabilidad de observar acontecimientos de escape de moléculas capturadas en nanoporos puede depender del potencial transmembrana aplicado y esta distribución de probabilidad puede ser la misma para muestras idénticas de moléculas en las mismas condiciones (por ejemplo, la muestra puede ser una mezcla de diferentes tipos de moléculas siempre que sus proporciones no varíen entre muestras). En algunos casos, la distribución de tensiones donde se producen acontecimientos de escape para un tipo de muestra fijo proporciona una medida de la compensación entre el potencial aplicado y el potencial transmembrana. Esta información se puede usar para calibrar la tensión aplicada a través del nanoporo, eliminando las fuentes sistemáticas de error provocadas por compensaciones de potencial dentro y entre experimentos y mejorando la exactitud de la identificación molecular y otras mediciones.

Para un aparato sensor de nanoporos dado que funciona con la misma muestra molecular y los mismos reactivos, el valor esperado de la distribución de las tensiones de escape se puede estimar a partir de una muestra estadística de los acontecimientos de escape molecular individuales (aunque cada acontecimiento individual puede ser un proceso estocástico sujeto a fluctuaciones aleatorias). Esta estimación puede depender del tiempo para justificar la deriva temporal de la compensación de potencial dentro del experimento. Esto puede corregir la diferencia variable entre los ajustes de tensión aplicada y la tensión real sentida en el poro, "alineando" eficazmente todas las mediciones horizontalmente cuando se representan gráficamente en el espacio I-V.

En algunos casos, la calibración de compensación de potencial (es decir, tensión) no justifica la ganancia de corriente y las variaciones de compensación de corriente, que también se pueden calibrar para mejorar la exactitud y la reproducibilidad de las mediciones de corriente de nanoporo. Sin embargo, la calibración de compensación de potencial se realiza en general antes de la corrección de ganancia y compensación para evitar errores al estimar la ganancia de corriente y las variaciones de compensación de corriente, ya que estas, a su vez, pueden implicar el ajuste de las curvas de corriente frente a tensión (I-V), y los resultados de estos ajustes se ven afectados por variaciones en la compensación de tensión. Es decir, desviar los datos de izquierda a derecha (horizontalmente) en el espacio I-V puede introducir errores en los ajustes posteriores de ganancia de corriente y compensación de corriente.

La **figura 30** muestra una curva de la corriente a través del nanoporo (líneas continuas) y la tensión aplicada (líneas discontinuas) frente al tiempo. La corriente puede disminuir cuando se captura una molécula en el nanoporo **3005**. A medida que la tensión aplicada disminuye con el tiempo **3010**, la corriente disminuye hasta que la molécula sale

del nanoporo **3015**, momento en el que la corriente se incrementa al nivel esperado en la tensión aplicada. La tensión aplicada a la que sale la molécula puede depender de la longitud de la molécula. Por ejemplo, una marca con 30 bases puede salir a alrededor de 40 mV, mientras que una marca con 50 bases puede salir a alrededor de 10 mV. Puede haber una variación de **3020** en la tensión de salida para diferentes nanoporos o para diferentes mediciones en el mismo nanoporo con el tiempo. Ajustar esta tensión de salida a un valor esperado puede hacer que los datos sean más fáciles de interpretar y/o más exactos.

En el presente documento se describe un procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor. El procedimiento puede incluir proporcionar nucleótidos marcados en una cámara de reacción que comprende un nanoporo, donde un nucleótido marcado individual de los nucleótidos marcados contiene una marca acoplada a un nucleótido que es detectable con la ayuda del nanoporo. El procedimiento puede incluir llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de los nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de la muestra de ácido nucleico. El procedimiento puede incluir a continuación detectar, con la ayuda del nanoporo, una marca asociada con el nucleótido marcado individual durante la incorporación del nucleótido marcado individual, donde la marca se detecta con la ayuda del nanoporo cuando el nucleótido está asociado con la polimerasa. En algunos casos, la detección comprende aplicar una tensión aplicada a través del nanoporo y medir una corriente con el electrodo sensor a la tensión aplicada.

En algunos casos se calibra la tensión aplicada. La calibración puede incluir la estimación de una distribución esperada de tensión de escape frente al tiempo para el electrodo sensor. La calibración puede calcular a continuación una diferencia entre la distribución de tensiones de escape esperadas y un punto de referencia (por ejemplo, un punto de referencia arbitrario, tal como cero). A continuación, la calibración puede desviar la tensión aplicada en la diferencia calculada. En algunos casos, la tensión aplicada disminuye con el tiempo.

En algunos casos, se estima una distribución de las tensiones de escape esperadas frente al tiempo. En algunos casos, el punto de referencia es cero voltios. El procedimiento puede eliminar las variaciones detectadas en la distribución de tensiones de escape esperadas. En algunos casos, el procedimiento se realiza en una pluralidad de nanoporos direccionables independientemente, cada uno contiguo a un electrodo sensor.

En algunos casos, la presencia de la marca en el nanoporo reduce la corriente medida con el electrodo sensor a la tensión aplicada. En algunos casos, los nucleótidos marcados comprenden una pluralidad de marcas diferentes y el procedimiento detecta cada una de la pluralidad de marcas diferentes.

En algunos casos, la calibración incrementa la exactitud del procedimiento en comparación con la realización del procedimiento sin calibración. En algunos casos, la calibración compensa los cambios en las condiciones electroquímicas con el tiempo. En algunos casos, la calibración compensa que diferentes nanoporos tengan diferentes condiciones electroquímicas en un dispositivo que tiene una pluralidad de nanoporos. En algunos casos, la calibración compensa diferentes condiciones electroquímicas para cada realización del procedimiento. En algunos casos, el procedimiento comprende además calibrar variaciones en una ganancia de corriente y/o variaciones en una compensación de corriente.

Procedimientos de secuenciación por extensión

La presente divulgación proporciona procedimientos para secuenciar una molécula de ácido nucleico usando secuenciación por extensión. La secuenciación por extensión implica una serie de etapas que crean un polímero extendido que es más largo que el ácido nucleico que se va a secuenciar y tiene una secuencia derivada de la molécula de ácido nucleico que se va a secuenciar. El polímero extendido se puede enhebrar a través de un nanoporo para determinar su secuencia. Como se describe en el presente documento, el polímero extendido puede tener puertas de modo que el polímero extendido pueda pasar a través del nanoporo en una única dirección. La etapas del procedimiento se ilustran en las **figuras 33 a 36**. Se puede encontrar más información con respecto a la secuenciación de ácidos nucleicos por extensión (es decir, extensores) en la patente de EE. UU. n.º 8.324.360.

Con referencia a la **figura 33**, en un aspecto, un procedimiento para secuenciar ácidos nucleicos comprende proporcionar un ácido nucleico monocatenario que se va a secuenciar y proporcionar una pluralidad de sondas. Las sondas comprenden un resto de hibridación **3305** capaz de hibridar con el ácido nucleico monocatenario, una estructura de bucle **3310** que tiene dos extremos, donde cada extremo se une al resto de hibridación, y un grupo escindible **3315** localizado en el resto de hibridación entre los extremos de la estructura de bucle. La estructura de bucle comprende una puerta **3320** que evita que la estructura de bucle se enhebre a través de un nanoporo en una dirección inversa.

Con referencia a la **figura 34**, el procedimiento puede incluir polimerizar **3405** la pluralidad de sondas en un orden determinado por la hibridación de los restos de hibridación con el ácido nucleico monocatenario que se va a secuenciar **3410**. Con referencia a la **figura 35**, el procedimiento puede incluir escindir **3505** los grupos escindibles para proporcionar una hebra extendida que se va a secuenciar.

Con referencia a la **figura 36**, el procedimiento puede incluir enhebrar **3605** la hebra extendida **3610** a través de un nanoporo **3615**, donde las puertas evitan que la hebra extendida se enhebre a través del nanoporo en la dirección inversa **3620**. El procedimiento puede incluir detectar, con la ayuda del nanoporo, las estructuras de bucle de la hebra extendida en el orden determinado por la hibridación de los restos de hibridación con el ácido nucleico monocatenario que se va a secuenciar, secuenciando de este modo el ácido nucleico monocatenario que se va a secuenciar.

En algunos casos, la estructura de bucle comprende un segmento estrecho y la puerta es un polímero que comprende dos extremos, donde un primer extremo se fija a la estructura de bucle contigua al segmento estrecho y un segundo extremo no se fija a la estructura de bucle. La estructura de bucle puede ser capaz de enhebrarse a través de un nanoporo en una primera dirección donde la puerta se alinea contigua al segmento estrecho. En algunos casos, la estructura de bucle no se puede enhebrar a través del nanoporo en la dirección inversa donde la puerta no se alinea contigua al segmento estrecho.

En algunos casos, la puerta comprende nucleótidos. La puerta se puede emparejar con la estructura de bucle cuando la puerta no se alinea contigua al segmento estrecho.

El segmento estrecho puede comprender cualquier polímero o molécula que sea lo suficientemente estrecho para que el polímero pueda pasar a través de un nanoporo cuando una puerta se alinea con el segmento estrecho. En algunos casos, el segmento estrecho comprende nucleótidos abásicos (es decir, una cadena lateral de ácido nucleico que no tiene bases de nucleótidos unidas a la misma) o una cadena de carbonos.

En algunos casos, un electrodo se recarga entre periodos de detección. En general, la hebra extendida no se enhebra a través del nanoporo en la dirección inversa cuando el electrodo se recarga.

Procedimientos y aplicaciones distintos de la secuenciación

Los dispositivos y procedimientos descritos en el presente documento se pueden usar para medir determinadas propiedades de muestras de ácido nucleico y/o moléculas de ácido nucleico distintas de su secuencia de ácido nucleico (por ejemplo, la longitud de la secuencia o cualquier medida de la calidad de la muestra de ácido nucleico que incluye, pero sin limitarse a, el grado de reticulación de los ácidos nucleicos en la muestra). En algunos casos, puede ser deseable no determinar la secuencia de una molécula de ácido nucleico. Por ejemplo, se pueden identificar individuos humanos (u otros organismos tales como caballos y similares) determinando las longitudes de determinadas secuencias repetitivas en el genoma (por ejemplo, conocidas como microsatélites, repeticiones de secuencia simple (SSR) o repeticiones cortas en tándem (STR)). Es posible que se desee saber la longitud de una o más STR (por ejemplo, para identificar paternidad o autoría de un delito) sin conocer la secuencia de la STR y/o la secuencia de ADN encontrada antes (5') o después (3') de la STR (por ejemplo, para no identificar la raza de la persona, la probabilidad de contraer una enfermedad y similares).

El procedimiento puede identificar una o más STR presentes en un genoma. Se puede identificar cualquier número de STR (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 o más). La STR puede comprender un segmento repetido (por ejemplo, 'AGGTCT' de la secuencia SEQ. ID. N.º 1 - AGGTCT AGGTCT AGGTCT AGGTCT AGGTCT AGGTCT AGGTCT) que tiene cualquier número de bases de ácido nucleico (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 o más bases). La STR puede comprender cualquier número de segmentos repetidos, en general, repetidos consecutivamente (por ejemplo, repetidos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 o más veces).

El número de acontecimientos de incorporación de nucleótidos y/o la longitud de un ácido nucleico o segmento del mismo se puede determinar usando nucleótidos que tienen la misma marca unida a algunos, la mayoría o todos los nucleótidos marcados. La detección de una marca (precargada en un nanoporo antes de la liberación o dirigida al nanoporo después de la liberación del nucleótido marcado) indica que se ha producido un acontecimiento de incorporación de nucleótidos, pero, en este caso, no identifica qué nucleótido se ha incorporado (por ejemplo, no se determina la información de secuencia).

En algunos casos, todos los nucleótidos (por ejemplo, todos los nucleótidos adenina (A), citosina (C), guanina (G), timina (T) y/o uracilo (U)) tienen la misma marca acoplada al nucleótido. En algunos casos, sin embargo, esto puede no ser necesario. Al menos algunos de los nucleótidos pueden tener marcas que identifiquen el nucleótido (por ejemplo, de modo que se determine cierta información de secuencia). En algunos casos, aproximadamente un 5 %, aproximadamente un 10 %, aproximadamente un 15 %, aproximadamente un 20 %, aproximadamente un 25 %, aproximadamente un 30 %, aproximadamente un 40 % o aproximadamente un 50 % de los nucleótidos tienen marcas que identifican el nucleótido (por ejemplo, de modo que algunas posiciones de ácido nucleico se secuencien). Las posiciones secuenciadas de ácido nucleico se pueden distribuir aleatoriamente a lo largo de la cadena de ácido nucleico. En algunos casos, todos los nucleótidos de un único tipo de nucleótido tienen una marca de identificación (por ejemplo, de modo que todas las adeninas se secuencien, por ejemplo). En algunos casos, al

menos un 5 %, al menos un 10 %, al menos un 15 %, al menos un 20 %, al menos un 25 %, al menos un 30 %, al menos un 40 % o al menos un 50 % de los nucleótidos tienen marcas que identifican el nucleótido. En algunos casos, como máximo un 5 %, como máximo un 10 %, como máximo un 15 %, como máximo un 20 %, como máximo un 25 %, como máximo un 30 %, como máximo un 40 % o como máximo un 50 % de los nucleótidos tienen marcas que identifican el nucleótido. En algunos casos, todo el ácido nucleico o segmento del mismo es una repetición corta en tándem (STR).

En un aspecto, un procedimiento para determinar la longitud de un ácido nucleico o segmento del mismo con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor comprende proporcionar nucleótidos marcados en una cámara de reacción que comprende el nanoporo. Los nucleótidos tienen bases diferentes, tal como al menos dos bases diferentes, que contienen la misma marca acoplada a un nucleótido, siendo dicha marca detectable con la ayuda del nanoporo. El procedimiento incluye además llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de los nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de la muestra de ácido nucleico. El procedimiento incluye además detectar, con la ayuda del nanoporo, una marca asociada con el nucleótido marcado individual durante la incorporación del nucleótido marcado individual.

También se describe un procedimiento para determinar la longitud de un ácido nucleico o segmento del mismo con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor que comprende proporcionar nucleótidos marcados en una cámara de reacción que comprende el nanoporo. Un nucleótido marcado individual de los nucleótidos marcados puede contener una marca acoplada a un nucleótido, pudiendo dicha marca reducir la magnitud de la corriente que fluye a través del nanoporo con respecto a la corriente cuando la marca no está presente.

El procedimiento puede comprender además llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de los nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de la muestra de ácido nucleico y reducir la magnitud de la corriente que fluye a través del nanoporo. La magnitud de la corriente se puede reducir en cualquier cantidad adecuada, que incluye aproximadamente un 5 %, aproximadamente un 10 %, aproximadamente un 15 %, aproximadamente un 20 %, aproximadamente un 25 %, aproximadamente un 30 %, aproximadamente un 40 %, aproximadamente un 50 %, aproximadamente un 60 %, aproximadamente un 70 %, aproximadamente un 80 %, aproximadamente un 90 %, aproximadamente un 95 % o aproximadamente un 99 %. En algunos casos, la magnitud de la corriente se reduce en al menos un 5 %, al menos un 10 %, al menos un 15 %, al menos un 20 %, al menos un 25 %, al menos un 30 %, al menos un 40 %, al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 %, al menos un 95 % o al menos un 99 %. En algunos casos, la magnitud de la corriente se reduce como máximo un 5 %, como máximo un 10 %, como máximo un 15 %, como máximo un 20 %, como máximo un 25 %, como máximo un 30 %, como máximo un 40 %, como máximo un 50 %, como máximo un 60 %, como máximo un 70 %, como máximo un 80 %, como máximo un 90 %, como máximo un 95 % o como máximo un 99 %.

El procedimiento puede comprender además detectar, con la ayuda del nanoporo, periodos de tiempo entre la incorporación de los nucleótidos marcados individuales (por ejemplo, periodos 605 en la figura 6). Los periodos de tiempo entre la incorporación de los nucleótidos marcados individuales pueden tener una gran magnitud de corriente. En algunos casos, la magnitud de la corriente que fluye a través del nanoporo entre los acontecimientos de incorporación de nucleótidos es (por ejemplo, vuelve a) aproximadamente un 50 %, aproximadamente un 60 %, aproximadamente un 70 %, aproximadamente un 80 %, aproximadamente un 90 %, aproximadamente un 95 % o aproximadamente un 99 % de la corriente máxima (por ejemplo, cuando no hay una marca presente). En algunos casos, la magnitud de la corriente que fluye a través del nanoporo entre los acontecimientos de incorporación de nucleótidos es al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 %, al menos un 90 %, al menos un 95 % o al menos un 99 % de la corriente máxima.

En algunos casos, una sección de ácido nucleico se secuenciar antes (5') o después (3') de la STR para identificar qué STR tiene determinada su longitud en un nanoporo (por ejemplo, en un contexto multiplexado donde se dirige una pluralidad de cebadores hacia una pluralidad de STR). En algunos casos, aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 o más ácidos nucleicos se secuencian antes (5') o después (3') de la STR.

Concordancia de patrones

Esta divulgación también proporciona lectores electrónicos para hacer coincidir patrones de señales detectadas por un dispositivo de nanoporos con señales conocidas (o de referencia). El dispositivo de nanoporos puede incluir un nanoporo en una membrana, como se describe en otra parte del presente documento. Las señales conocidas se pueden mantener en una localización de memoria, tal como una base de datos remota o una localización de memoria localizada en un chip que comprende el dispositivo de nanoporos. Un lector electrónico puede hacer coincidir patrones con la ayuda de un algoritmo de coincidencia de patrones, que se puede implementar con la ayuda de un procesador informático del lector electrónico. El lector electrónico se puede localizar en el chip.

La coincidencia de patrones se puede implementar en tiempo real, tal como mientras el dispositivo de nanoporos recopila datos. Como alternativa, la coincidencia de patrones se puede implementar al recoger primero los datos y posteriormente procesar los datos para que coincidan con los patrones.

En algunos casos, el lector contiene una lista de una o más secuencias de ácido nucleico que son de interés para un usuario (también "lista blanca" en el presente documento) y una lista (o una pluralidad de listas) de una o más de otras secuencias de ácido nucleico que no son de interés para un usuario (también "lista negra" en el presente documento). Durante la detección de ácidos nucleicos, que incluye los acontecimientos de incorporación de ácidos nucleicos, el lector puede detectar y registrar secuencias de ácido nucleico que están en la lista blanca, y no detectar o registrar secuencias de ácido nucleico que están en la lista negra.

EJEMPLOS

Ejemplo 1: conducción no faradaica

La **figura 37** muestra que la conducción no faradaica puede desacoplar el nanoporo de la modulación. El eje vertical de la figura es la corriente medida en picoamperios (pA) que varía de -30 a 30. El eje horizontal es el tiempo medido en segundos (s) que varía de 0 a 2. La forma de onda tiene un ciclo de trabajo de un 40 %. Los puntos de datos **3705** son la corriente medida con un electrodo de trabajo de platino esponjoso en presencia de KCl 150 mM, pH 7,5 con tampón HEPES 20 mM y SrCl_2 3 mM por encima y por debajo de la bicapa. Hay 240 nM de polimerasa pegajosa y un sándwich de 5GS con una absorbancia (densidad óptica) de 0,0464. El lípido es un 75 % de fosfatidiletanolamina (PE) y un 25 % de fosfatidilcolina (PC). La tensión simulada **3710** a través del electrodo de trabajo y el contraelectrodo (sedimento de AgCl) se muestra multiplicado por 100 para ajustarse a la curva. El potencial electroquímico simulado a través del complejo nanoporo-polimerasa **3715** se muestra multiplicado por 100 para ajustarse a la curva. La corriente **3720** se simula usando un modelo de programa de simulación con énfasis de circuitos integrados (SPICE).

Ejemplo 2: captura de marcas

La **figura 38** muestra dos marcas capturadas en un nanoporo en un sistema de corriente alterna (CA). El eje vertical de la figura es la corriente medida en picoamperios (pA) que varía de 0 a 25. El eje horizontal es el tiempo medido en segundos (s) que varía de aproximadamente 769 a 780. Una primera marca **3805** se captura a aproximadamente 10 pA. Una segunda marca **3810** se captura a aproximadamente 5 pA. La corriente de canal abierto **3815** es de aproximadamente 18 pA. Pocos puntos de datos se ven en la corriente de canal abierto debido a la captura rápida de las marcas. La forma de onda es de 0 a 150 mV a 10 Hz y un ciclo de trabajo de un 40 %. La solución contiene KCl 150 mM.

Ejemplo 3: secuenciación de marcas

La **figura 39** muestra un ejemplo de un complejo ternario **3900** formado entre una molécula de ADN molde que se va a secuenciar **3905**, una fusión de un nanoporo de hemolisina **3910** y una ADN polimerasa **3915**, y un nucleótido marcado **3920**. La polimerasa **3915** se une al nanoporo **3910** con un conector de proteínas **3925**. La construcción nanoporo/polimerasa se forma de modo que solo uno de los siete monómeros polipeptídicos del nanoporo tenga una polimerasa unida. Parte del nucleótido marcado se enhebra en el nanoporo **3920** y afecta a la corriente que pasa a través del nanoporo.

La **figura 40** muestra la corriente que fluye a través del nanoporo en presencia de ADN molde que se va a secuenciar, pero sin nucleótidos marcados. La solución en contacto con el nanoporo tiene KCl 150 mM, SrCl_2 0,7 mM, MgCl_2 3 mM y tampón HEPES 20 mM a pH 7,5, a una tensión aplicada de 100 mV. La corriente permanece cerca de 18 picoamps (pA) con algunas excepciones 4005. Las excepciones pueden ser ruido electrónico y pueden ser solo un punto de datos en el eje horizontal de tiempo. El ruido electrónico se puede mitigar usando un algoritmo que distingue el ruido de la señal, tal como, por ejemplo, un algoritmo de procesamiento de señal adaptativo.

La **figura 41**, la **figura 42** y la **figura 43** muestran que diferentes marcas proporcionan diferentes niveles de corriente. En todos los ejemplos, la solución en contacto con el nanoporo tiene KCl 150 mM, SrCl_2 0,7 mM, MgCl_2 3 mM y tampón HEPES 20 mM a pH 7,5, a una tensión aplicada de 100 mV. La **figura 41** muestra una guanina (G) **4105** que se distingue de una timina (T) **4110**. Las marcas son dT6P-T6-dSp8-T16-C3 (para T) que tiene un nivel de corriente de aproximadamente 8 a 10 pA y dG6P-Cy3-30T-C6 (para G) y que tiene un nivel de corriente de aproximadamente 4 o 5 pA. La **figura 42** muestra una guanina (G) **4205** que se distingue de una adenina (A) **4210**. Las marcas son dA6P-T4-(Sp18)-T22-C3 (para A) que tiene un nivel de corriente de aproximadamente 6 a 7 pA y dG6P-Cy3-30T-C6 (para G) y que tiene un nivel de corriente de aproximadamente 4 o 5 pA. La **figura 43** muestra una guanina (G) **4305** que se distingue de una citosina (C) **4310**. Las marcas son dC6P-T4-(Sp18)-T22-C3 (para C) que tiene un nivel de corriente de aproximadamente 1 a 3 pA y dG6P-Cy3-30T-C6 (para G) y que tiene un nivel de corriente de aproximadamente 4 o 5 pA.

