I803855

(19)中華民國智慧財產局

(12)發明說明書公告本 **(11)證書號數：TW I803855 B**

(45)公告日： 中華民國 112 (2023) 年 06 月 01 日

(21)申請案號：110114376 (22)申請日： 中華民國 110 (2021) 年 04 月 21 日

(51)Int. Cl.： *C12Q1/6869 (2018.01)* *C12Q1/6837 (2018.01)*
B01L3/00 (2006.01)

(30)優先權：2020/04/21 美國 63/013,236

(71)申請人：美商西方數位科技公司 (美國) WESTERN DIGITAL TECHNOLOGIES, INC. (US)
美國
瑞士商赫孚孟拉羅股份公司 (瑞士) F. HOFFMANN-LA ROCHE AG (CH)
瑞士

(72)發明人：阿斯鐵爾 揚恩 ASTIER, YANN (FR)；帕拉都古 斯里 PALADUGU, SRI (IN)；托普蘭奇克 朱菈 TOPOLANCIK, JURAJ (SK)；布拉干薩 派翠克 M BRAGANCA, PATRICK M. (US)

(74)代理人：陳長文

(56)參考文獻：

TW	200306167A	TW	201219770A
CN	102686997B	US	10260095B2
US	2010/0039105A1	US	2018/0237850A1

審查人員：周永泰

申請專利範圍項數：128 項 圖式數：50 共 202 頁

(54)名稱

用於定序核酸之系統及裝置、定序複數個 S 個核酸股之方法及減輕由於使用單分子感測器陣列之核酸定序程序而產生之定序資料之錯誤之方法

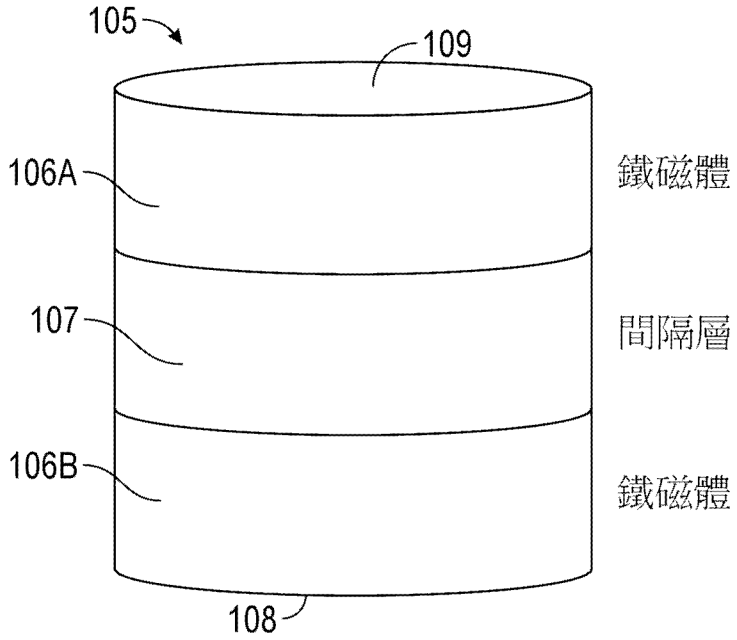
(57)摘要

本文揭示單分子陣列定序(SMAS)裝置及系統之實施例。該 SMAS 裝置之感測器陣列中之各感測器能夠偵測附接至併入與各別結合位點結合之單個核酸股中之核苷酸之標記。各感測器可偵測附接至該併入的核苷酸之單個標記(例如螢光、磁性、有機金屬、帶電荷分子等)。亦揭示使用 SMAS 裝置及系統基於固定於此種 SMAS 裝置上的選殖擴增之 DNA 之多種例項之合成定序(SBS)進行高度可調之核酸(例如 DNA)定序之方法。亦揭示錯誤校正方法，該方法減輕在定序個別核酸股中產生的錯誤(例如偵測到或未偵測到錯誤標記)。

Disclosed herein are embodiments of single-molecule array sequencing (SMAS) devices and systems. Each sensor of an array of sensors of the SMAS device is capable of detecting labels attached to nucleotides incorporated into a single nucleic acid strand bound to a respective binding site. Each sensor can detect a single label (e.g., fluorescent, magnetic, organometallic, charged molecule, etc.) attached to the incorporated nucleotide. Also disclosed are methods of using SMAS devices and systems for highly-scalable nucleic acid (e.g., DNA) sequencing based on sequencing by synthesis (SBS) of multiple instances of clonally amplified DNA immobilized on such SMAS devices. Also disclosed are error correction methods that mitigate errors (e.g., errant label detections or non-detections) made in sequencing individual nucleic acid strands.

TW I803855 B

指定代表圖：



【圖1】

符號簡單說明：

105:磁感測器

106A:鐵磁層

106B:鐵磁層

107:非磁性間隔層

108:底表面

109:頂表面

I803855

【發明摘要】

【中文發明名稱】

用於定序核酸之系統及裝置、定序複數個S個核酸股之方法及減輕由於使用單分子感測器陣列之核酸定序程序而產生之定序資料之錯誤之方法

【英文發明名稱】

SYSTEM AND DEVICE FOR SEQUENCING NUCLEIC ACID, METHOD OF SEQUENCING A PLURALITY OF S NUCLEIC ACID STRANDS, AND METHOD OF MITIGATING ERRORS IN SEQUENCING DATA GENERATED AS A RESULT OF A NUCLEIC ACID SEQUENCING PROCEDURE USING A SINGLE-MOLECULE SENSOR ARRAY

【中文】

本文揭示單分子陣列定序(SMAS)裝置及系統之實施例。該SMAS裝置之感測器陣列中之各感測器能夠偵測附接至併入與各別結合位點結合之單個核酸股中之核苷酸之標記。各感測器可偵測附接至該併入的核苷酸之單個標記(例如螢光、磁性、有機金屬、帶電荷分子等)。亦揭示使用SMAS裝置及系統基於固定於此種SMAS裝置上的選殖擴增之DNA之多種例項之合成定序(SBS)進行高度可調之核酸(例如DNA)定序之方法。亦揭示錯誤校正方法，該方法減輕在定序個別核酸股中產生的錯誤(例如偵測到或未偵測到錯誤標記)。

【英文】

Disclosed herein are embodiments of single-molecule array sequencing (SMAS) devices and systems. Each sensor of an array of sensors of the SMAS device is capable of detecting labels attached to nucleotides incorporated into a single nucleic acid strand bound to a respective binding site. Each sensor can detect a single label (e.g., fluorescent, magnetic, organometallic, charged molecule, etc.) attached to the incorporated nucleotide. Also disclosed are methods of using

SMAS devices and systems for highly-scalable nucleic acid (e.g., DNA) sequencing based on sequencing by synthesis (SBS) of multiple instances of clonally amplified DNA immobilized on such SMAS devices. Also disclosed are error correction methods that mitigate errors (e.g., errant label detections or non-detections) made in sequencing individual nucleic acid strands.

【指定代表圖】

圖1

【代表圖之符號簡單說明】

105: 磁感測器

106A: 鐵磁層

106B: 鐵磁層

107: 非磁性間隔層

108: 底表面

109: 頂表面

【發明說明書】

【中文發明名稱】

用於定序核酸之系統及裝置、定序複數個S個核酸股之方法及減輕由於使用單分子感測器陣列之核酸定序程序而產生之定序資料之錯誤之方法

【英文發明名稱】

SYSTEM AND DEVICE FOR SEQUENCING NUCLEIC ACID, METHOD OF SEQUENCING A PLURALITY OF S NUCLEIC ACID STRANDS, AND METHOD OF MITIGATING ERRORS IN SEQUENCING DATA GENERATED AS A RESULT OF A NUCLEIC ACID SEQUENCING PROCEDURE USING A SINGLE-MOLECULE SENSOR ARRAY

【技術領域】

【先前技術】

【0001】 商業上成功的DNA定序方法涉及選殖去氧核糖核酸(DNA)簇之合成及分析或個別DNA分子之偵測。儘管簇定序儀展現對於診斷應用而言足夠低之錯誤率，但由於分子集體(molecular ensemble)中之錯誤傳播性質，其讀取長度受到很大限制。單分子定序儀可產生顯著較長讀段，但通常展現靜態及動態異質性，該異質性導致對於高精度診斷而言過於大的錯誤。

【0002】 因此，一般而言需要改良DNA定序及核酸定序，以實現具有較低錯誤率之較長讀段。

【發明內容】

【0003】 本[發明內容]表示本發明之非限制性實施例。

【0004】 本文揭示單分子陣列定序(SMAS)裝置及系統之實施例。SMAS裝置之感測器陣列內的複數個感測器中之各感測器偵測附接至併入與各別結合位點結合之單個核酸股中之核苷酸之標記。各感測器可偵測附接至併入的核苷酸之單個標記(例如螢光、磁性、有機金屬、帶電荷分子等)。亦揭示使用SMAS裝置及系統基於固定於此種SMAS裝置上的選殖擴

增之DNA之多種例項之合成定序(SBS)進行高度可調之核酸(例如DNA)定序之方法。亦揭示錯誤校正方法，該方法減輕在定序個別核酸股中產生的錯誤(例如偵測到或未偵測到錯誤標記)。