La **figura 44**, la **figura 45**, la **figura 46** y la **figura 47** muestran ejemplos de secuenciación usando nucleótidos marcados. La molécula de ADN que se va a secuenciar es monocatenaria y tiene la secuencia AGTCAGTC (SEQ ID N.º: 36) y se estabiliza mediante dos estructuras de horquilla flanqueantes. En todos los ejemplos, la solución en contacto con el nanoporo tiene KCl 150 mM, SrCl₂ 0,7 mM, MgCl₂ 3 mM y tampón HEPES 20 mM a pH 7,5, a una tensión aplicada de 100 mV. En la solución se incluyen cuatro marcas correspondientes a guanina (dG6P-Cy3-30T-C6), adenina (dA6P-T4-(Sp18)-T22-C3), citosina (dC6P-T4-(Sp18)-T22-C3) y timina (dT6P-T6-dSp8-T16-C3). La **figura 44** muestra un ejemplo donde se identifican cuatro nucleótidos marcados consecutivos (es decir, C **4405**, A **4410**, G **4415** y T **4420**) correspondientes a la secuencia GTCA en SEQ ID N.º: 36. La marca puede entrar y salir del nanoporo varias veces antes de incorporarse a la hebra en crecimiento (por ejemplo, para cada acontecimiento de incorporación, el nivel de corriente puede cambiar varias veces entre la corriente de canal abierto y el nivel de corriente reducido que distingue la marca).

La duración de la reducción de corriente puede ser diferente entre ensayos, por cualquier motivo, incluyendo, pero sin limitarse a, que el número de veces que la marca entra y sale del nanoporo sea diferente y/o que la marca se mantenga durante poco tiempo por la polimerasa pero no se incorpore completamente en una hebra de ácido nucleico en crecimiento. En algunos casos, la duración de la reducción de corriente es aproximadamente consecuente entre ensayos (por ejemplo, varía en no más de aproximadamente un 200 %, un 100 %, un 50% o un 20 %). En algunos casos, la enzima, la forma de onda de la tensión aplicada, la concentración de iones divalentes y/o monovalentes, la temperatura y/o el pH se eligen de modo que la duración de la reducción de corriente sea aproximadamente consecuente entre ensayos. La **figura 45** muestra la identificación de la misma secuencia GTCA en SEQ ID N.º: 36 que se muestra en la **figura 44** (es decir, en este caso, los nucleótidos marcados identificados son C **4505**, A **4510**, G **4515** y T **4520**). En algunos casos, la corriente permanece reducida durante un periodo prolongado de tiempo (por ejemplo, aproximadamente 2 segundos, como se muestra en **4520**).

La **figura 46** muestra la identificación de cinco nucleótidos marcados consecutivos (es decir, T **4605**, C **4610**, A **4615**, G **4620**, T **4625**) correspondientes a la secuencia AGTCA en SEQ ID N.º: 36. La **figura 47** muestra la identificación de cinco nucleótidos marcados consecutivos que se identifican (es decir, T **4705**, C **4710**, A **4715**, G **4720**, T **4725**, C **4730**) correspondientes a la secuencia AGTCAG en SEQ ID N.º: 36.

Aunque se muestran y describen modos de realización preferentes de la presente invención en el presente documento, será obvio para los expertos en la técnica que dichos modos de realización se proporcionan solo a modo de ejemplo. No se pretende que la invención esté limitada por los ejemplos específicos proporcionados dentro de la memoria descriptiva. Aunque la invención se ha descrito con referencia a la memoria descriptiva mencionada anteriormente, las descripciones e ilustraciones de los modos de realización en el presente documento no se deben interpretar en un sentido limitante. A los expertos en la técnica se les ocurrirán ahora numerosas variaciones, cambios y sustituciones sin apartarse de la invención. Además, se debe entender que todos los aspectos de la invención no se limitan a las representaciones, configuraciones o proporciones relativas específicas expuestas en el presente documento que dependen de una variedad de condiciones y variables. Se debe entender que se pueden emplear diversas alternativas a los modos de realización de la invención descritos en el presente documento al poner en práctica la invención. Por lo tanto, se contempla que la invención también cubrirá dichas alternativas, modificaciones, variaciones o equivalentes. Se pretende que las siguientes reivindicaciones definan el alcance de la invención y que los procedimientos dentro del alcance de estas reivindicaciones estén cubiertos de este modo.

REIVINDICACIONES

1. Un procedimiento para secuenciar una muestra de ácido nucleico con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor, comprendiendo el procedimiento:

(a) proporcionar nucleótidos marcados en una cámara de reacción que comprende dicho nanoporo, en el que un nucleótido marcado individual de dichos nucleótidos marcados contiene una marca acoplada a un nucleótido, siendo dicha marca detectable con la ayuda de dicho nanoporo;

(b) llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de dichos nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de dicha muestra de ácido nucleico; y

(c) detectar, con la ayuda de dicho nanoporo, una marca asociada con dicho nucleótido marcado individual durante la incorporación de dicho nucleótido marcado individual, en el que dicha marca se detecta con la ayuda de dicho nanoporo mientras dicho nucleótido marcado está asociado con dicha polimerasa.

2. El procedimiento de la reivindicación 1,

i) en el que dicha marca se detecta una pluralidad de veces mientras está asociada con dicha polimerasa, en particular, en el que un electrodo se recarga entre periodos de detección de marcas; o

ii) en el que, en (c), un nucleótido marcado incorporado se distingue de un nucleótido marcado no incorporado en base al periodo de tiempo en el que dicho nanoporo detecta dicho nucleótido marcado, en particular, en el que la proporción entre el tiempo en el que el nanoporo detecta un nucleótido marcado incorporado y el tiempo en el que el nanoporo detecta un nucleótido marcado no incorporado es al menos 1,5, o, en particular, en el que la proporción entre el tiempo en el que el nanoporo detecta un nucleótido marcado incorporado y el tiempo en el que el nanoporo detecta un nucleótido marcado no incorporado es al menos 10; o

iii) en el que dicho nucleótido marcado se asocia con dicha polimerasa durante un periodo de tiempo promedio de al menos aproximadamente 1 milisegundo; o

iv) en el que dicho nucleótido marcado pasa a través de dicho nanoporo en menos de 1 milisegundo cuando dicho nucleótido marcado no está asociado con dicha polimerasa; o

v) en el que dicho nanoporo puede distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos un 95 %; o

vi) en el que dicho nanoporo puede distinguir entre un nucleótido marcado incorporado y un nucleótido marcado no incorporado con una exactitud de al menos un 99 %; o

vii) en el que, en (b), una marca asociada con dicho nucleótido marcado individual se enhebra a través de al menos una parte de dicho nanoporo, en particular, en el que, en (c), dicha marca se detecta mientras se enhebra a través de al menos la parte de dicho nanoporo.

3. El procedimiento de la reivindicación 1,

en el que, durante la incorporación de dicho nucleótido marcado individual, se diferencia, con la ayuda de dicho nanoporo, una marca asociada con dicho nucleótido marcado individual de una o más marcas asociadas con uno o más nucleótidos marcados individuales no incorporados, en particular, en el que dicho nucleótido marcado individual incorporado en (b) se diferencia de los nucleótidos marcados individuales no incorporados en base a los periodos de tiempo en los que dicho nucleótido marcado individual incorporado en (b) y dichos nucleótidos marcados individuales no incorporados se detectan con la ayuda de dicho nanoporo.

4. Un procedimiento para secuenciar un ácido nucleico, comprendiendo el procedimiento: aplicar una forma de onda de corriente alterna (CA) a un circuito próximo a un nanoporo y un electrodo sensor, en el que una marca asociada con un nucleótido marcado que se incorpora a una hebra de ácido nucleico en crecimiento complementaria a una hebra de ácido nucleico molde se detecta cuando la forma de onda tiene una primera polaridad y el electrodo se recarga cuando la forma de onda tiene una segunda polaridad,

en particular,

i) en el que la forma de onda de CA no agota dicho electrodo durante un periodo de tiempo de al menos 1 hora; o

ii) en el que la identidad de la marca se determina mediante la relación entre la corriente medida y la tensión aplicada por la forma de onda de CA a diversas tensiones; o

5 iii) en el que la metilación de una base de la hebra de ácido nucleico molde se determina mediante la detección de la marca durante un periodo de tiempo mayor cuando la base está metilada que cuando la base no está metilada.

5. El procedimiento de la reivindicación 4,

10 en el que la marca asociada con dicho nucleótido marcado individual se detecta una o más veces, en el que dicha marca se detecta mientras dicha marca está unida a dicho nucleótido marcado individual incorporado en dicha hebra;

en particular,

15 i) en el que la forma de onda de CA no agota dicho electrodo durante un periodo de tiempo de al menos 1 hora; o

20 ii) en el que la identidad de la marca se determina mediante la relación entre una corriente medida y las diversas tensiones aplicadas por la forma de onda de CA; o

25 iii) en el que al menos algunos de dichos nucleótidos marcados incluyen bases metiladas, en especial, en el que una marca asociada con un nucleótido marcado y que tiene una base que está metilada se detecta durante un periodo de tiempo mayor en comparación con una marca asociada con un nucleótido marcado que no tiene una base que está metilada.

6. Un procedimiento para determinar la longitud de un ácido nucleico o un segmento del mismo con la ayuda de un nanoporo en una membrana contigua a un electrodo sensor, comprendiendo el procedimiento:

30 (a) proporcionar nucleótidos marcados en una cámara de reacción que comprende dicho nanoporo, en el que los nucleótidos que tienen al menos dos bases diferentes contienen la misma marca acoplada a un nucleótido, siendo dicha marca detectable con la ayuda de dicho nanoporo;

35 (b) llevar a cabo una reacción de polimerización con la ayuda de una polimerasa, incorporando de este modo un nucleótido marcado individual de dichos nucleótidos marcados en una hebra en crecimiento complementaria a una molécula de ácido nucleico monocatenario de dicha muestra de ácido nucleico; y

40 (c) detectar, con la ayuda de dicho nanoporo, una marca asociada con dicho nucleótido marcado individual durante la incorporación de dicho nucleótido marcado individual,

en el que la longitud se correlaciona con el número de nucleótidos marcados incorporados,

en particular,

45 i) en el que todos los nucleótidos tienen la misma marca acoplada al nucleótido; o

ii) en el que al menos algunos de los nucleótidos tienen marcas que identifican el nucleótido; o

50 iii) en el que el ácido nucleico o segmento del mismo es una repetición corta en tándem (STR).

7. El procedimiento de la reivindicación 1,

55 en el que dicha detección comprende (i) proporcionar una tensión aplicada a través de dicho nanoporo y (ii) medir una corriente con dicho electrodo sensor a dicha tensión aplicada; y comprendiendo el procedimiento (d) calibrar dicha tensión aplicada.

8. El procedimiento de la reivindicación 7,

60 i) en el que dicha calibración comprende (i) medir una pluralidad de tensiones de escape para dicha molécula marcadora, (ii) calcular una diferencia entre las tensiones de escape medidas y un punto de referencia, y (iii) desviar la tensión aplicada por la diferencia calculada; o

65 ii) en el que se estima una distribución de las tensiones de escape esperadas frente al tiempo, en particular, en el que el punto de referencia es la media o mediana de las tensiones de escape medidas; o

iii) en el que (a)-(d) se realizan en una pluralidad de nanoporos direccionables independientemente, cada uno contiguo a un electrodo sensor; o

iv) en el que la tensión aplicada disminuye con el tiempo; o

v) que comprende además (e) calibrar variaciones en una ganancia de corriente y/o variaciones en una compensación de corriente; o

vi) en el que dicha marca se detecta una pluralidad de veces mientras está asociada con dicha polimerasa, en particular, en el que un electrodo se recarga entre periodos de detección de marca.

9. El procedimiento de la reivindicación 1,

en el que el procedimiento comprende además eliminar una secuencia de ácido nucleico repetitiva de la muestra de ácido nucleico para proporcionar una molécula de ácido nucleico monocatenario para la secuenciación;

en particular,

i) en el que la secuencia de ácido nucleico repetitiva se elimina mediante hibridación con una secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva, en especial, en el que la secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva se inmoviliza sobre un soporte sólido; o

ii) en el que la secuencia de ácido nucleico complementaria a la secuencia de ácido nucleico repetitiva comprende ADN Cot-1.

10. El procedimiento de la reivindicación 1,

en el que la polimerasa se une al nanoporo mediante un conector.

11. El procedimiento de la reivindicación 1,

en el que la marca comprende (i) una primera cadena polimérica que comprende un primer segmento y un segundo segmento, en la que el segundo segmento es más estrecho que el primer segmento; y (ii) una segunda cadena polimérica que comprende dos extremos, en la que un primer extremo se fija a la primera cadena polimérica contigua al segundo segmento y un segundo extremo no se fija a la primera cadena polimérica, en la que la molécula marcadora se puede enhebrar a través de un nanoporo en una primera dirección cuando la segunda cadena polimérica se alinea contigua al segundo segmento.