【0005】 在一些實施例中，用於定序核酸之裝置包含流體腔室、經結構設計成偵測存在於流體腔室中之標記之複數個S個磁感測器、及至少一個處理器。流體腔室包含複數個S個結合位點，該等S個結合位點中之各者經結構設計成結合不超過一個核酸股。S個磁感測器中之各者感測與S個結合位點之各別結合位點結合之核酸之各別股。該至少一個處理器經結構設計成執行一或多個機器可執行之指令，該等指令在執行時導致至少一個處理器在定序程序之複數個M個查詢步驟中之各查詢步驟且針對S個磁感測器中之各者(a)獲得各別磁感測器之各別特性，其中該各別特性指示至少一個標記之存在或不存在，且(b)至少部分地基於所獲得的各別特性，來判定在查詢步驟期間各別磁感測器是否偵測到至少一個標記之存在或不存在。

【0006】 在一些實施例中，系統包含複數個S個結合位點(該等S個結合位點中之各者經結構設計成結合不超過一個核酸股)、經結構設計成偵測標記之複數個S個感測器(例如磁性、光學感測器等)、及至少一個處理器。S個感測器中之各者經結構設計成感測與S個結合位點之各別結合位點結合之核酸之各別股。該至少一個處理器經結構設計成執行一或多個機器可執行之指令，該等指令在執行時導致至少一個處理器在定序程序之複數個M個查詢步驟中之各查詢步驟且針對S個感測器中之各者(a)獲得各別感測器之各別特性，其中該各別特性指示至少一個標記之存在或不存在，且(b)至少部分地基於所獲得的各別特性，來判定在查詢步驟期間各

別感測器是否偵測到至少一個標記之存在或不存在。另外，在執行時，該一或多個機器可執行之指令進一步導致至少一個處理器進行於至少一個記錄上之錯誤校正程序，該至少一個記錄包含在M個查詢步驟中之各步驟針對至少一個S個感測器子組之定序程序之結果。

【0007】 在一些實施例中，一種使用SMAS裝置定序複數個S個核酸股之方法包括(a)使S個核酸股與S個結合位點結合，(b)進行包括M個查詢步驟之定序程序以產生S個記錄，該等S個記錄中之各者捕捉S個感測器中各別感測器之M個偵測結果，該M個偵測結果中之各者指示在M個查詢步驟中之各別步驟期間，S個感測器中之各別感測器在流體腔室中是否偵測到至少一個標記，及(c)對至少一個S個記錄子組應用錯誤校正程序以估計S個核酸股中之至少一個股之核酸序列。

【0008】 一些實施例係一種減輕由於使用單分子感測器陣列之核酸定序程序產生之定序資料之錯誤之方法，該單分子感測器陣列具有複數個感測器，該複數個感測器中之各者與複數個結合位點中之各別結合位點相關聯，該複數個結合位點中之各者經結構設計成結合不超過一個待定序的核酸股。在一些此類實施例中，該方法包括(a)識別定序資料中之複數個記錄，該複數個記錄中之各者捕捉核酸之第一股之各別例項之各別定序結果，該複數個記錄中之各者具有複數個條目，該複數個條目中之各者指示對於核酸定序程序之複數個查詢步驟之各別步驟，(i)藉由與核酸之第一股之各別例項相關之各別感測器偵測到標記，或(ii)藉由與核酸之第一股之各別例項相關之各別感測器沒有偵測到標記；(b)基於複數個記錄，判定核酸之第一股之複數個候選序列，該複數個候選序列中之各者估計核酸之第一股之核酸序列之至少一部分；及(c)識別該複數個候選序列中之特定

候選序列為核酸之第一股之至少一部分核酸序列，該特定候選序列自該複數個候選序列中為最可能正確的。

【0009】 與基於簇之方法相比，所揭示的定序及錯誤校正裝置、系統及方法有望實現更高通量，更低錯誤率，及更長讀段長度。

【圖式簡單說明】

【0010】 自結合附圖進行的某些實施例之以下描述當可輕易地明瞭本發明之目標、特徵及優點，其中：

【0011】 圖1說明根據一些實施例之磁感測器之一部分。

【0012】 圖2A及2B說明磁阻(MR)感測器之電阻，其可根據一些實施例進行使用。

【0013】 圖3A說明自旋轉矩振盪器(STO)感測器，其可根據一些實施例進行使用。

【0014】 圖3B顯示STO在實例條件下之實驗反應。

【0015】 圖3C及3D說明STO之短奈秒場脈衝，其可根據一些實施例進行使用。

【0016】 圖4A說明簇定序裝置之單個感測器，該簇定序裝置用於感測在其附近的一些*N*個選殖擴增之DNA股。

【0017】 圖4B說明示例性的複數個*S*個單分子感測器，根據一些實施例，各感測器藉由SMAS裝置用於監測各別單股DNA (ssDNA)。

【0018】 圖5A為顯示根據一些實施例之用於核酸定序之示例性SMAS裝置之組件之方塊圖。

【0019】 圖5B、5C及5D說明根據一些實施例之用於核酸定序之例示性SMAS裝置之部分。

【0020】 圖5E說明根據一些實施例之感測器之正方形格網(或格子)圖案。

【0021】 圖6A說明根據一些實施例之感測器、處於螺旋狀態中之DNA股及標記。

【0022】 圖6B說明根據一些實施例之感測器、長形DNA股及標記之示例性尺寸。

【0023】 圖7A說明根據一些實施例之用於估計SMAS裝置之感測器陣列封裝極限之示例性幾何配置。

【0024】 圖7B說明根據一些實施例之以正方形格子配置之SMAS裝置之感測器。

【0025】 圖8A及8B說明根據一些實施例之以六邊形圖案配置之SMAS裝置之感測器。

【0026】 圖9A說明根據一些實施例之用於估計SMAS裝置之感測器陣列封裝極限之示例性幾何配置。

【0027】 圖9B說明根據一些實施例之以六邊形格子配置之SMAS裝置之感測器。

【0028】 圖10比較示例性SMAS實施方案之密度與當前最先進技術簇定序裝置。

【0029】 圖11說明根據一些實施例之使用SMAS裝置定序複數個核酸股之示例性方法。

【0030】 圖12為根據一些實施例之使用加性方法之定序程序之流程圖。

【0031】 圖13說明根據一些實施例之加性定序方案。

【0032】 圖14為根據一些實施例之使用減性方法之定序程序之流程圖。

【0033】 圖15說明根據一些實施例之減性定序方案。

【0034】 圖16為根據一些實施例之使用改進型加性方法之定序程序之流程圖。

【0035】 圖17說明根據一些實施例之改進型加性定序方案。

【0036】 圖18A說明簇定序裝置之失敗之核苷酸併入(FNI)。

【0037】 圖18B說明用於SMAS裝置之FNI。

【0038】 圖18C說明簇定序裝置之失敗之標記移除(FLR)。

【0039】 圖18D說明用於SMAS裝置之FLR。

【0040】 圖18E說明簇定序裝置之失敗之核苷酸移除(FNR)。

【0041】 圖18F說明用於SMAS裝置之FNR。

【0042】 圖18G說明簇定序裝置之失敗之核苷酸刪除(FLD)。

【0043】 圖18H說明用於SMAS裝置之FLD。

【0044】 圖19為根據一些實施例之使用具有FLR及FNI錯誤偵測之改進型加性方法之示例性定序程序之流程圖。

【0045】 圖20顯示具有FNI及FLR錯誤之實例記錄。

【0046】 圖21說明藉由簇定序裝置感測器偵測到的預期信號位準，該感測器捕捉定序程序期間分子集體之行為。

【0047】 圖22說明根據一些實施例之當使用錯誤校正技術時SMAS裝置如何提供更佳準確度。

【0048】 圖23說明根據一些實施例之藉由在來自定序程序之偵測結果之記錄中刪除若干串四個「未偵測到標記」條目之FNI錯誤校正。

【0049】 圖24說明根據一些實施例之示例性SBS反應之結果。

【0050】 圖25說明較大簇尺寸於簇定序裝置之鹼基識別(base-calling)準確度之影響。

【0051】 圖26說明根據一些實施例之FLR及FNI錯誤之判定性錯誤校正。

【0052】 圖27說明偵測資料中之FNI、FLR及FNR錯誤。

【0053】 圖28說明根據一些實施例之藉由SMAS裝置產生之資料之FLR錯誤校正及鹼基識別。

【0054】 圖29說明根據一些實施例之藉由SMAS裝置產生之資料之FNI錯誤校正及鹼基識別。

【0055】 圖30說明根據一些實施例之藉由SMAS裝置產生之資料之錯誤校正及鹼基識別。

【0056】 圖31說明來自SMAS裝置之示例性偵測結果中之FNI、FLR、FNR及FLD錯誤。