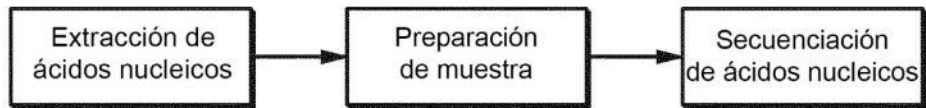


FIG. 1

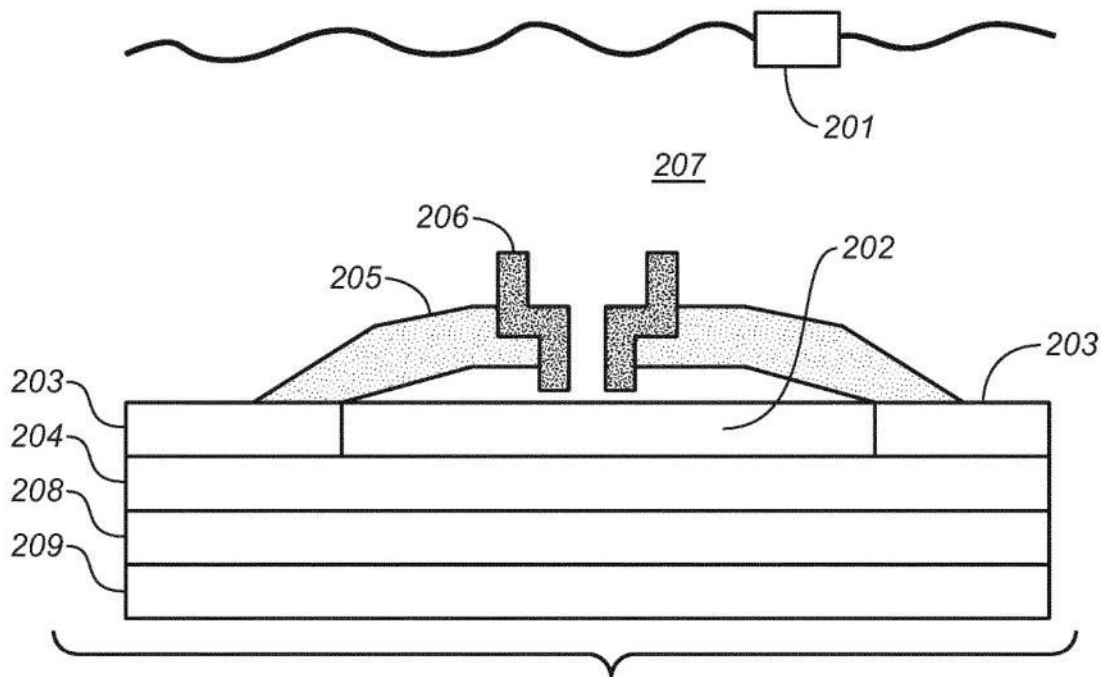


FIG. 2A

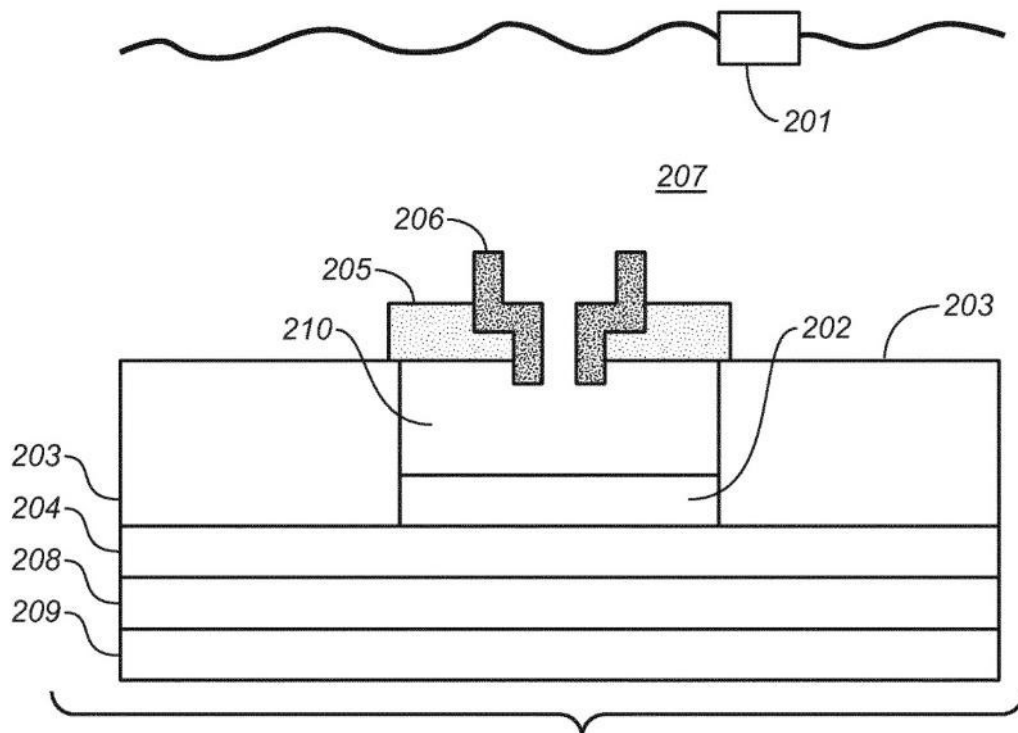


FIG. 2B

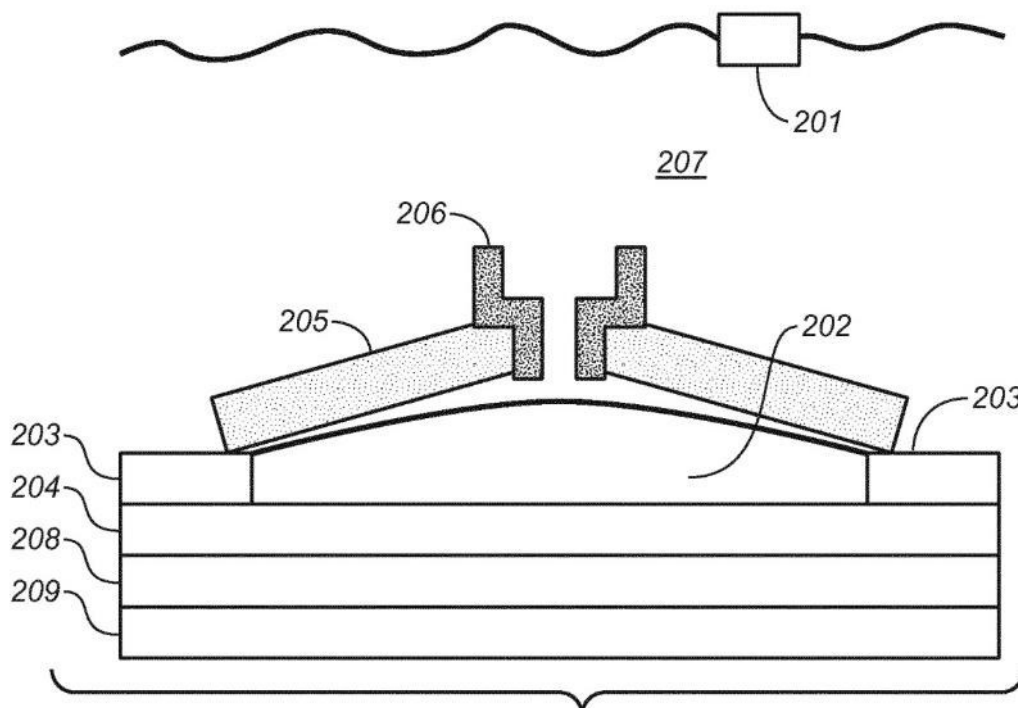


FIG. 2C

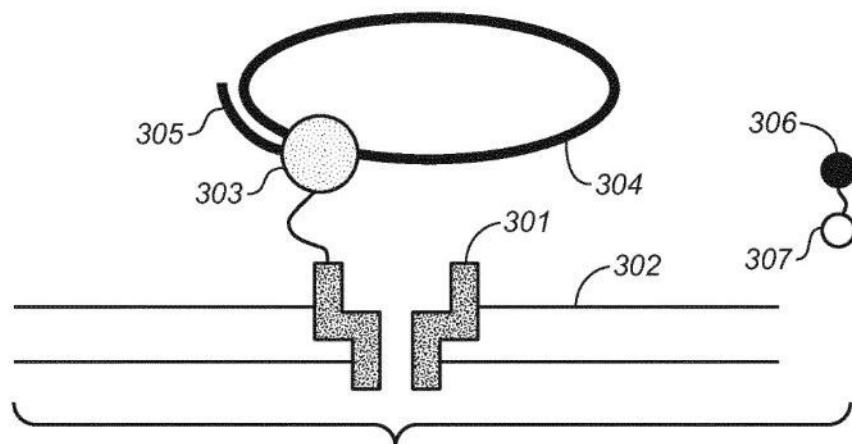


FIG. 3

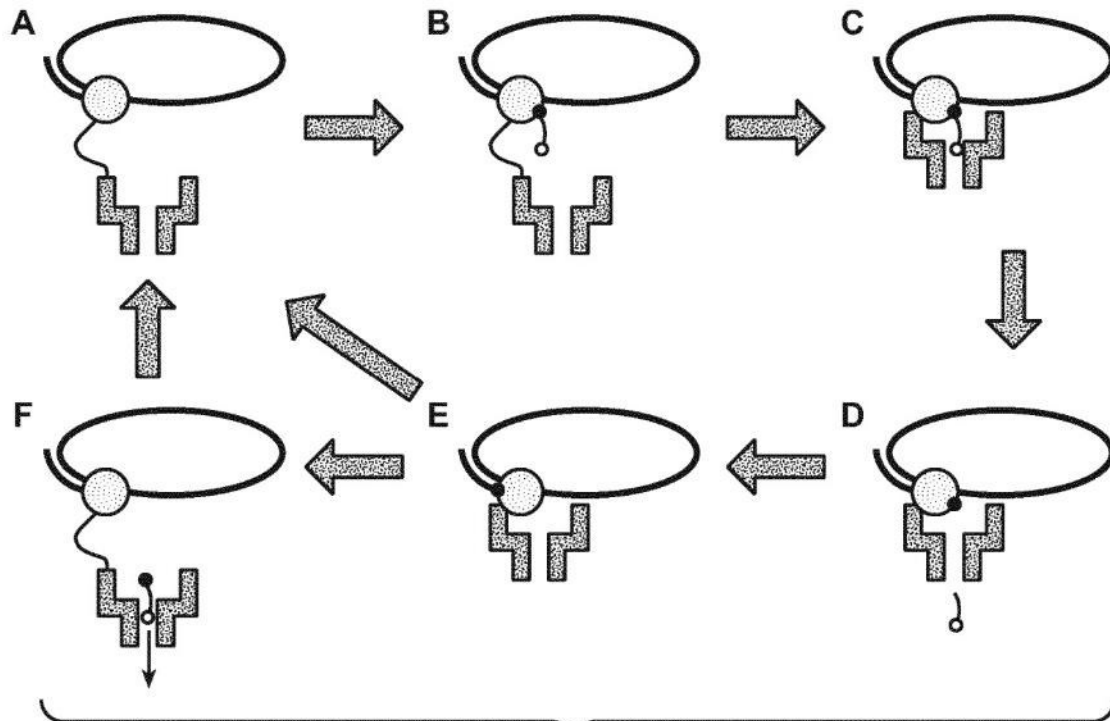


FIG. 4

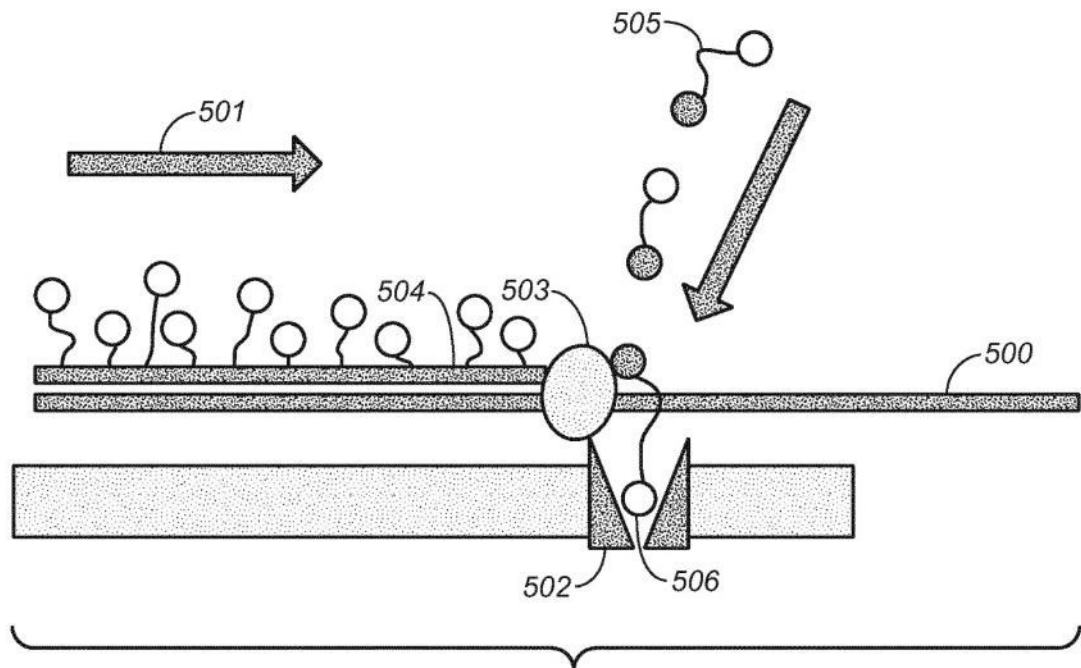


FIG. 5

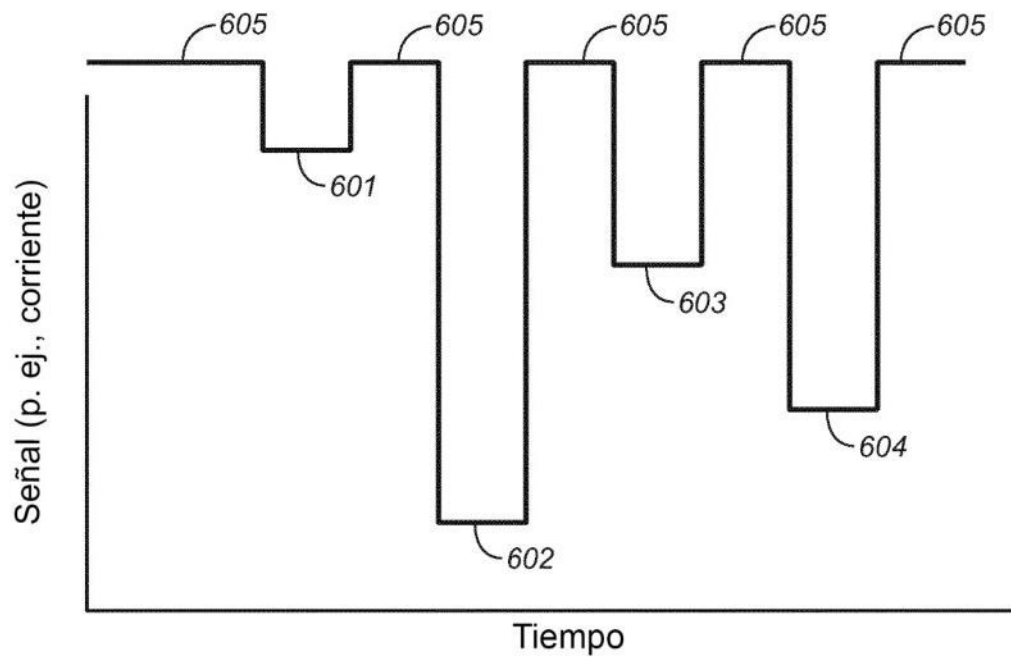


FIG. 6

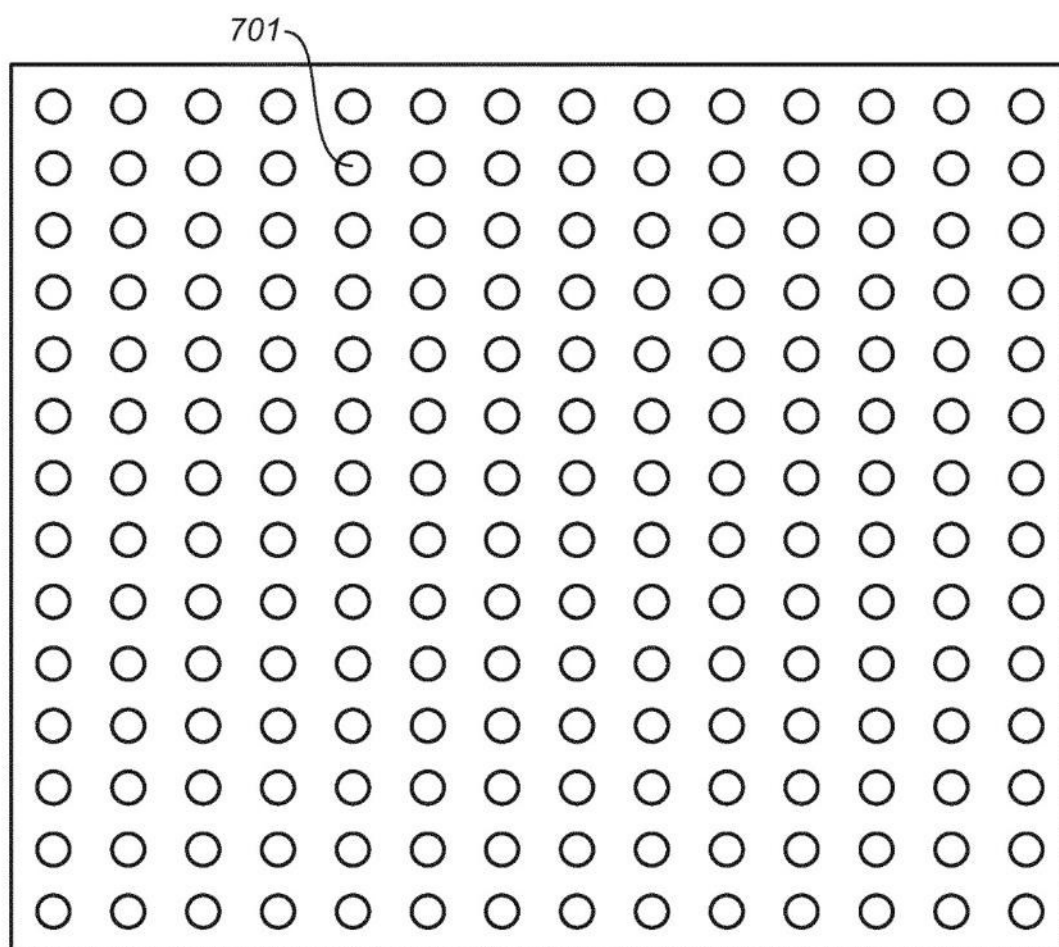


FIG. 7

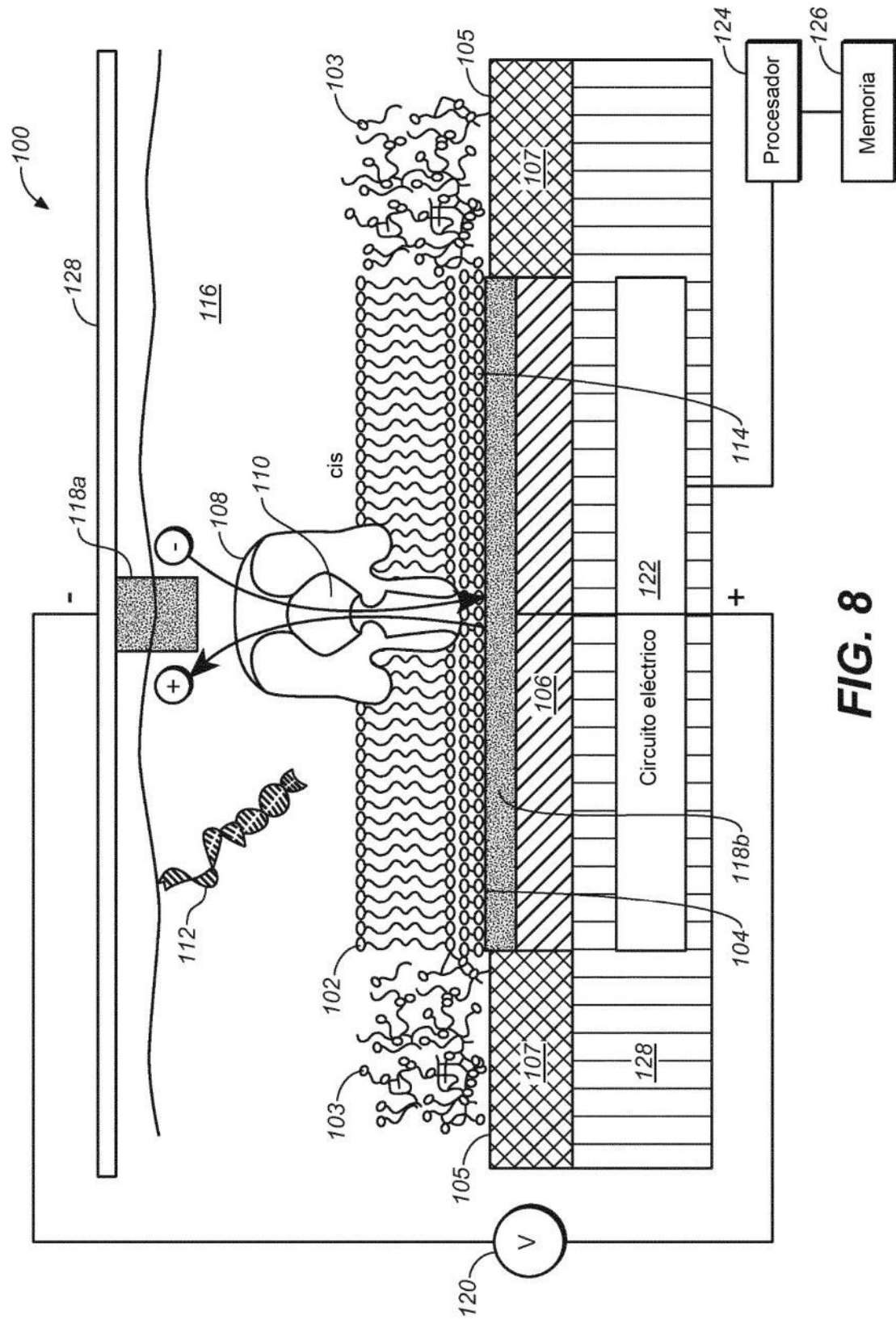


FIG. 8

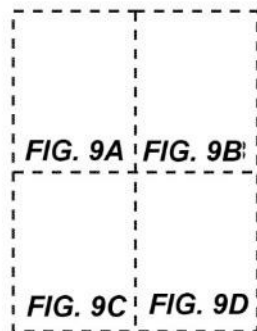


FIG. 9

FIG. 9A

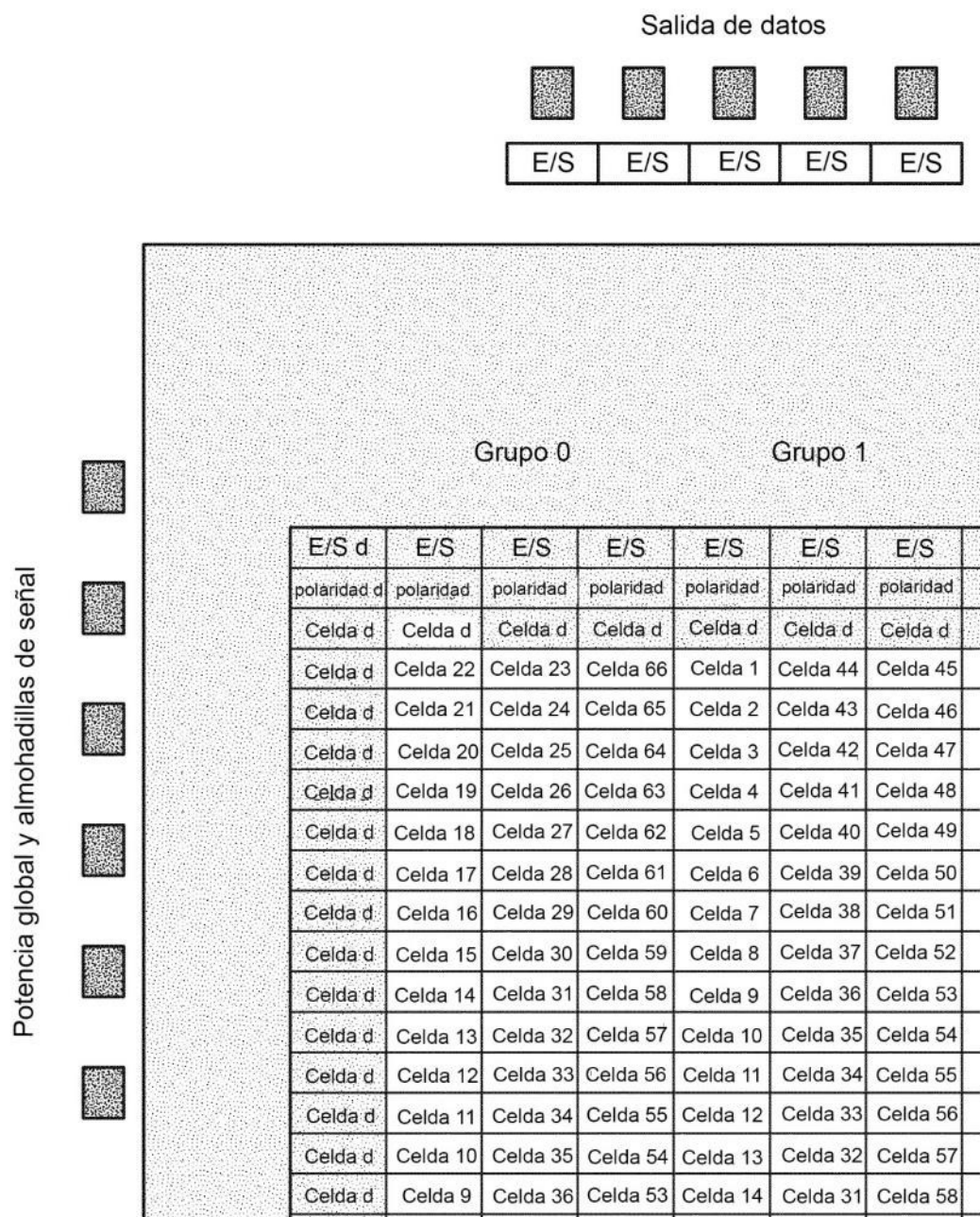


FIG. 9B

Salida de datos



E/S	E/S	E/S	E/S	E/S	E/S
-----	-----	-----	-----	-----	-----

Grupo 2

Grupo 3

E/S	E/S	E/S	E/S	E/S	E/S	E/S d
polaridad	polaridad	polaridad	polaridad	polaridad	polaridad	polaridad d
Celda d	Celda d	Celda d	Celda d	Celda d	Celda d	Celda d
Celda 22	Celda 23	Celda 66	Celda 1	Celda 44	Celda 45	Celda d
Celda 21	Celda 24	Celda 65	Celda 2	Celda 43	Celda 46	Celda d
Celda 20	Celda 25	Celda 64	Celda 3	Celda 42	Celda 47	Celda d
Celda 19	Celda 26	Celda 63	Celda 4	Celda 41	Celda 48	Celda d
Celda 18	Celda 27	Celda 62	Celda 5	Celda 40	Celda 49	Celda d
Celda 17	Celda 28	Celda 61	Celda 6	Celda 39	Celda 50	Celda d
Celda 16	Celda 29	Celda 60	Celda 7	Celda 38	Celda 51	Celda d
Celda 15	Celda 30	Celda 59	Celda 8	Celda 37	Celda 52	Celda d
Celda 14	Celda 31	Celda 58	Celda 9	Celda 36	Celda 53	Celda d
Celda 13	Celda 32	Celda 57	Celda 10	Celda 35	Celda 54	Celda d
Celda 12	Celda 33	Celda 56	Celda 11	Celda 34	Celda 55	Celda d
Celda 11	Celda 34	Celda 55	Celda 12	Celda 33	Celda 56	Celda d
Celda 10	Celda 35	Celda 54	Celda 13	Celda 32	Celda 57	Celda d
Celda 9	Celda 36	Celda 53	Celda 14	Celda 31	Celda 58	Celda d

Potencia global y almohadillas de señal

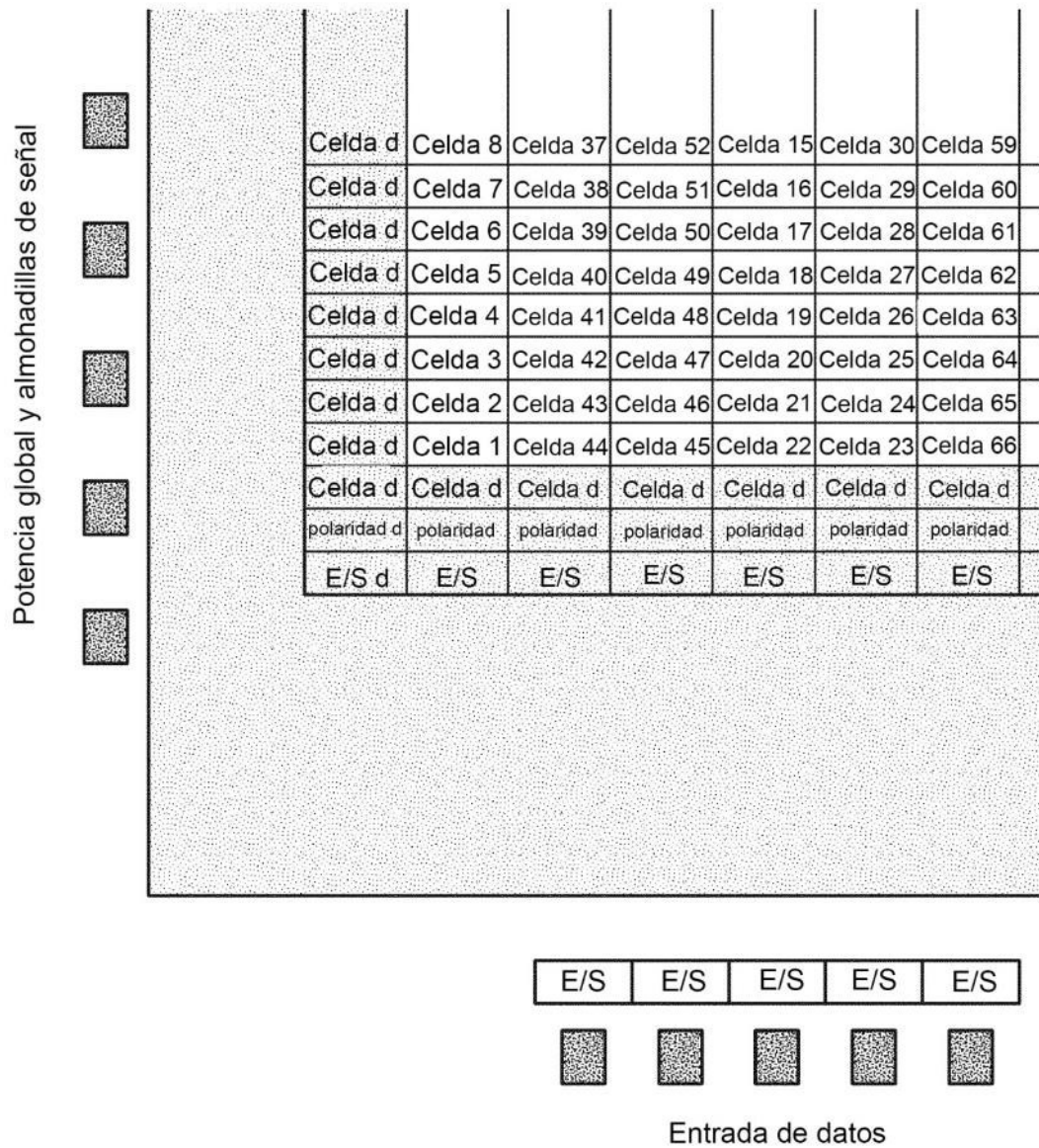


FIG. 9C

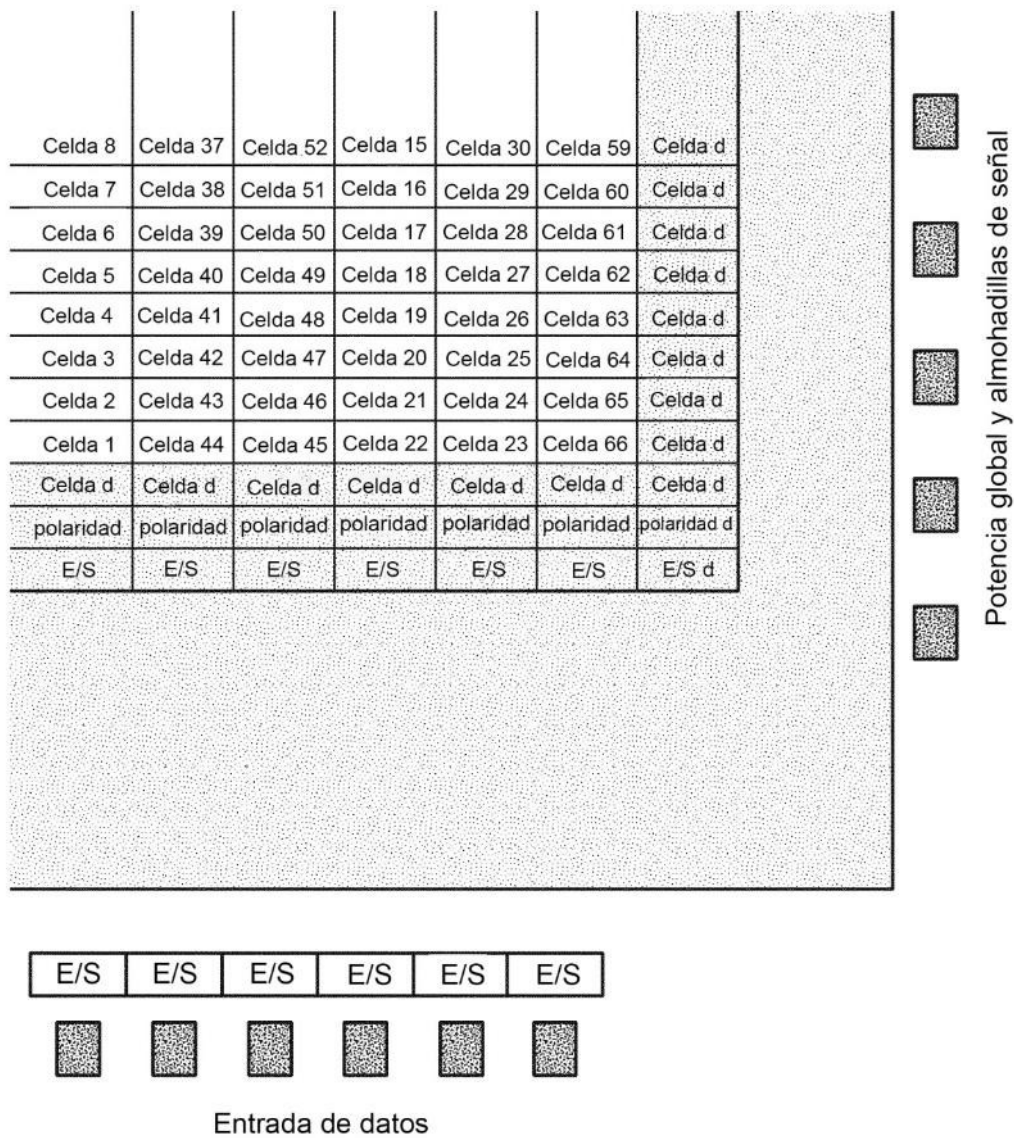


FIG. 9D

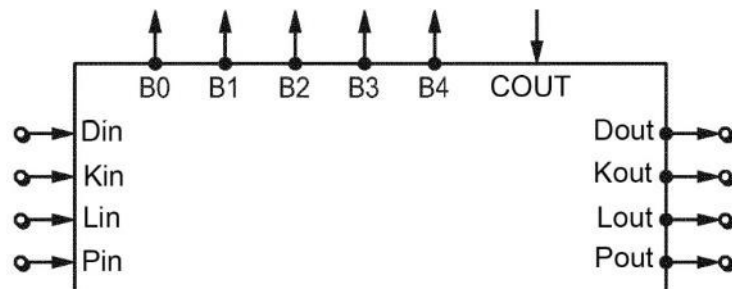
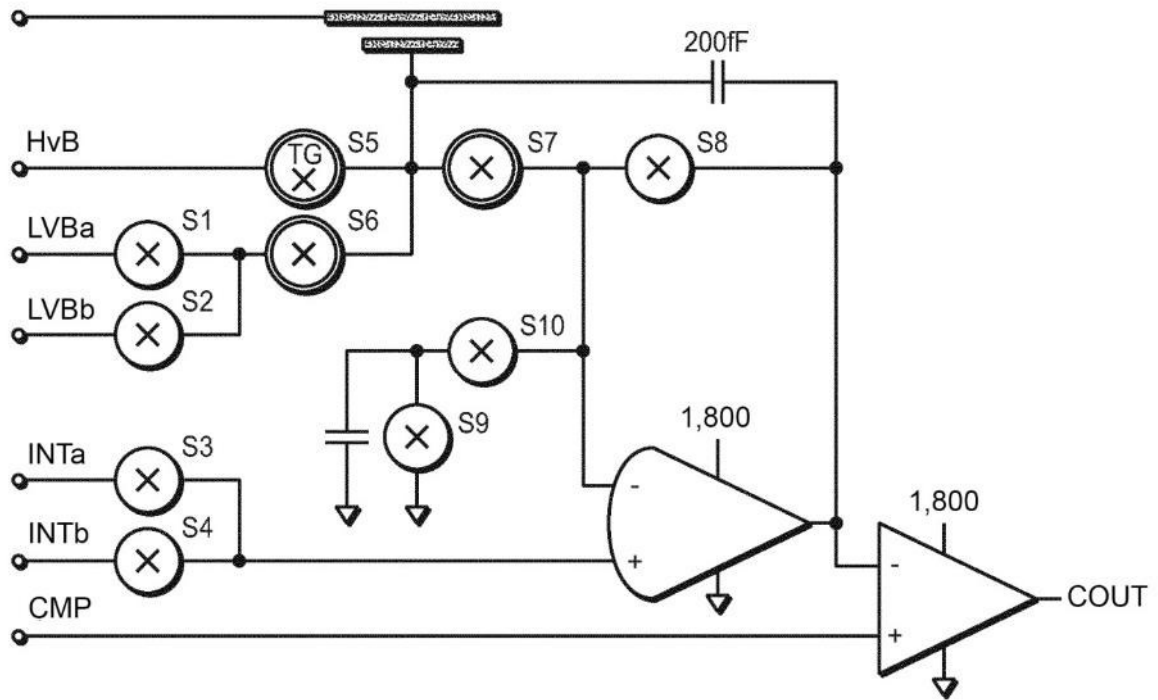


FIG. 10

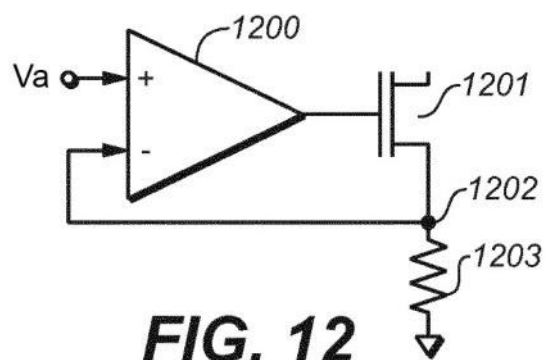


FIG. 12

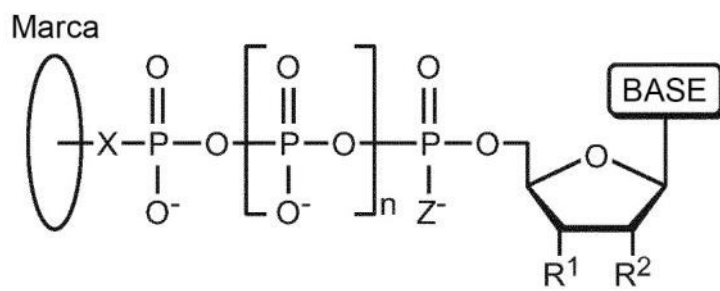
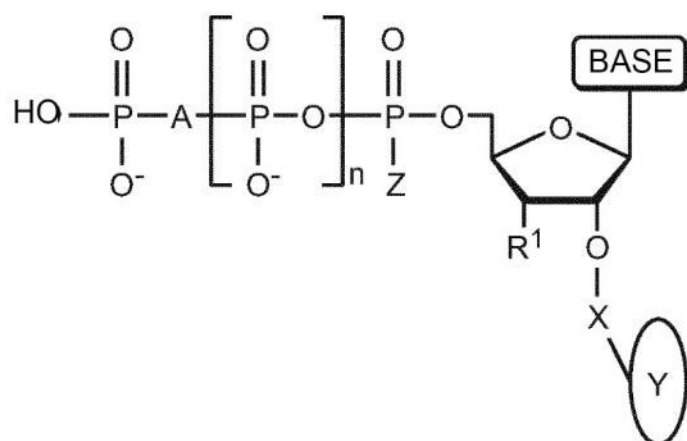
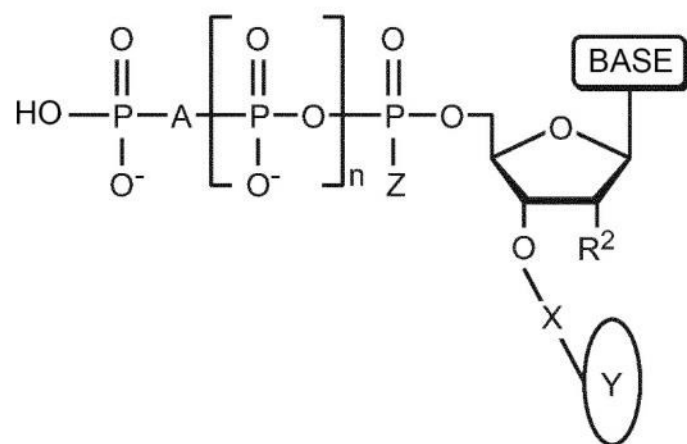
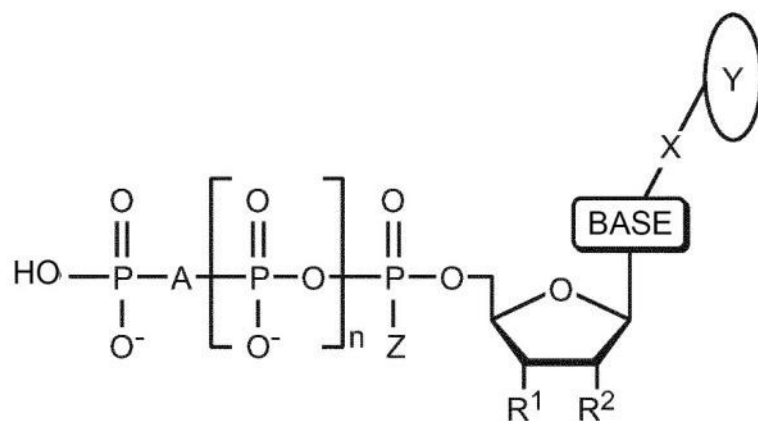