【0057】 圖32說明根據一些實施例之對SBS期間藉由SMAS裝置捕捉的資料之錯誤校正程序之應用。

【0058】 圖33為說明根據一些實施例之錯誤校正程序之流程圖。

【0059】 圖34A說明在查詢步驟之平均信號強度，在該步驟應偵測到標記，因為引入且成功地併入匹配的核苷酸。

【0060】 圖34B說明自簇模型測得的強度之函數擬合。

【0061】 圖35繪示簇定序裝置之概率函數。

【0062】 圖36說明簇定序裝置之離散概率函數。

【0063】 圖37A說明簇定序裝置之強度曲線。

【0064】 圖37B說明簇定序裝置之概率分佈函數。

【0065】 圖38A及38B繪示簇定序裝置之概率函數。

【0066】 圖39說明在各種條件下簇定序裝置之N-r參數空間。

【0067】 圖40A顯示針對各種N-r組合之沿著Q30等高線之簇定序裝置之計算概率。

【0068】 圖40B繪示簇定序裝置之計算累積錯誤概率。

【0069】 圖41說明用於簇定序裝置之N-r參數空間，其中該在位置150處的不正確鹼基識別之累積概率為小於或等於100分之一($\tilde{Q}20$)、1,000分之一($\tilde{Q}30$)、10,000分之一($\tilde{Q}40$)、及100,000分之一($\tilde{Q}50$)。

【0070】 圖42說明用於SMAS裝置之K-r參數空間之計算結果，其中根據一些實施例，在各查詢步驟的不正確鹼基識別之概率為小於100分之一(Q20)、1,000分之一(Q30)、10,000分之一(Q40)及100,000分之一(Q50)。

【0071】 圖43A及43B顯示根據一些實施例之針對簇定序裝置及SMAS裝置之位置150處的不正確鹼基識別之累積概率。

【0072】 圖44及45說明根據一些實施例之示例性樣品製備及加載過程。

【0073】 圖46A、46B及46C說明根據一些實施例之示例性SMAS裝置之模擬偵測結果。

【0074】 圖47說明根據一些實施例之說明於圖46A、46B及46C中之偵測資料可如何進行重新配置以識別鹼基且顯示不同DNA股之位置。

【0075】 圖48A及48B繪示根據查詢步驟數量C及化學失敗率r進行不正確鹼基識別之計算概率。

【0076】 圖49繪示根據一些實施例之樣品製備及DNA加載中條碼之用途。

【0077】 圖50說明根據一些實施例之示例性系統160。

【0078】 為了便於理解，在可能的地方使用相同參考數字表示圖中共有的相同元件。經考慮揭示於一個實施例中之元件可在沒有特定敘述下有益地用於其他實施例中。此外，在一個附圖之內文中對元件之描述可應用於說明該元件之其他附圖。

【實施方式】

相關申請案之交叉參考

【0079】 本申請案主張2020年4月21日申請且題為「HIGH-THROUGHPUT DNA SEQUENCING WITH SINGLE-MOLECULE SENSOR-ARRAYS」(代理人檔案編號ROA-1002P-US / P36083-US)之美國臨時申請案第63/013,236號之優先權，且其內容係以其全文引用之方式併入本文中。本申請案亦出於所有目的以引用方式併入2020年4月8日申請之題為「NUCLEIC ACID SEQUENCING BY SYNTHESIS USING MAGNETIC SENSOR ARRAYS」(代理人檔案編號ROA-1000-WO / P35097-WO)之PCT申請案第PCT/US20/27290號(其於2020年10月15日以WO 2020/210370公開)及2021年3月7日申請且題為「MAGNETIC SENSOR ARRAYS FOR NUCLEIC ACID SEQUENCING AND METHODS OF MAKING AND USING THEM」(代理人檔案編號ROA-1001-WO/P35967-WO)之PCT申請案第PCT/US2021/021274號之全文。

【0080】 本文之一些描述及實例係在DNA定序之內文中，但應明瞭，本發明一般適用於核酸定序。

【0081】

術語及註記

如本文所用，術語「股」係指單個核酸股(例如ssDNA)。當指核酸時，術語「股」及「片段」可互換使用。

【0082】 如本文所用，術語「複數個」意指兩個或更多個，但不一定是全部。因此，複數個感測器僅意指至少兩個感測器，而不一定是感測器陣列或定序裝置/系統中之所有感測器。同樣地，複數個結合位點僅意指至少兩個結合位點，不一定是定序裝置/系統中之所有結合位點。

【0083】 如本文所用，術語「例項」在指核酸股時意指模板核酸股或其複本(例如藉由擴增或複製過程產生)。理想地，模板核酸股之複本與模板股相同，但如此項技術中已知，複本由於複製/擴增錯誤而不一定相同。應明瞭，即使擴增程序引入錯誤，藉由擴增產生之重複體仍視作原始核酸股之複本。因此，股之所有例項理想上彼此相同但可能是不相同。