FIG. 13

**FIG. 14**

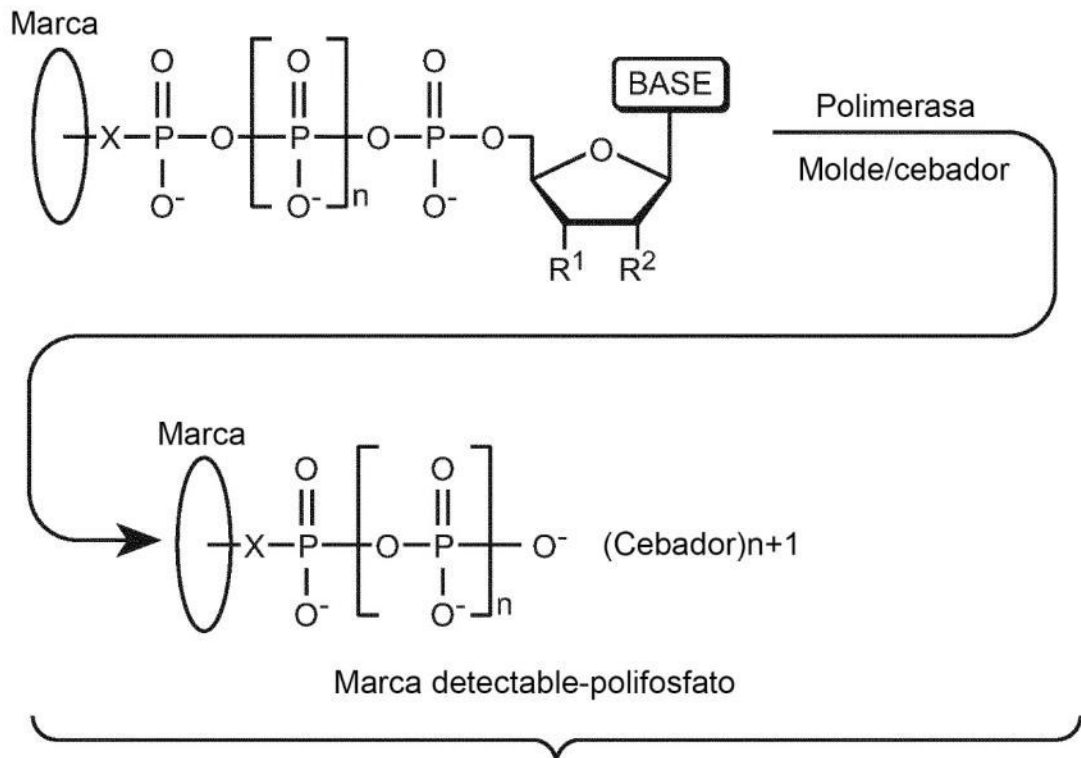


FIG. 15

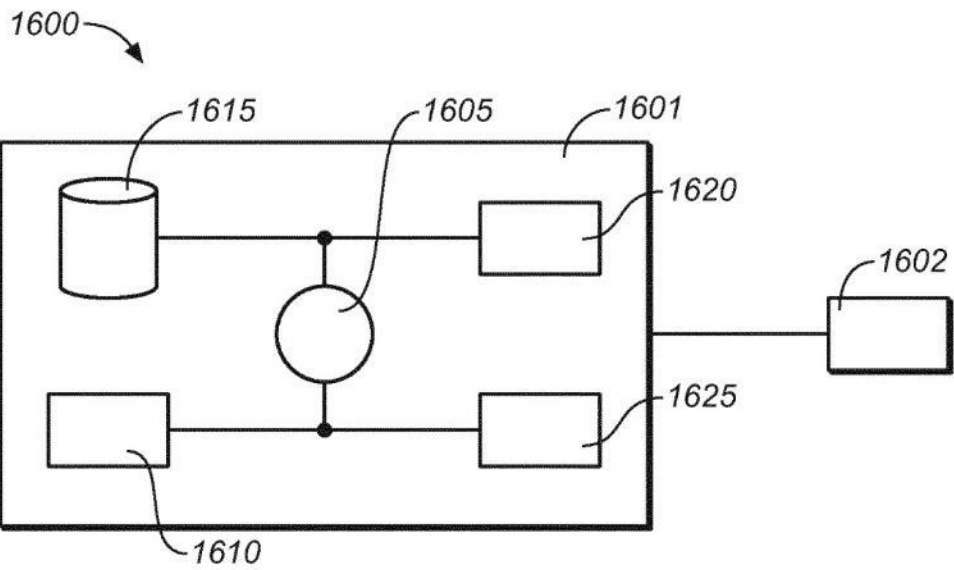


FIG. 16

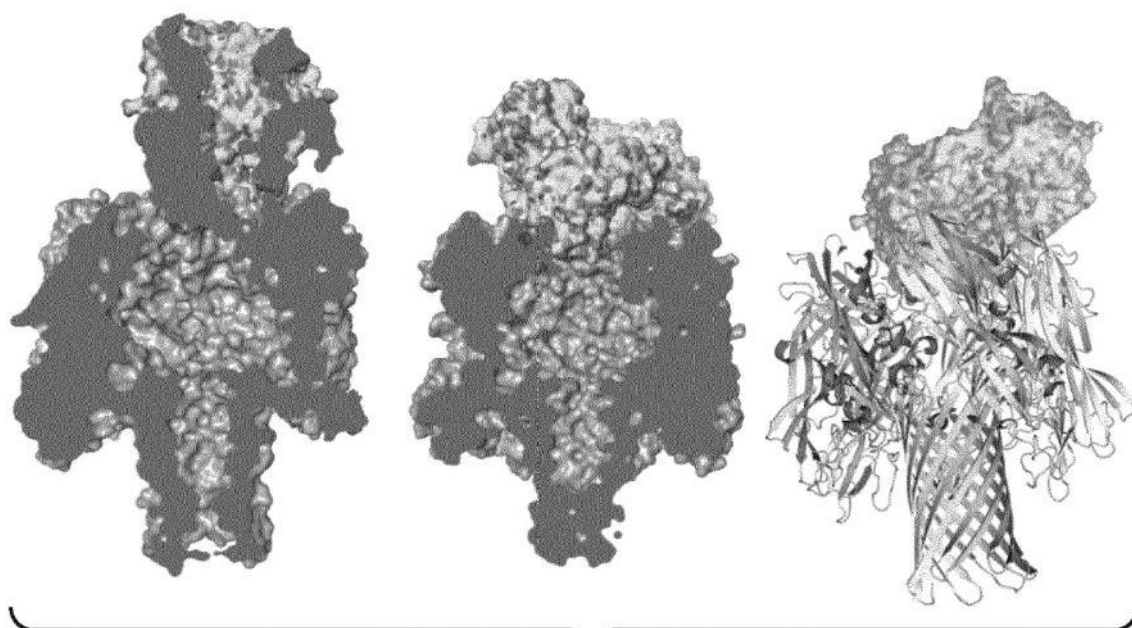


FIG. 17

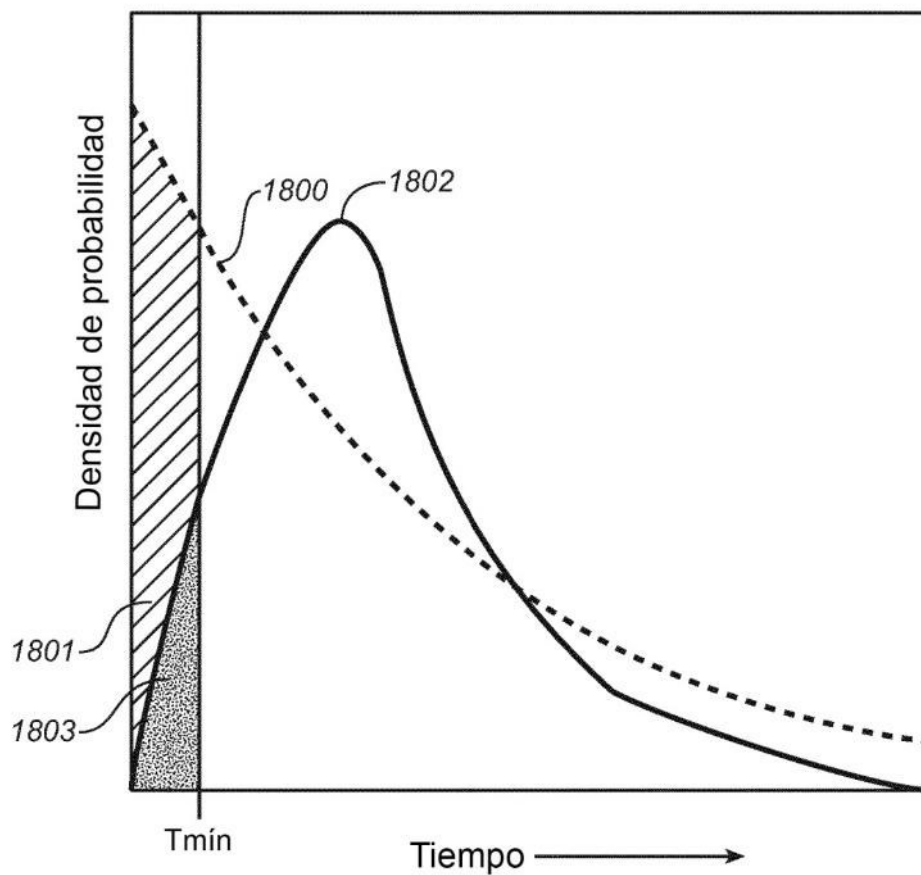


FIG. 18

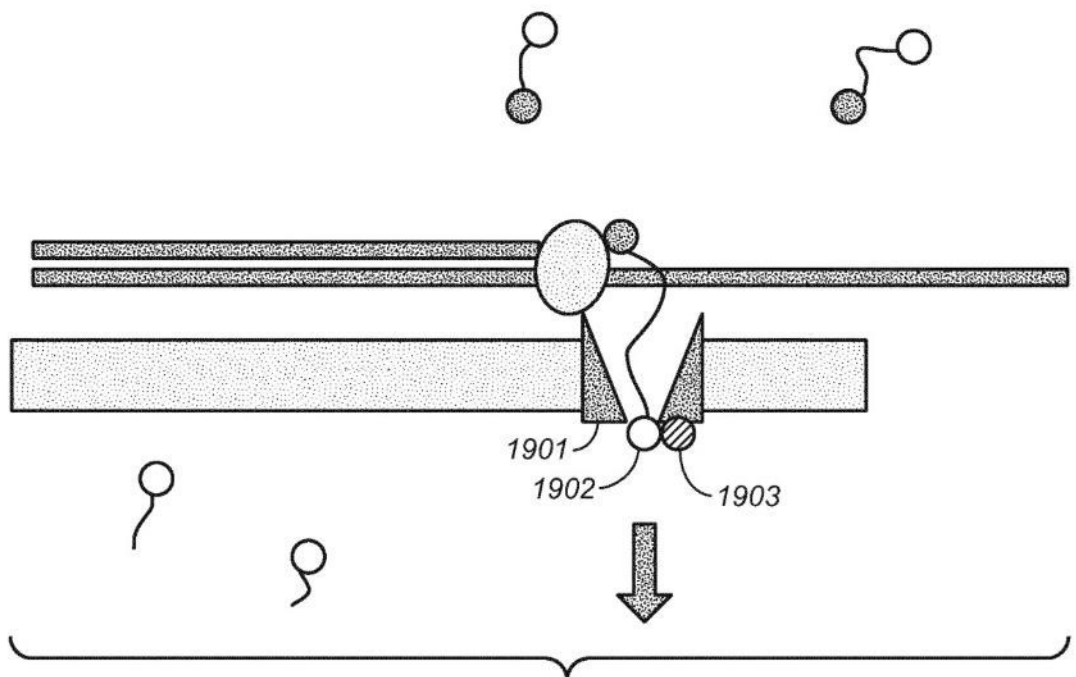


FIG. 19

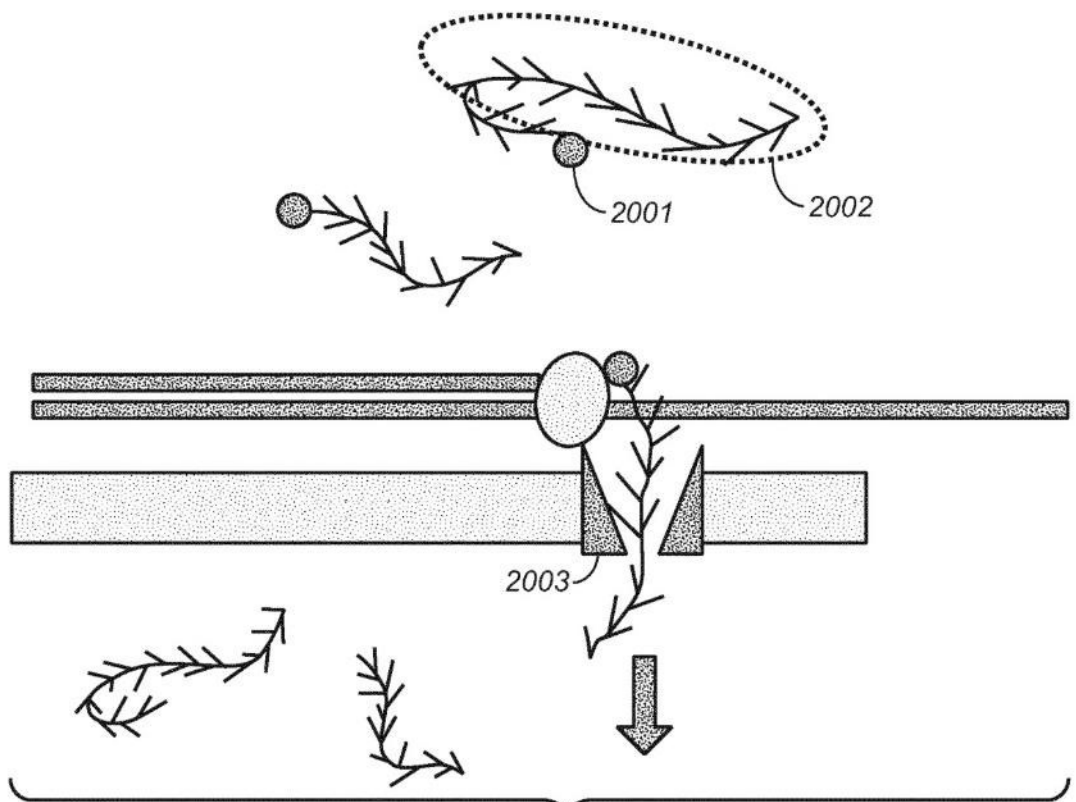


FIG. 20

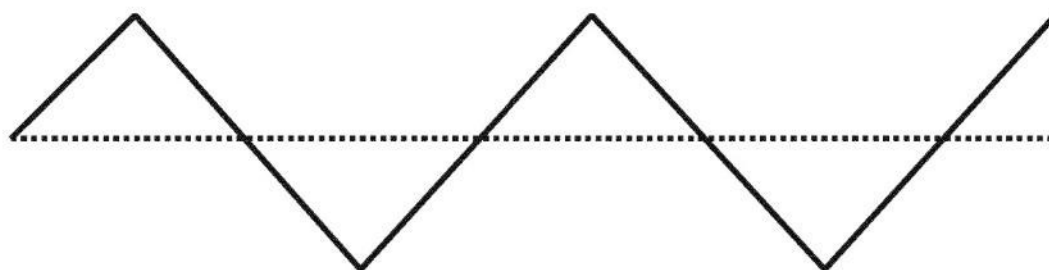


FIG. 21A

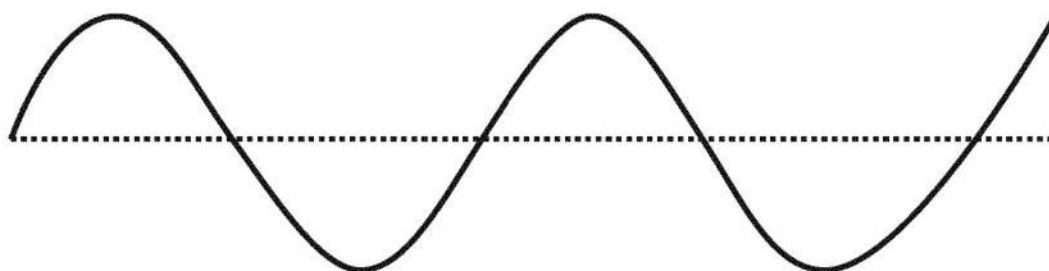


FIG. 21B

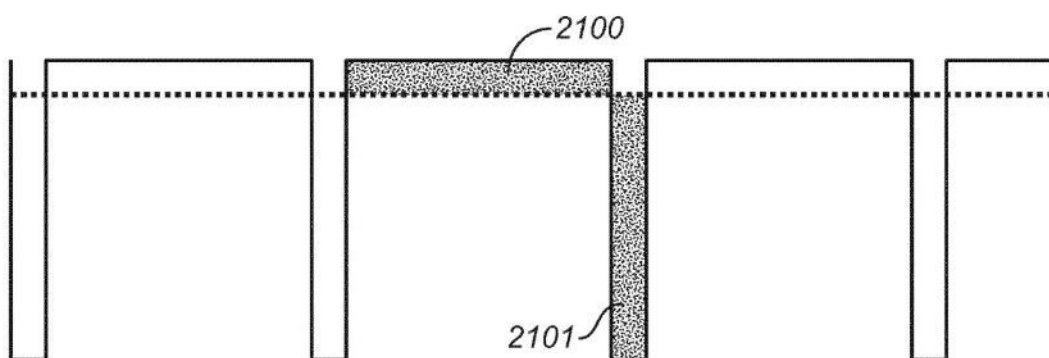


FIG. 21C

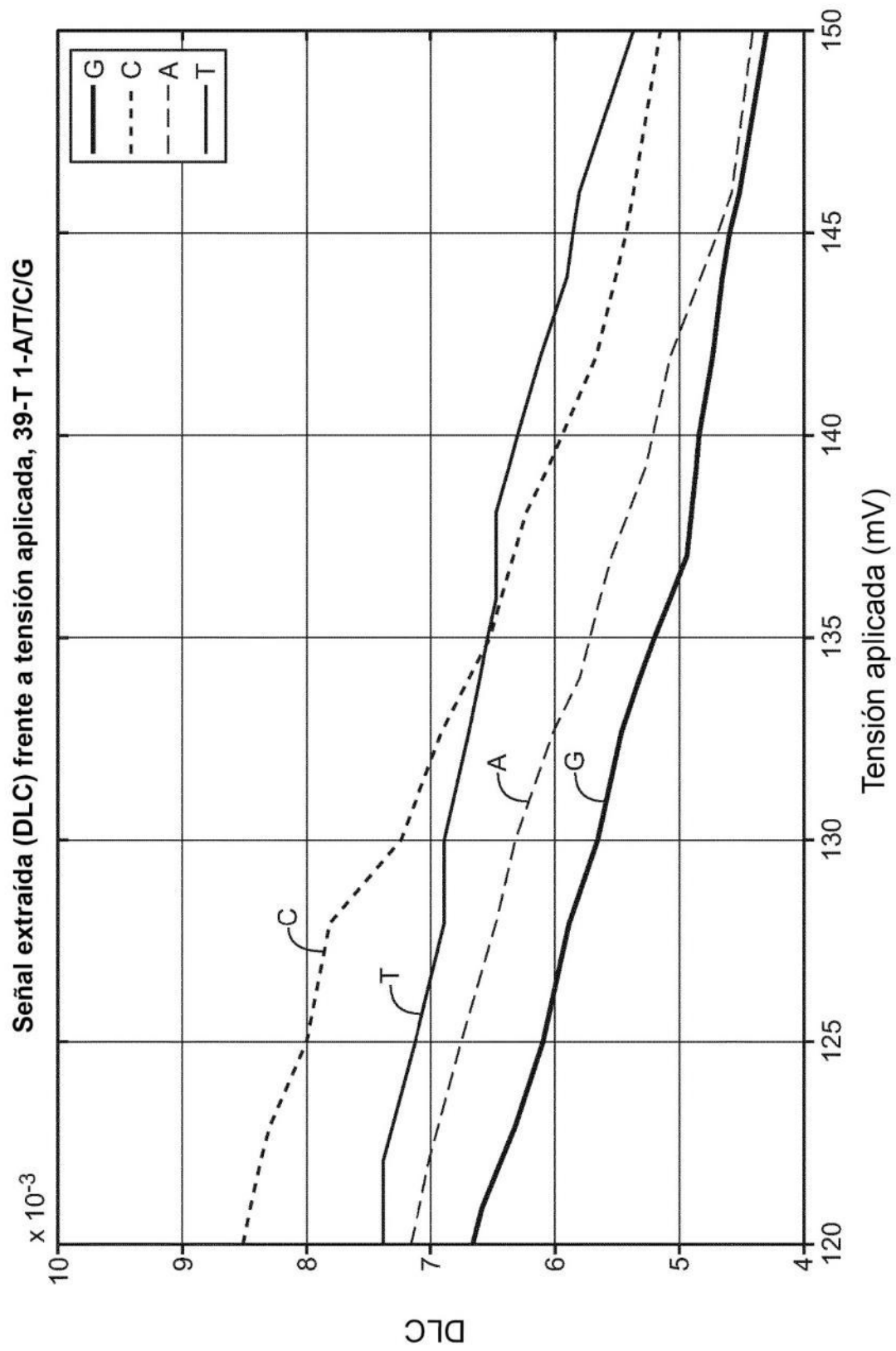


FIG. 22

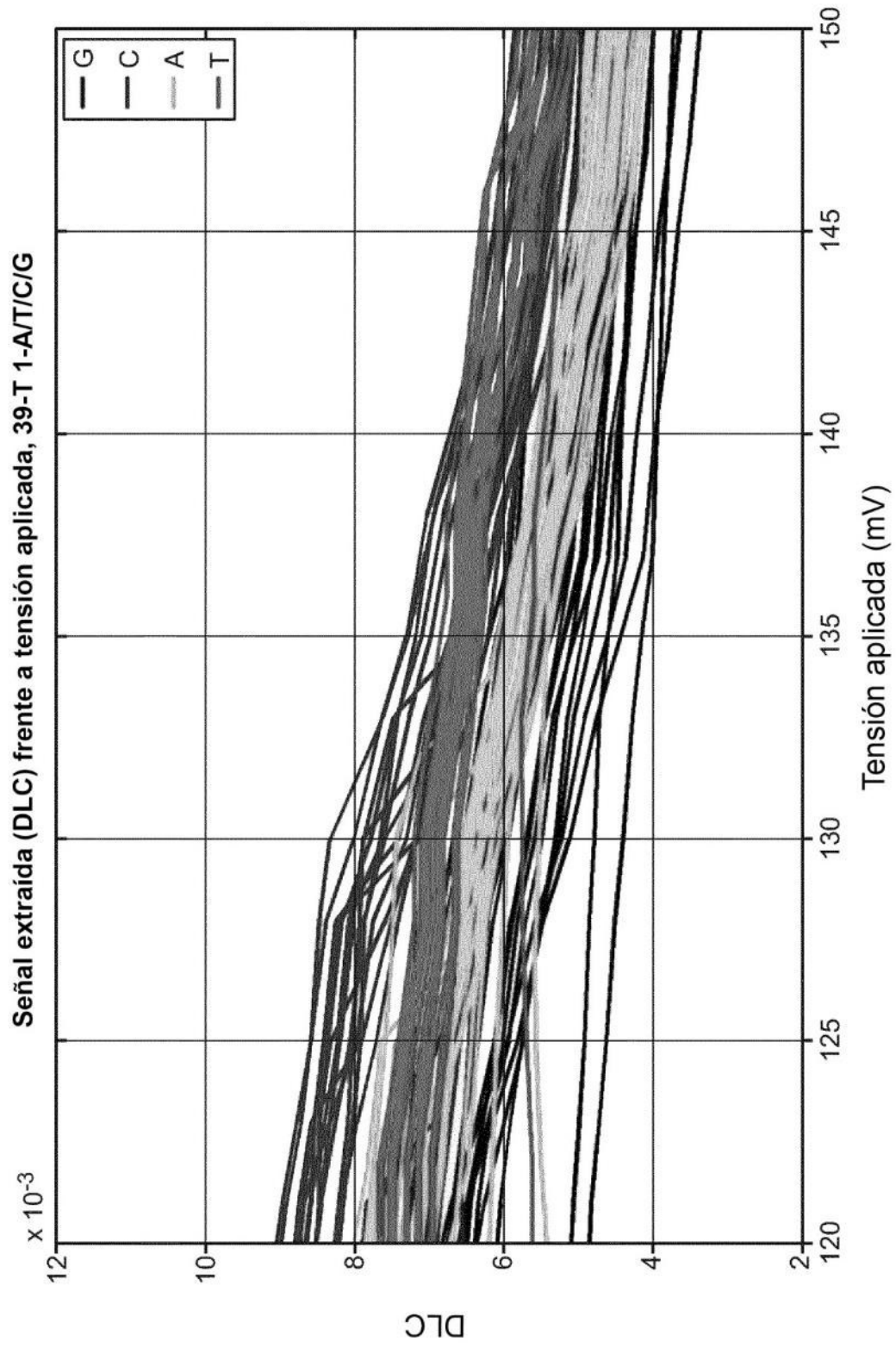


FIG. 23

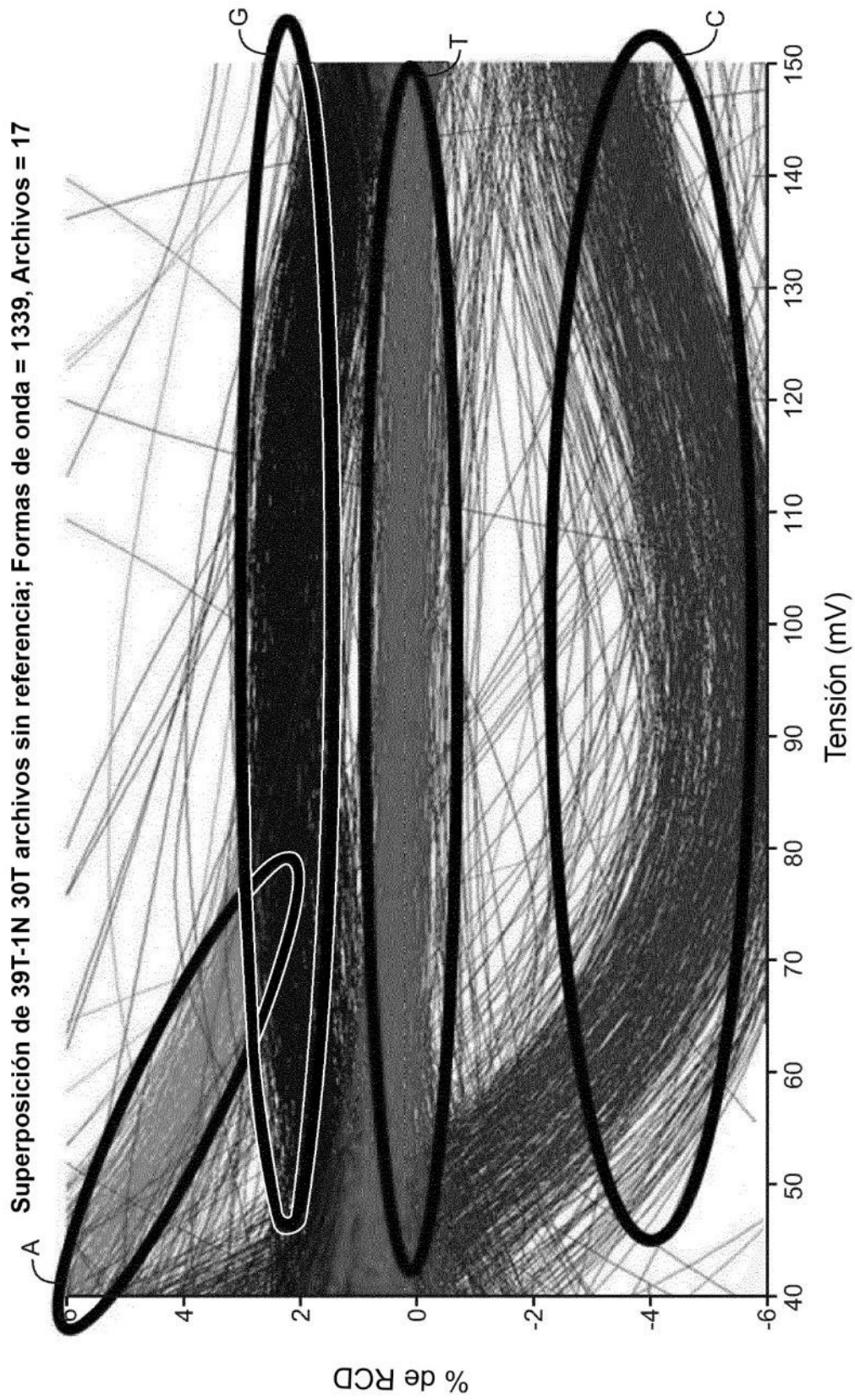


FIG. 24

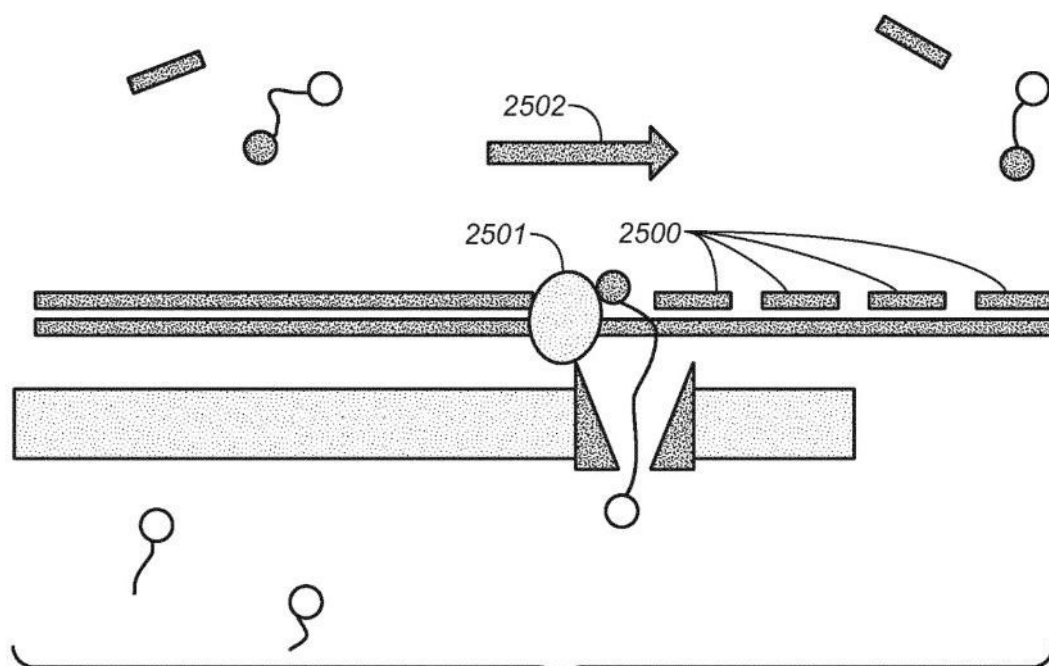


FIG. 25

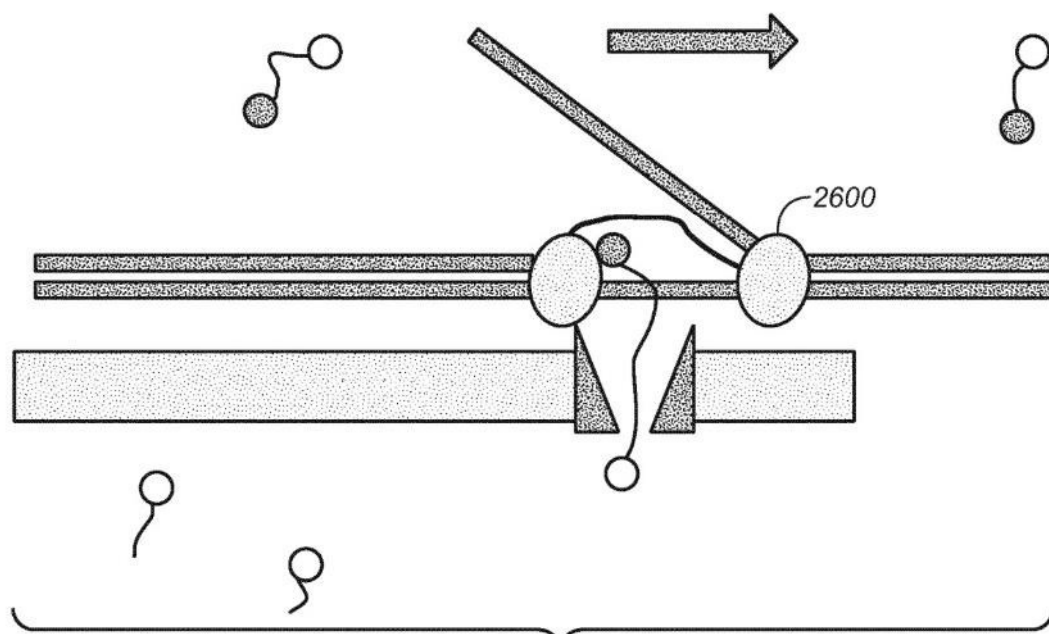
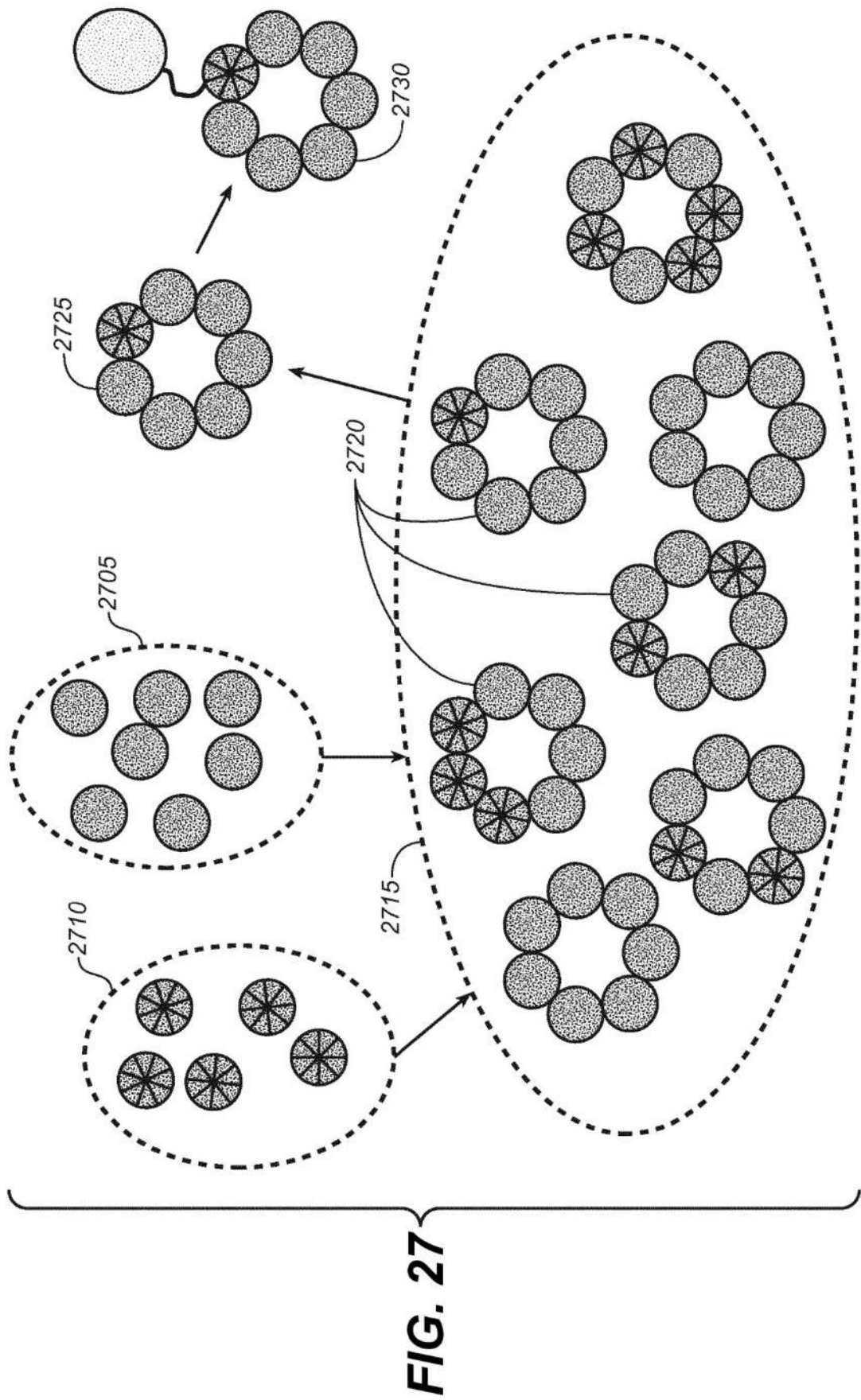


FIG. 26



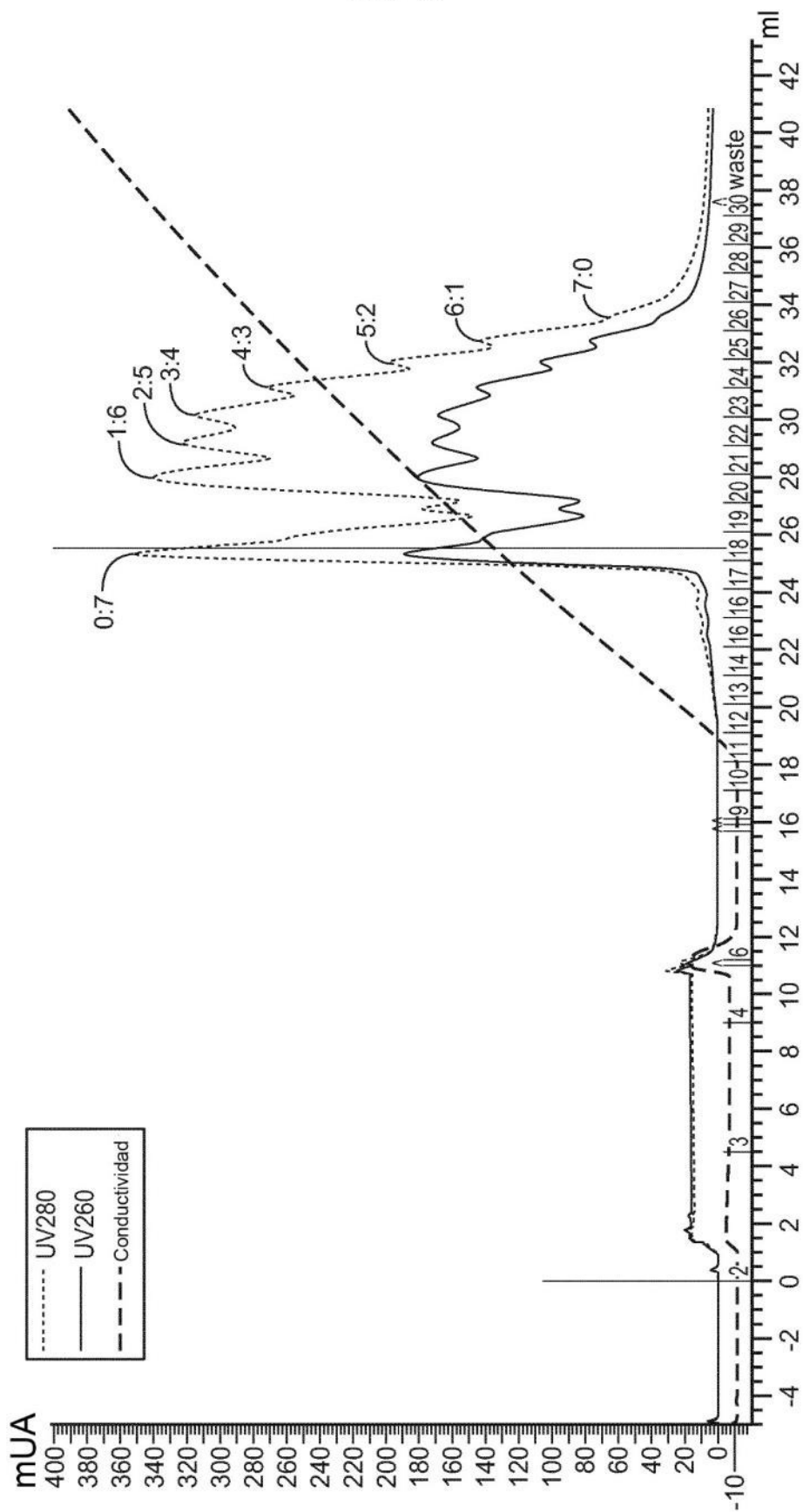


FIG. 28

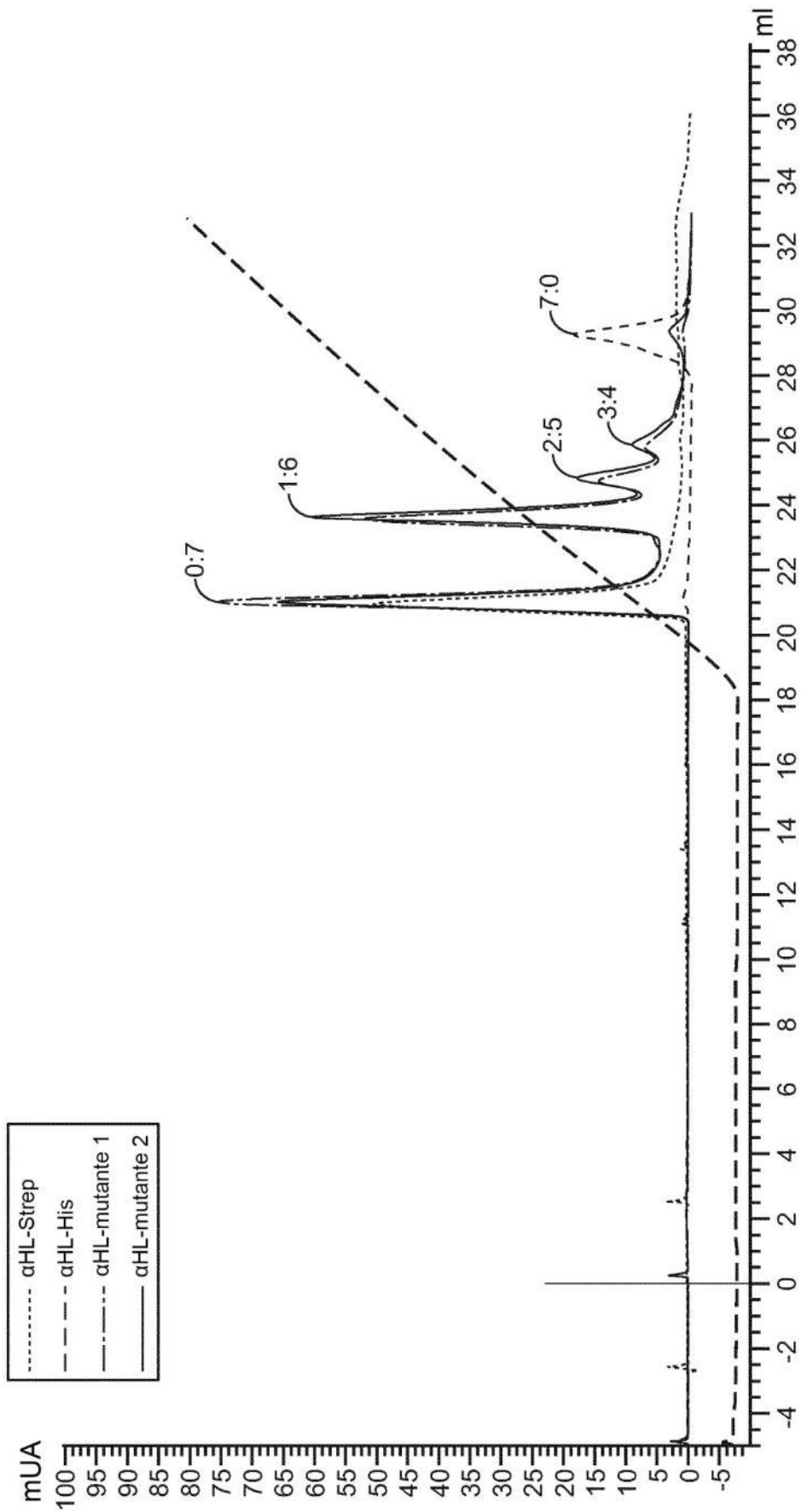


FIG. 29

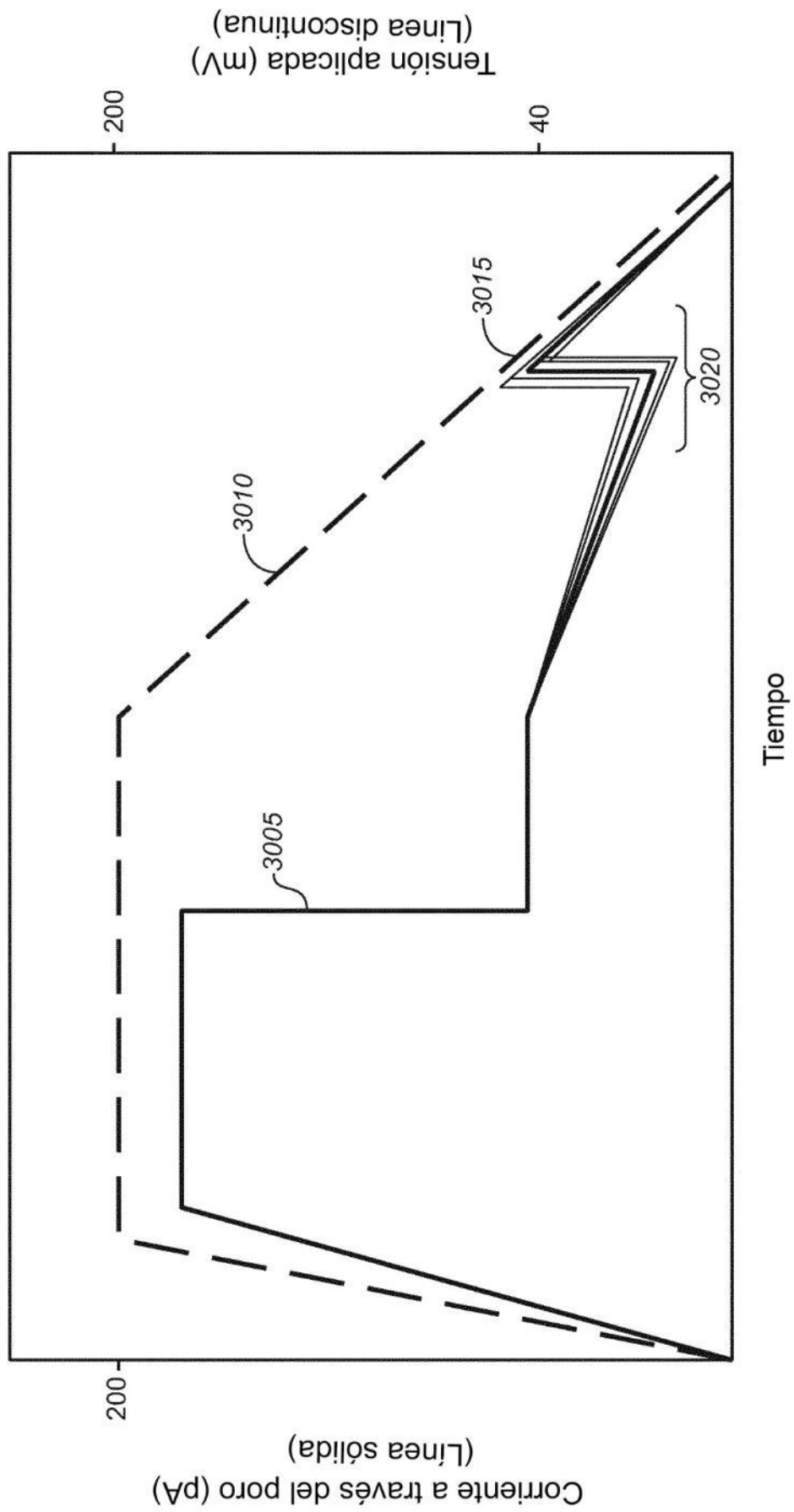


FIG. 30

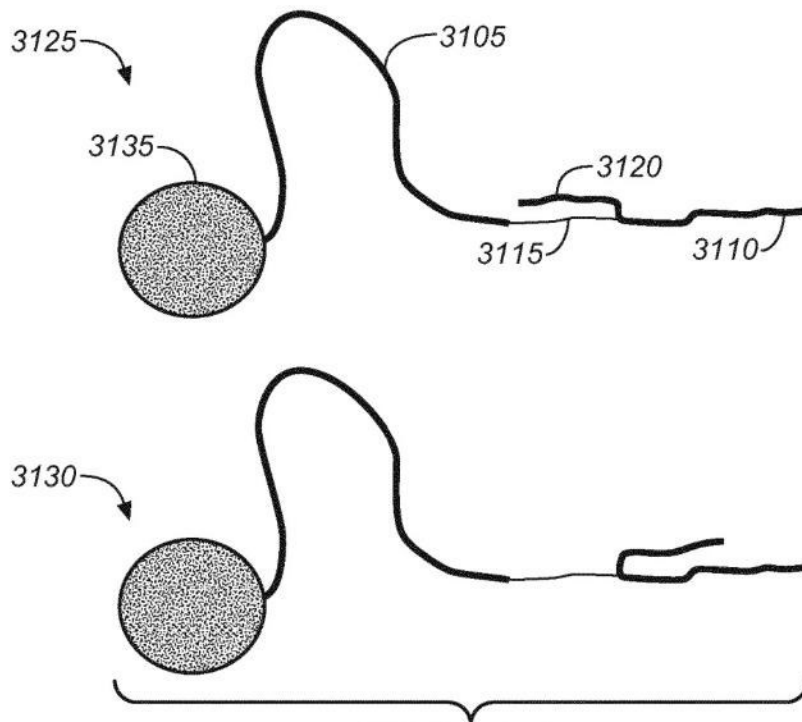


FIG. 31

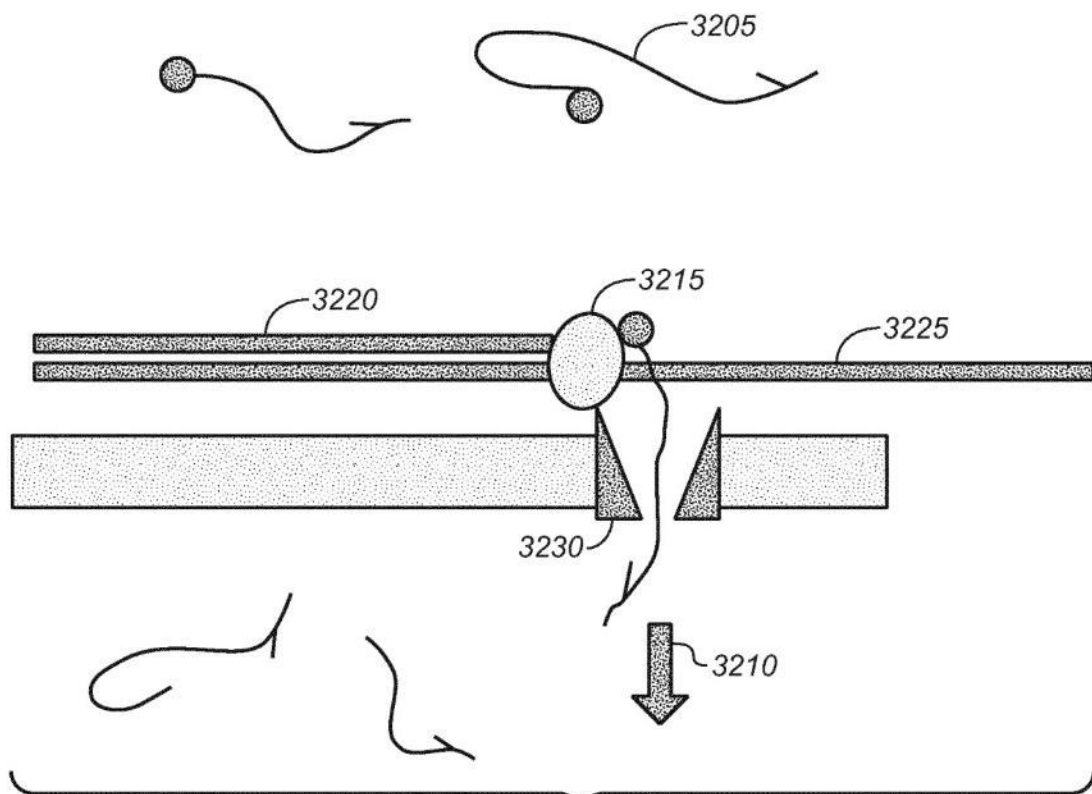


FIG. 32

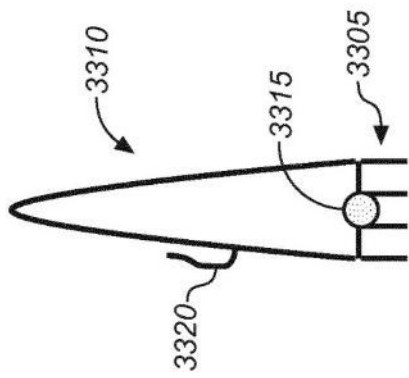


FIG. 33

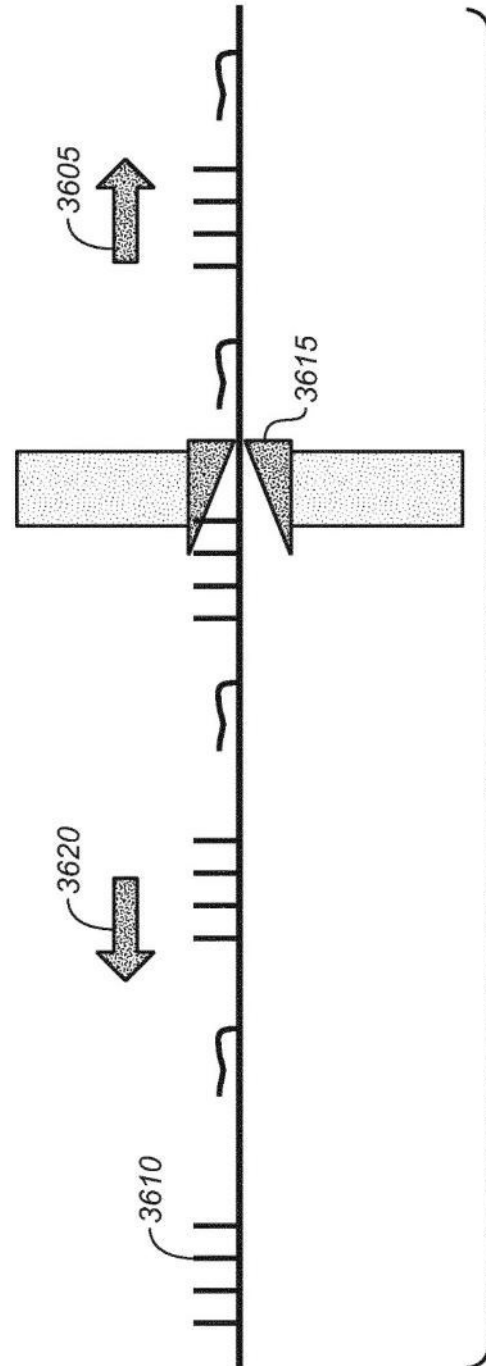
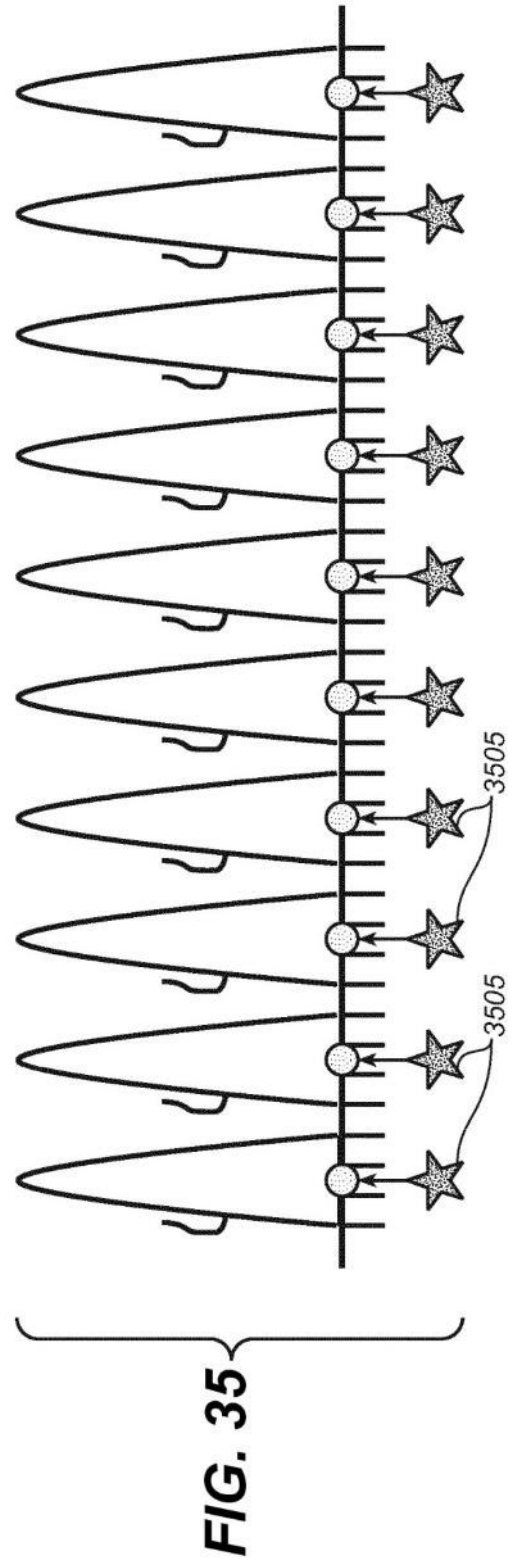
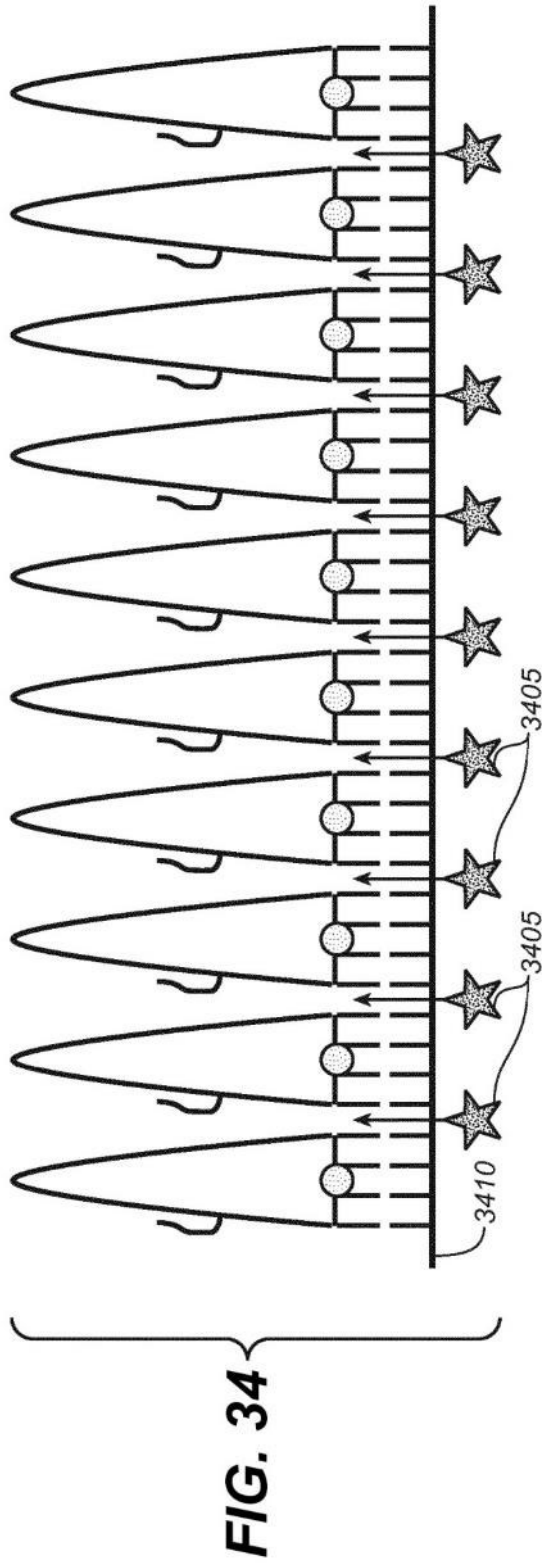