【0084】 如本文所用，術語「查詢循環」係指核酸定序程序之單個循環，在該循環期間，引入所有可能的核苷酸以判定哪一個(若有的話)被引入至經定序之股中。例如，對於DNA定序程序，所有腺嘌呤(**A**)、胸腺嘧啶(**T**)、胞嘧啶(**C**)及鳥嘌呤(**G**)均以某種(任意)順序(該順序不需要各查詢循環都相同)進行測試。如下文所詳細說明，取決於所選擇的定序程序，在單個定序循環期間每個股可偵測到超過一個標記。

【0085】 如本文所用，術語「查詢步驟」係指定序程序之步驟或步驟集合，在該程序期間，判定定序裝置之一或多個感測器是否偵測標記。對於遍及所有**A**、**T**、**C**及**G**之DNA定序循環，每個查詢循環有四個查詢步驟(各核苷酸一個)。對於使用中的感測器，各查詢步驟產生單次判定該

感測器是否在偵測標記。

【0086】 如本文所用，術語「偵測結果」係指指示以下之值：(a)在查詢步驟期間偵測到標記或(b)在查詢步驟期間沒有偵測到標記。在一些實施例中，偵測結果係二進制值(例如0或1)。偵測結果可自其他資料(例如表示電阻、頻率、強度等之信號；電阻、頻率、強度等之測量)得出。

【0087】 如本文所用，術語「記錄」係指單個感測器之偵測結果之儲存的表示。若所選擇的定序程序具有M個查詢步驟，則在定序程序完成後，各記錄具有M個偵測結果。可將S個感測器之記錄儲存在單個文件中(例如以具有S個行及M個列、或S個列及M個行之表)，或可針對各別感測器的記錄建立單獨文件。

【0088】 如本文所使用，關於包含在記錄中之偵測結果，術語「串」意指連續相同值序列。

【0089】 術語「感測器」及「感測元件」在本文中可互換使用。

【0090】 該變數S在本文中用於指複數個感測器中之感測器之數量。該等S個感測器可為感測相同股之例項，或其可為感測不同股之例項。

【0091】 該變數K在本文中用於指複數個感測器中全部感測相同股之例項之感測器之數量。

【0092】

標記

用於本文描述的核酸定序之方法使用包含可切割之標記之經標記之核苷酸前驅物。此等可切割之標記可為例如磁性、螢光、有機金屬或帶電荷分子。

【0093】 各標記可包含例如磁性奈米粒子，諸如例如分子、超順磁性奈米粒子或鐵磁性粒子。磁標記可為具有高磁性各向異性之奈米粒子。具有高磁性各向異性之奈米粒子之實例包括但不限於Fe_3O_4、FePt、FePd及CoPt。為了促進對核苷酸之化學結合，該等粒子可合成且經SiO_2塗佈。參見，例如M. Aslam、L. Fu、S. Li及V.P. Dravid，「Silica encapsulation and magnetic properties of FePt nanoparticles」，Journal of Colloid and Interface Science，第290卷，第2期，2005年10月15日，第444至449頁。因為此種尺寸之磁標記具有永久磁矩，其方向在極短時間標度上隨機波動，故以下進一步描述之一些實施例仰賴於敏感感測方案，該等方案偵測由於磁標記之存在引起之磁場之波動。

【0094】 各標記可包含例如螢光團。螢光標記係此項技術中熟知的且適合與本文揭示內容一起使用。

【0095】 標記可包含例如有機金屬化合物。如所瞭解，有機金屬化合物為包含至少一個金屬-對-碳鍵(其中該碳為有機基團之一部分)之一類物質之任何成員。有機金屬化合物之實例包括吉爾曼試劑(Gilman reagent) (其包含鋰及銅)、格林納試劑(Grinard reagent) (其包含鎂)、四羰基鎳及二茂鐵(其包含過渡金屬)、有機鋰化合物(例如正丁基鋰(n-BuLi))、有機鋅化合物(例如二乙基鋅 (Et_2Zn))、有機錫化合物(例如氫化三丁基錫(Bu_3SnH))、有機硼烷化合物(例如三乙基硼烷(Et_3B))及有機鋁化合物(例如三甲基鋁(Me_3Al))。

【0096】 標記可包含例如帶電荷分子。

【0097】 有多種方法可將標記附接至核苷酸前驅物且在併入核苷酸前驅物之後切割標記。例如，可將標記附接至鹼基，在該情況下，其可以

化學方式切割。作為另一個實例，可將標記附接至磷酸酯，在該情況下，其可藉由聚合酶切割，或若經連接子附接，則藉由切割連接子進行切割。

【0098】 在一些實施例中，將標記連接至核苷酸前驅物之含氮鹼基(例如A、C、T、G或衍生物)。在併入核苷酸前驅物且藉由(例如如以下進一步詳細描述之)定序裝置偵測之後，自所併入的核苷酸切割標記。

【0099】 在一些實施例中，標記係經由可切割之連接子附接。可切割之連接子係此項技術中已知的且已描述於例如美國專利第7,057,026號、第7,414,116號及其連續案及改進案中。在一些實施例中，標記係經包含烯丙基或疊氮基之連接子附接至嘧啶中之5位置或嘌呤中之7位置。在其他實施例中，連接子包含二硫鍵、吲哚或Sieber基團。該連接子可進一步包含一或多個選自烷基(C_{1-6})或烷氧基(C_{1-6})、硝基、氰基、氟基團或具有類似性質之基團之取代基。簡言之，該連接子可藉由水溶性膦或含膦基過渡金屬之觸媒切割。其他連接子及連接子切割機制係此項技術中已知的。例如，包含三苯甲基、對烷氧基苄基酯及對烷氧基苄基醯胺及第三丁氧基羰基(Boc)基團及縮醛系統之連接子可在酸性條件下藉由質子釋放切割劑切割。硫縮醛或其他含硫連接子可使用親硫金屬(諸如鎳、銀或汞)切割。可切割之保護基亦可考慮用於製備適宜連接子分子。含酯連接子及含二硫鍵連接子可在還原條件下切割。含有三異丙基矽烷(TIPS)或第三丁基二甲基矽烷(TBDMS)之連接子可在F離子之存在下切割。可藉由不影響反應混合物之其他組分之波長切割之可光切割之連接子包括包含O-硝基苄基基團之連接子。包含苄基氧基羰基基團之連接子可藉由基於Pd之觸媒切割。

【0100】 在一些實施例中，核苷酸前驅物包含附接至磷酸酯部分之

標記，如例如美國專利第7,405,281號及第8,058,031號中所述。簡言之，核苷酸前驅物包含核苷部分及3個或更多個磷酸酯基團之鏈，其中該等氧原子中之一者或多者係視需要經例如S取代。該標記可直接或經連接子附接至α、β、γ或更高之磷酸酯基團(若存在的話)。在一些實施例中，標記係經非共價連接子附接至磷酸酯基團，如例如美國專利第8,252,910號中所述。在一些實施例中，該連接子為選自以下之烴：經取代或未經取代之烷基、經取代或未經取代之雜烷基、經取代或未經取代之芳基、經取代或未經取代之雜芳基、經取代或未經取代之環烷基及經取代或未經取代之雜環烷基；參見，例如美國專利第8,367,813號。該連接子亦可包含核酸股；參見，例如美國專利第9,464,107號。

【0101】 在其中將標記連接至磷酸酯基團之實施例中，核苷酸前驅物係藉由核酸聚合酶併入至新生鏈中，該核酸聚合酶亦切割且釋放可偵測之標記。在一些實施例中，標記係藉由切割連接子來移除，例如，如美國專利第9,587,275號中所述。

【0102】 在一些實施例中，核苷酸前驅物為不可延伸之「終止子」核苷酸，亦即3'端經阻斷「終止子」基團阻斷而無法添加下一個核苷酸之核苷酸。該等阻斷基團係可逆終止子，其可經移除以便繼續如本文所述的股合成過程。將可移除之阻斷基團附接至核苷酸前驅物係此項技術中已知的。參見，例如美國專利第7,541,444號、第8,071,739號及其連續案及改進案。簡言之，阻斷基團可包含烯丙基基團，其可藉由在水溶液中與金屬-烯丙基錯合物在膦或氮-膦配體之存在下反應來切割。用於合成定序中之可逆終止子核苷酸之其他實例包括經修飾之核苷酸，該經修飾之核苷酸描述於2019年12月16日申請且題為「3'-protected Nucleotides」之國際申請

案第PCT/US2019/066670號中，該申請案經公開為WO/2020/131759。

【0103】

感測器

用於本文描述的核酸定序裝置、系統及方法中之感測器之特性及能力取決於所使用的標記之選擇。感測器可為例如磁感測器(以偵測例如磁性奈米粒子、有機金屬化合物等)或光學感測器(以偵測例如螢光團)。應明瞭，其他類型之感測器可適合於偵測各種類型之標記，且本文描述的實例無意為限制性的。一般而言，所揭示的裝置、系統及方法可使用可藉由所選擇的類型之感測器偵測之任何類型之標記，且相反地，所揭示的裝置、系統及方法可使用可偵測所選擇的類型之標記之存在(及不存在)之任何類型之感測器。