FIG. 36



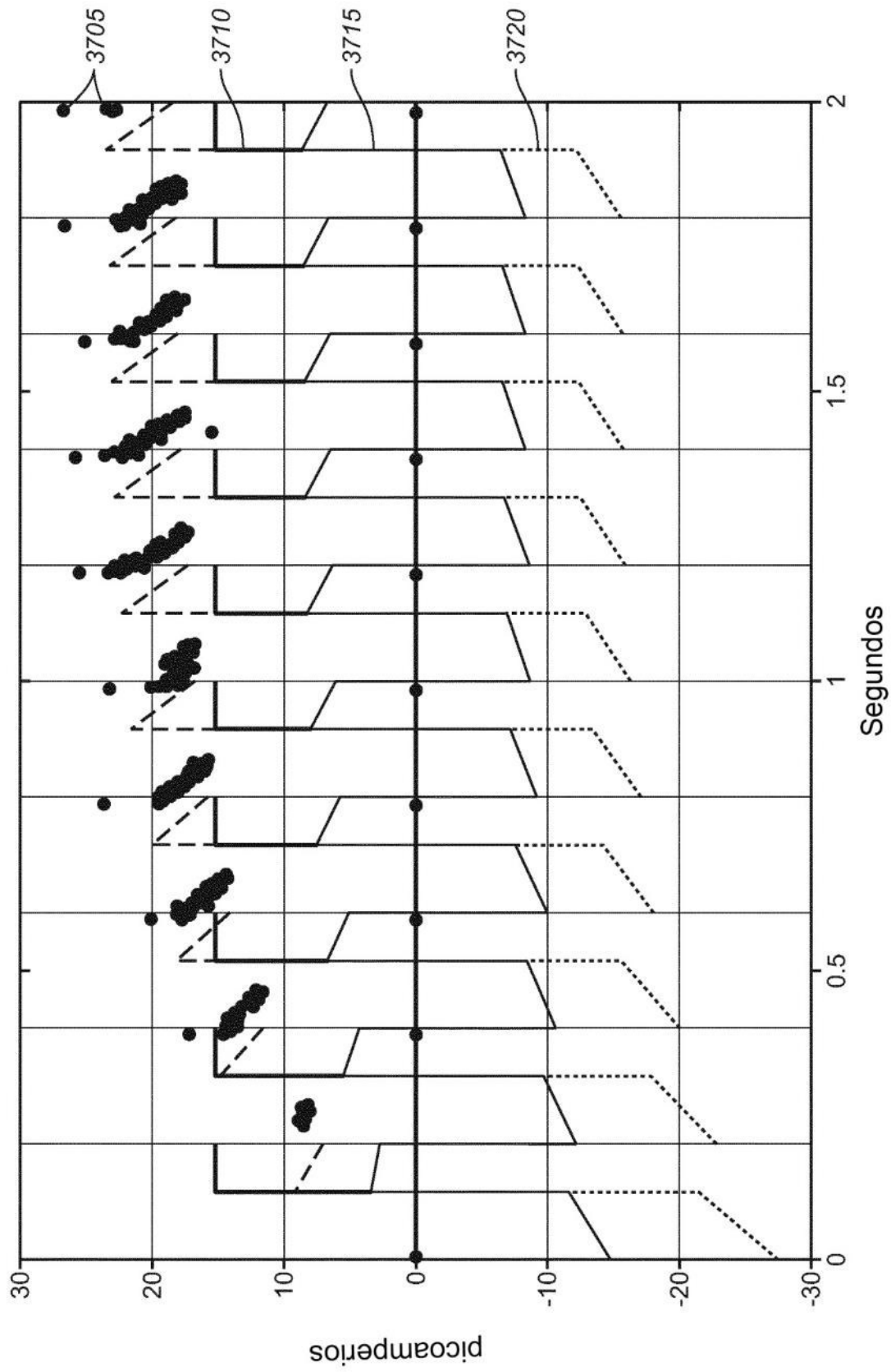


FIG. 37

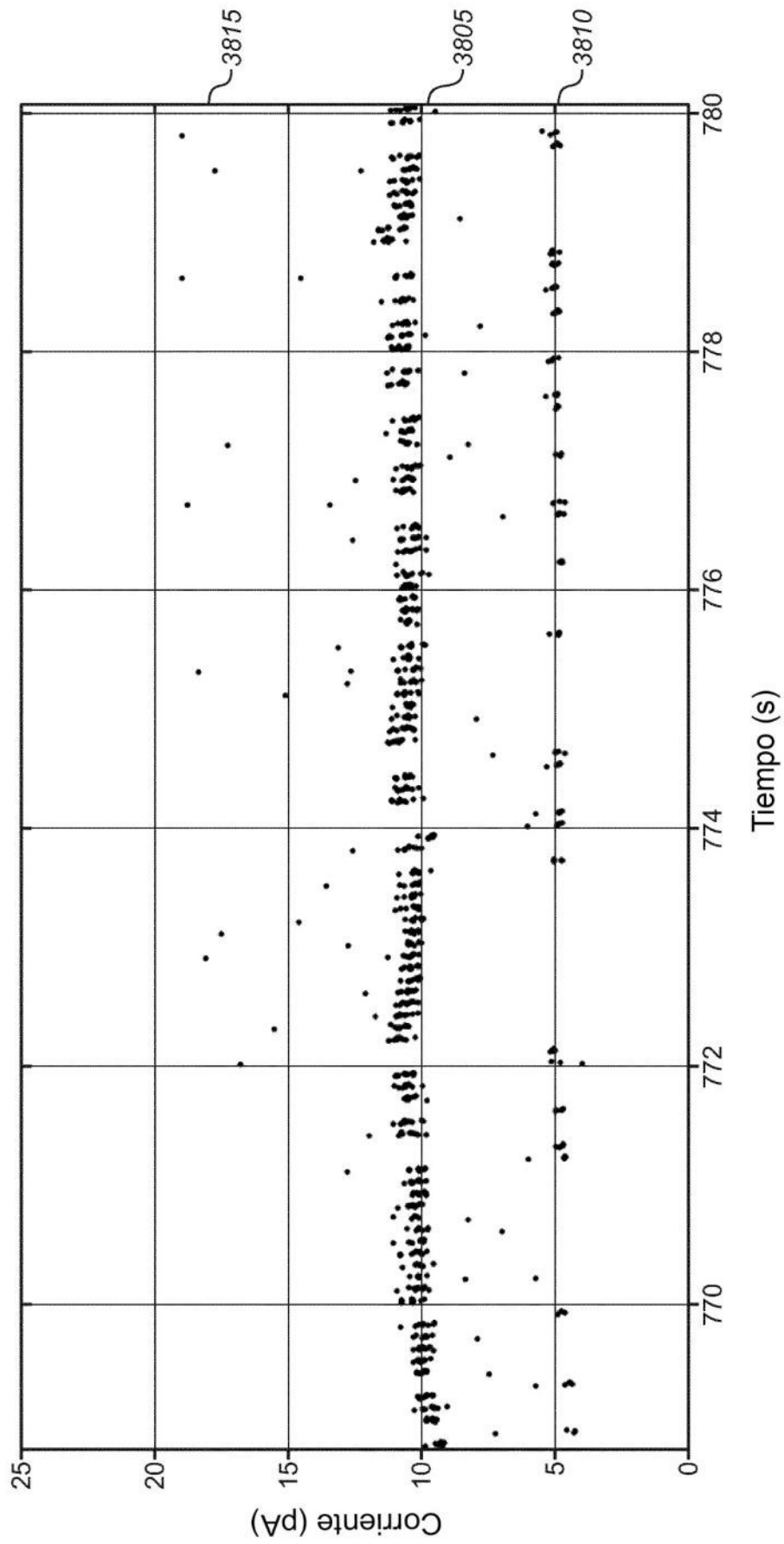


FIG. 38

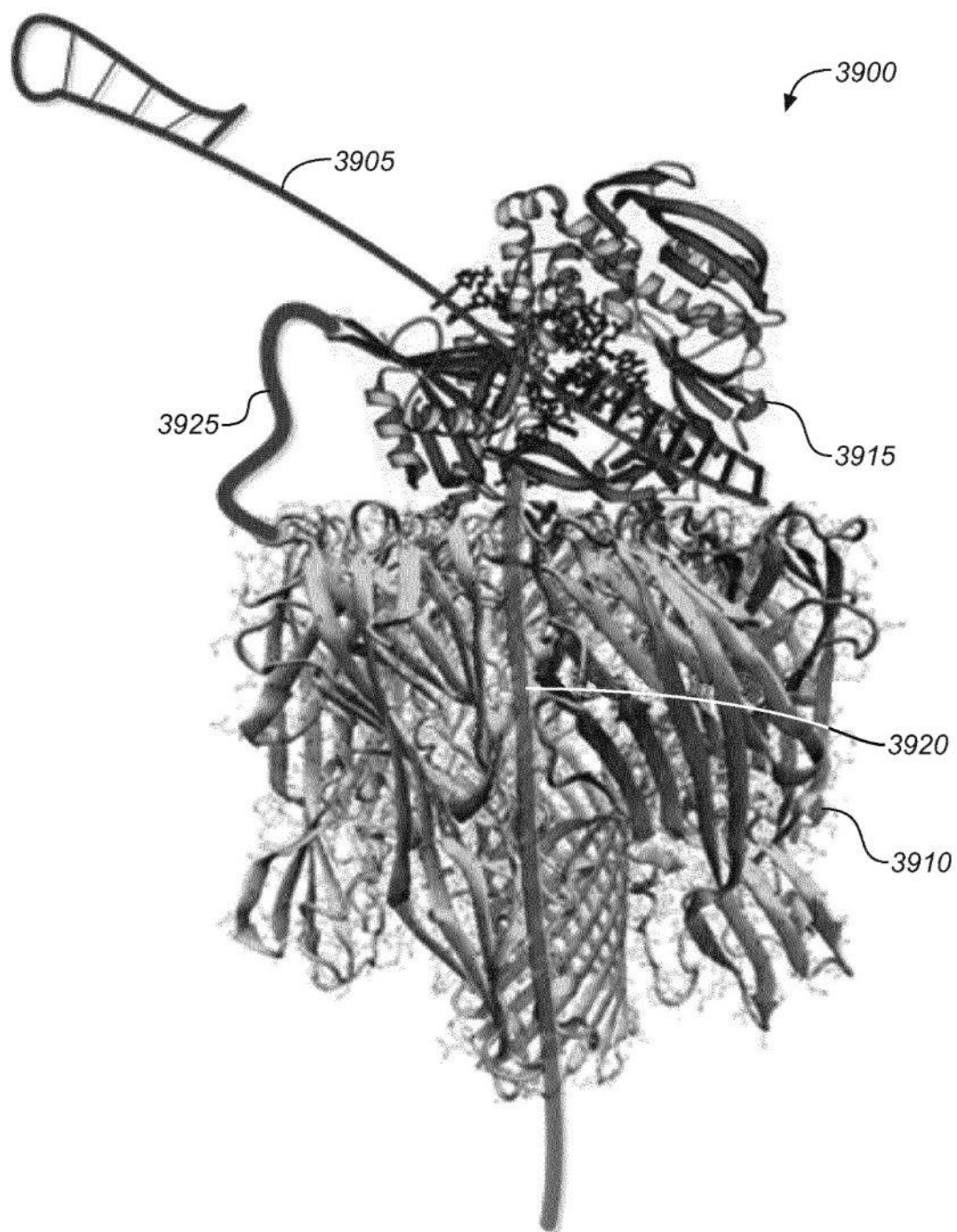


FIG. 39

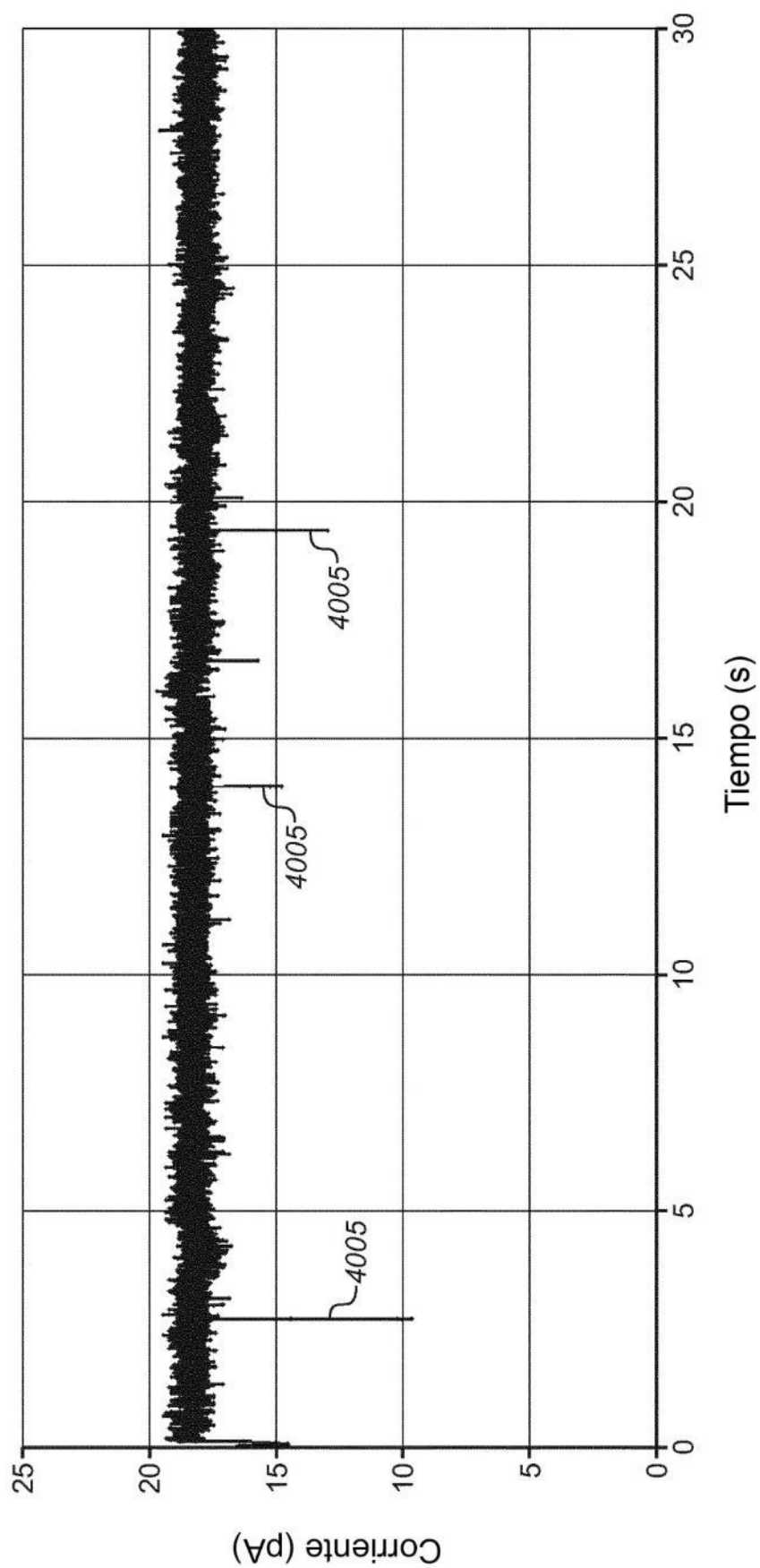


FIG. 40

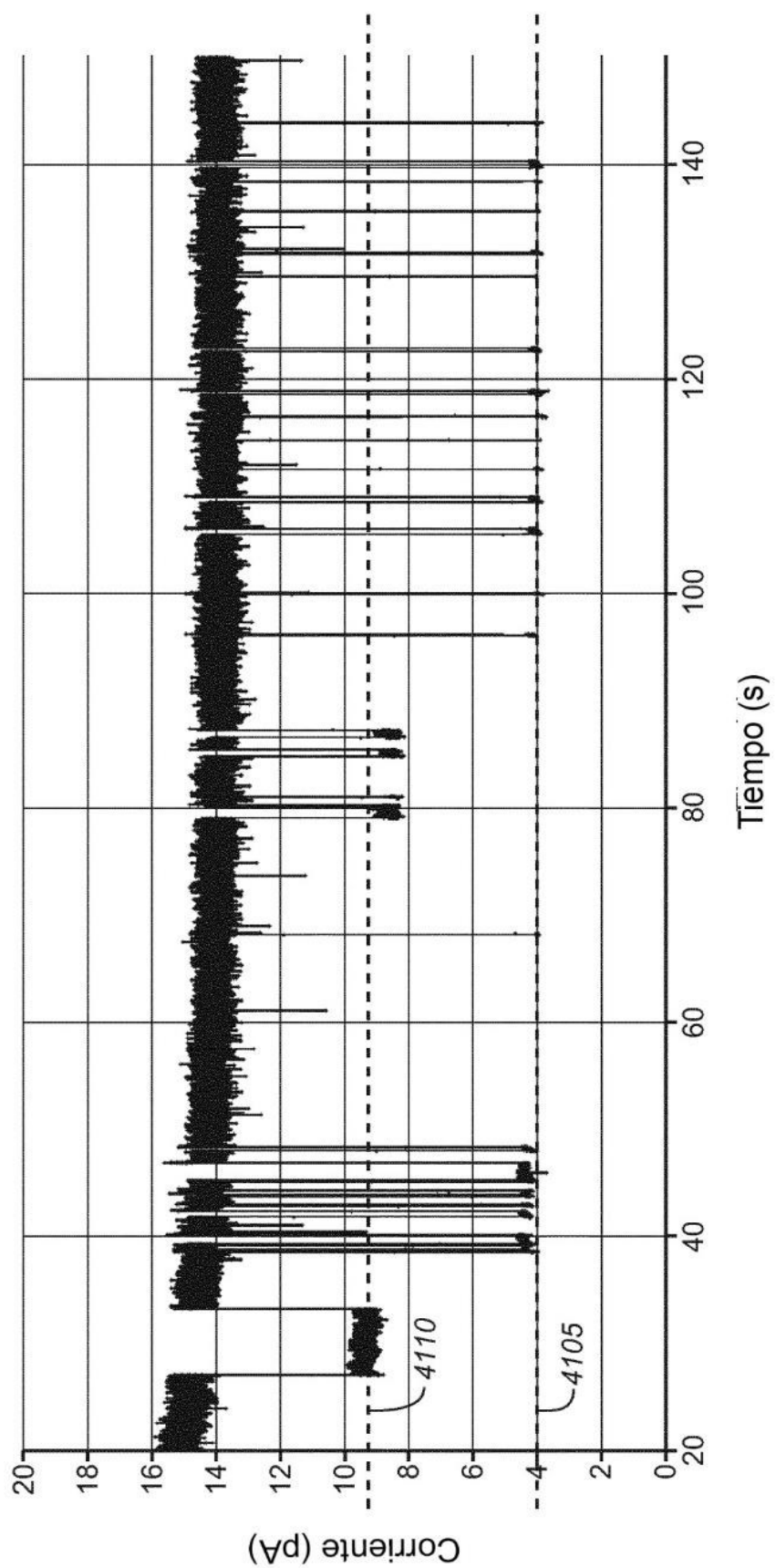


FIG. 41

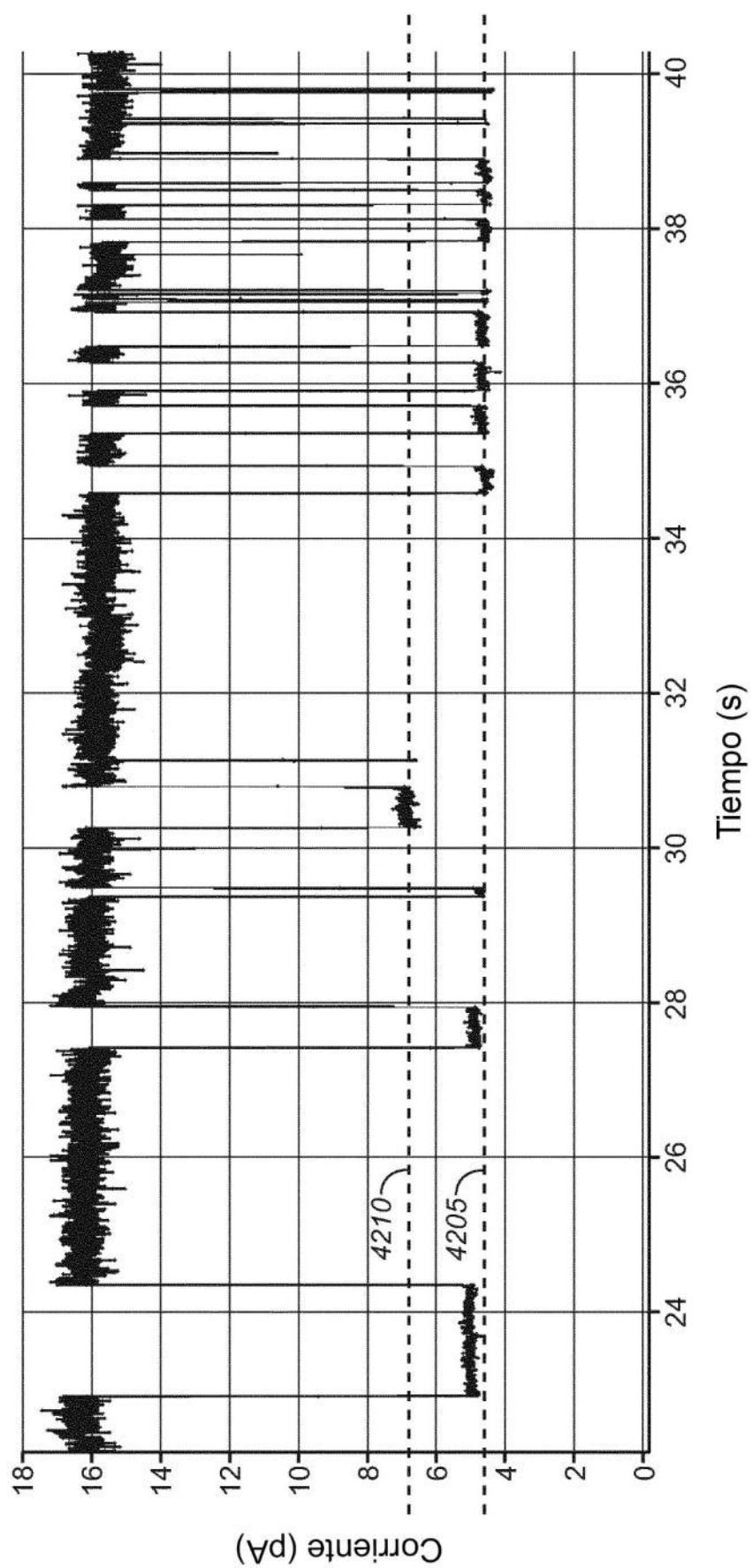


FIG. 42

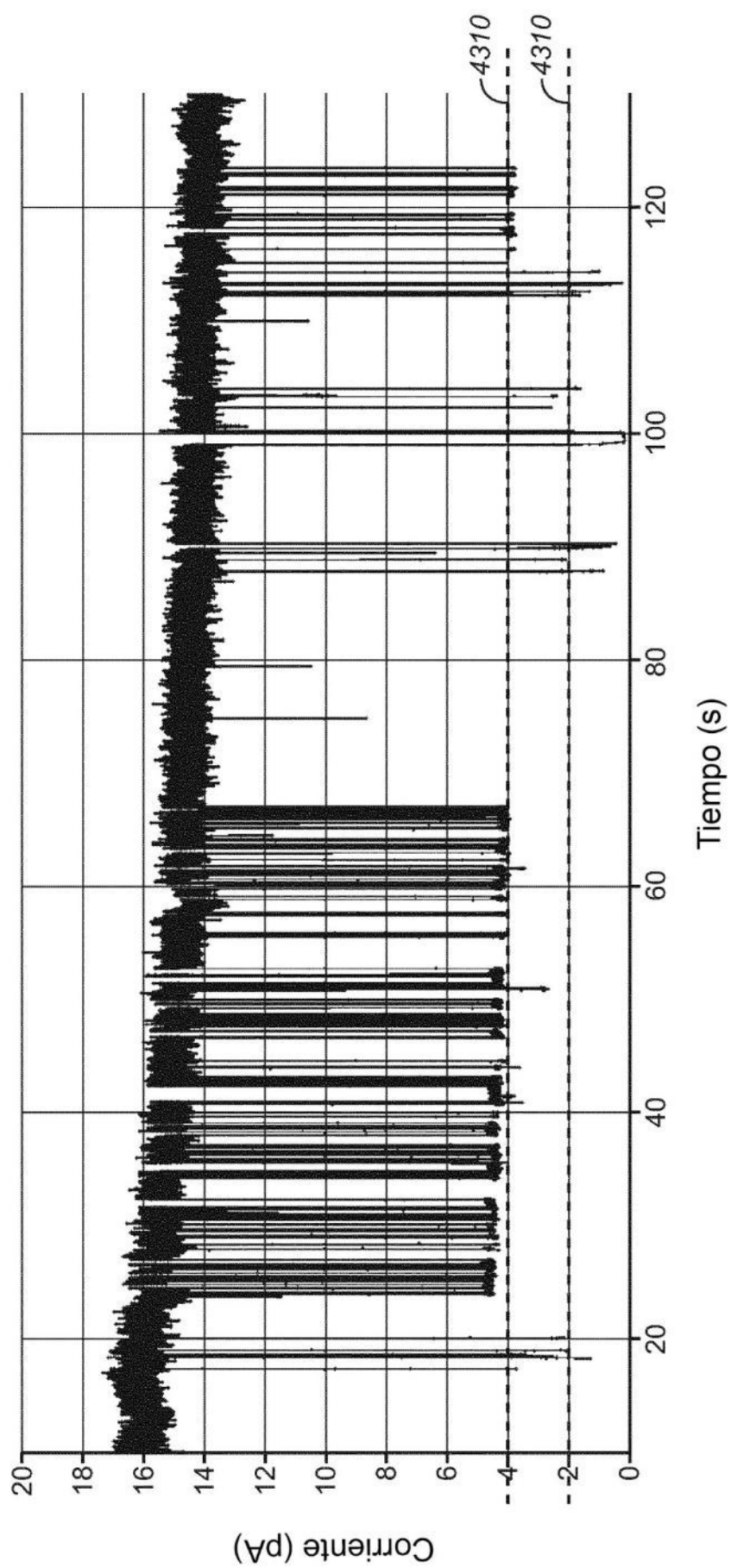


FIG. 43

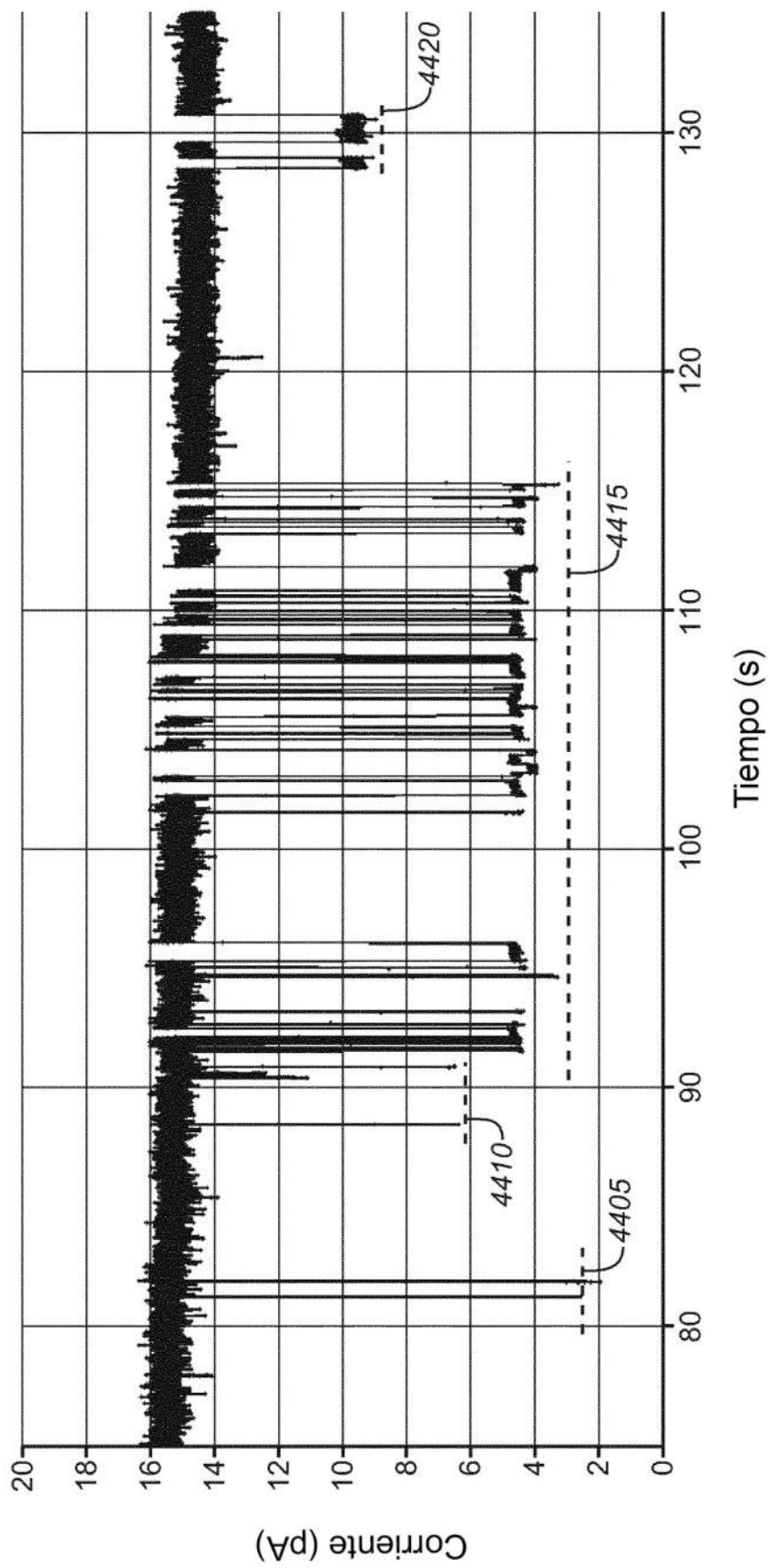


FIG. 44

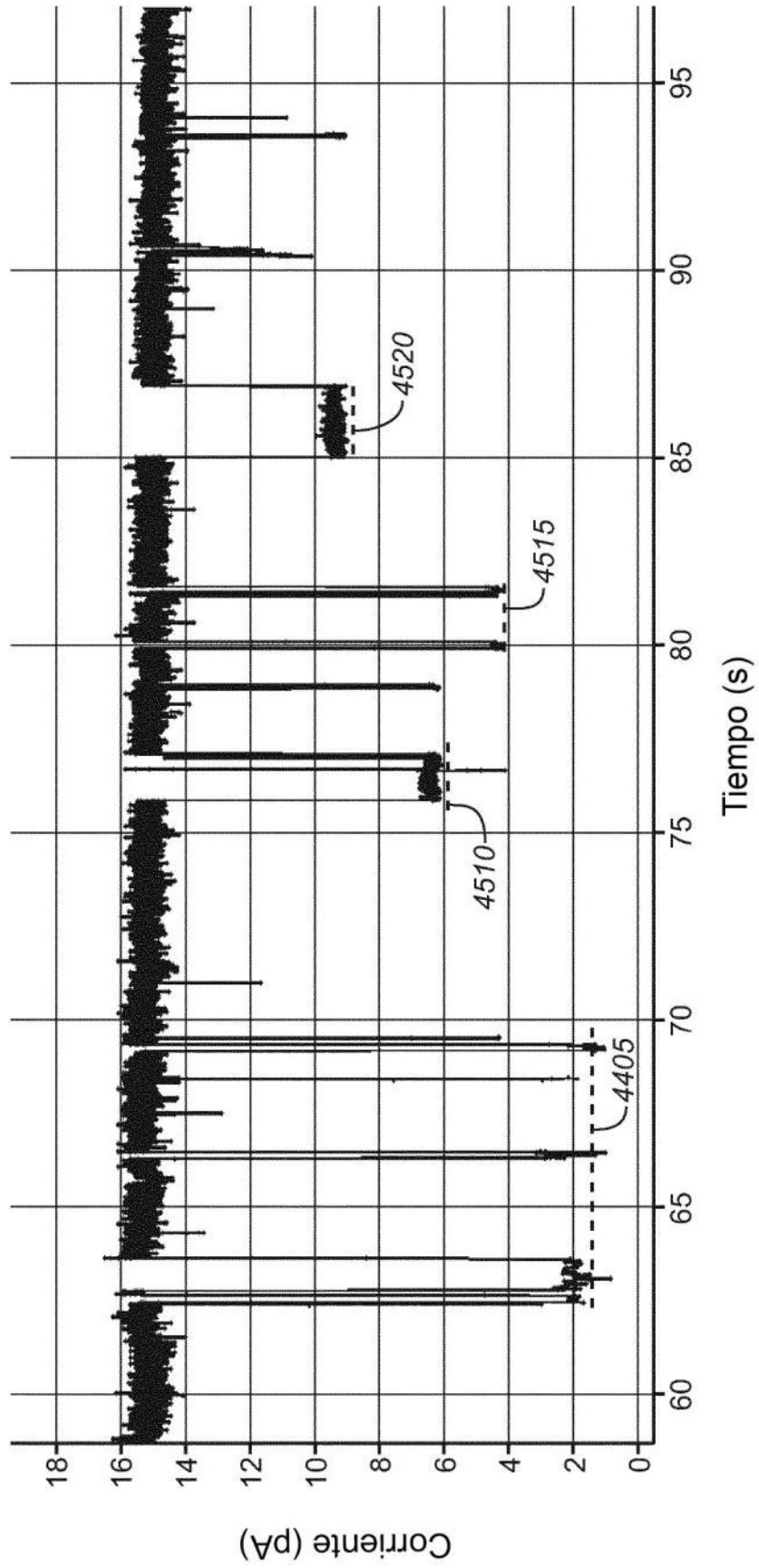


FIG. 45

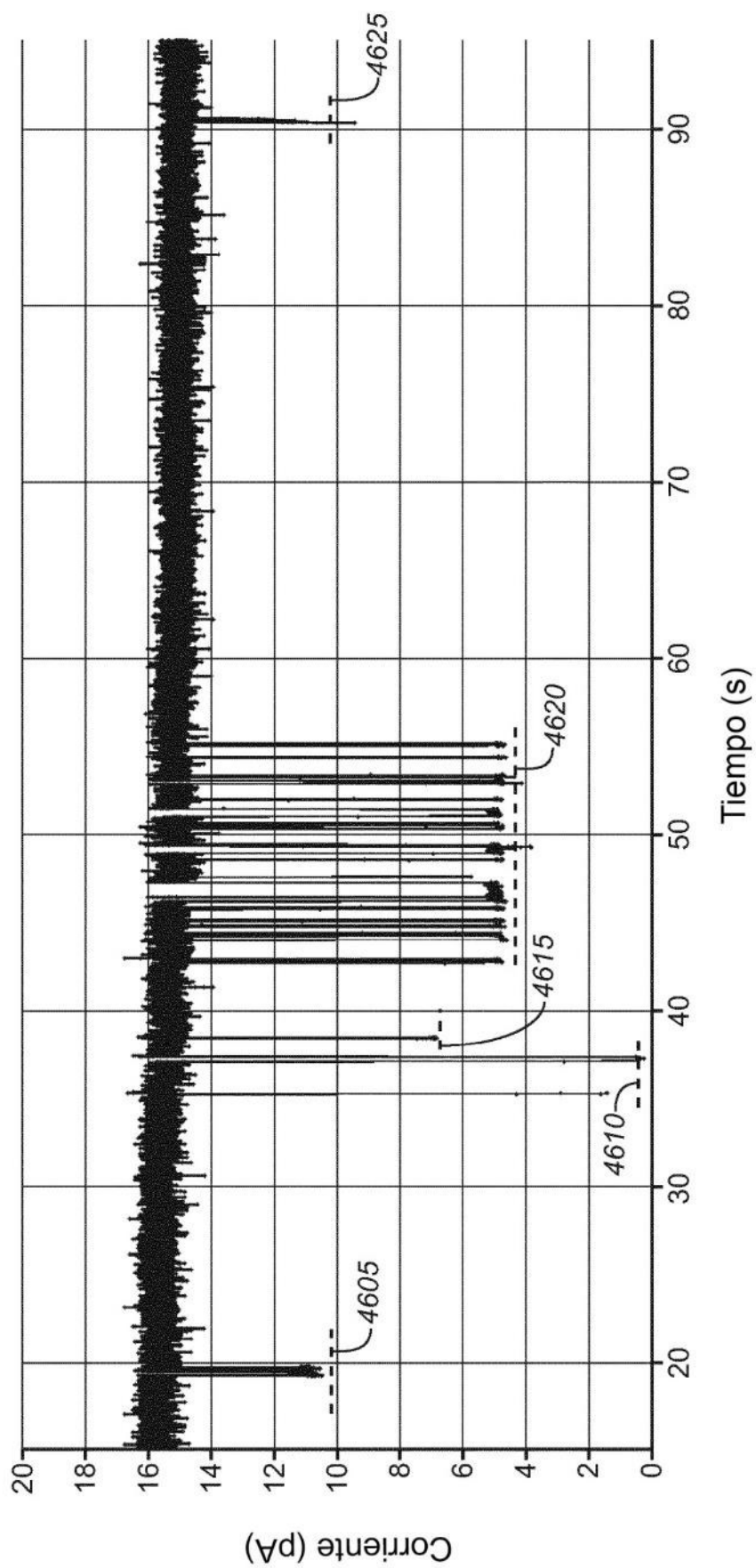


FIG. 46

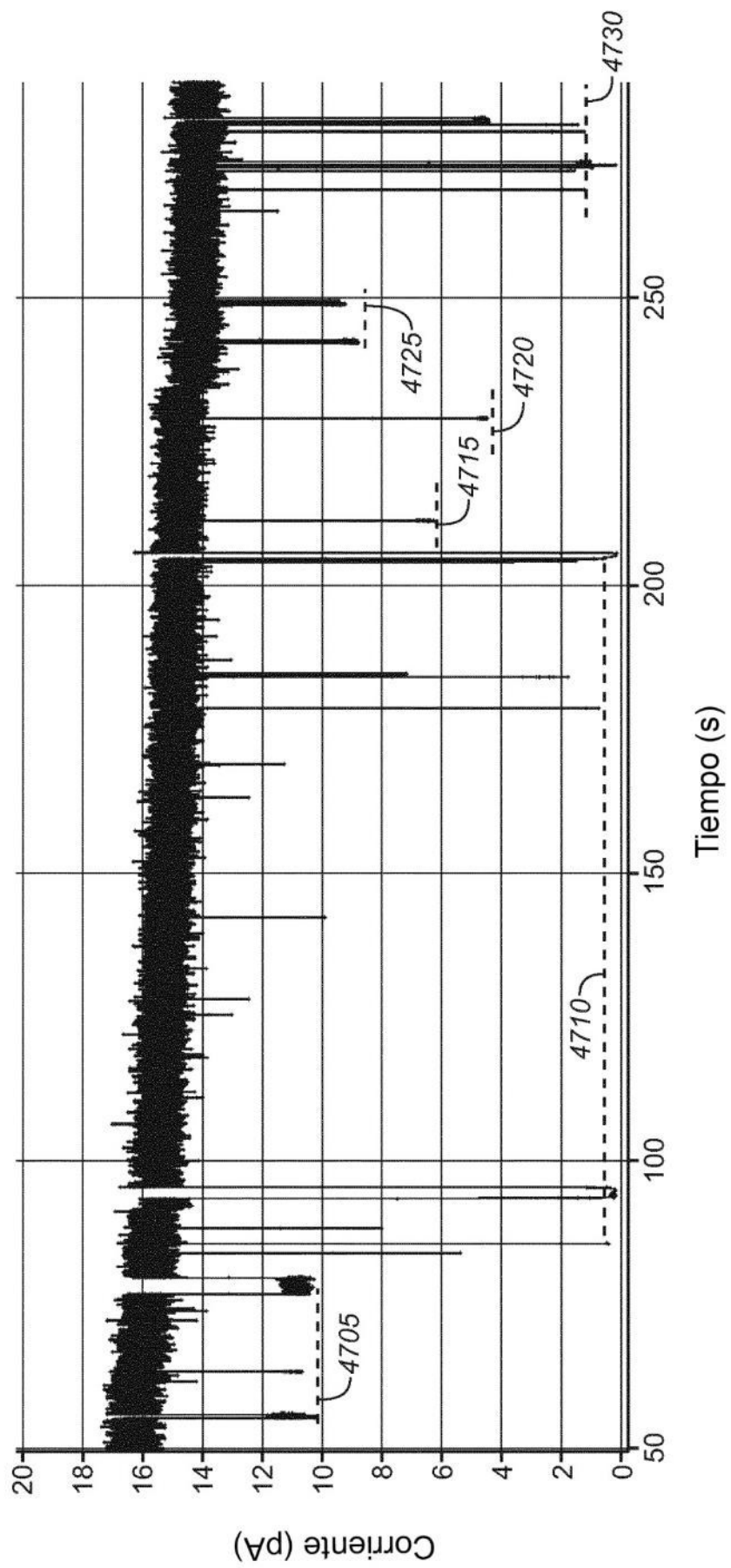


FIG. 47