【0104】 參考數字105在本文中一般用於單分子感測器，而與彼等單分子感測器之類型無關(且與其偵測的標記之類型無關)。參考數字15用於感測核酸股簇之感測器。

【0105】

磁感測器

本文揭示的一些實施例使用磁感測器來偵測偶聯至核苷酸前驅物之磁標記(例如磁性奈米粒子、有機金屬錯合物、帶電荷分子等)之存在。圖1說明根據一些實施例之磁感測器105之一部分。圖1之示例性磁感測器105具有底表面108及頂表面109且包含三個層，例如兩個鐵磁層106A、106B間隔非磁性間隔層107。非磁性間隔層107可為例如金屬性材料，諸如，例如銅或銀，在該情況下，該結構稱為自旋閥(SV)，或其可為絕緣體，諸如，例如氧化鋁或氧化鎂，在該情況下，該結構稱為磁穿隧接面

(MTJ)。用於鐵磁層106A、106B中之適宜材料包括例如Co、Ni及Fe (有時與其他元素混合)之合金。在一些實施例中，鐵磁層106A、106B係經工程化以使其磁矩定向於膜之平面中或垂直於膜之平面。另外材料可在顯示於圖1中之三個層106A、106B及107的下方及上方沉積以實現目的，諸如界面平滑、紋理化及保護免受用於圖案化其中併入感測器105之裝置之處理，但該磁感測器105之活性區域位於此三層結構中。因此，與磁感測器105接觸之組件可與三個層106A、106B或107中之一者接觸，或其可與磁感測器105之另一部分接觸。

【0106】 如圖2A及2B中所顯示，MR感測器之電阻係與1-cos(θ)成比例，其中θ為顯示於圖1中之兩個鐵磁層106A、106B之磁矩之間的角度。為了最大化由磁場產生之信號且提供磁感測器105對施加的磁場之線性反應，磁感測器105可設計成使得兩個鐵磁層106A、106B之磁矩定向為π/2弧度或在不存在磁場下相對於彼此成90度。此種定向可藉由此項技術中已知的許多方法來達成。例如，一種解決辦法係使用反鐵磁體透過稱為交換偏壓之作用來「固定」鐵磁層(106A或106B，指定為「FM1」)中之一者之磁化方向且然後用具有絕緣層及永久磁鐵之雙層塗覆感測器。絕緣層避免磁感測器105之電短路，且永久磁鐵提供垂直於FM1之固定方向之「硬偏壓」磁場，該磁場然後會使第二鐵磁體(106B或106A，指定為「FM2」)旋轉且產生所期組態。平行於FM1之磁場然後使FM2繞著此90度組態旋轉，且電阻變化導致電壓信號可經校準以測量作用於磁感測器105上之磁場。依此方式，磁感測器105用作磁場-與-電壓轉換器。

【0107】 應注意，儘管以上剛剛論述的實例描述其磁矩在膜平面中相對於彼此成90度定向之鐵磁體之用途，但垂直組態可替代地藉由將鐵磁

層106A、106B中之一者之磁矩定向於膜平面之外來達成，此定向可使用稱為垂直磁各向異性(PMA)者來達成。

【0108】 在一些實施例中，磁感測器105使用稱為自旋轉移轉矩之量子機械效應。在此種裝置中，通過SV或MTJ中之一個鐵磁層106A (或106B)之電流優先允許具有平行於層的磁矩之自轉之電子傳輸通過，而具有反平行自旋之電子更可能被反射。依此方式，電流變得自旋極化，其中一種自旋類型之電子比另一種自旋類型之電子更多。此種自旋極化之電流然後與第二鐵磁層106B (或106A)相互作用，從而於層的磁矩上施加轉矩。此轉矩可在不同情況下導致第二鐵磁層106B (或106A)之磁矩圍繞作用於鐵磁體時之有效磁場進動，或其可導致磁矩在藉由在系統中感應之單軸各向異性限定的兩種定向之間可逆地切換。所產生的自旋轉矩振盪器(STO)藉由改變作用於其時之磁場而係頻率可調諧的。因此，其具有充當磁場-與-頻率(或相位)轉換器(藉此產生具有頻率之AC信號)之能力，如圖3A中所顯示，其說明使用STO感測器之概念。圖3B顯示當跨STO施加具有1 GHz之頻率及5 mT之峰-與-峰幅度之AC磁場時透過延遲偵測電路之STO之實驗反應。此結果及彼等顯示於短奈秒場脈衝之圖3C及3D中之結果說明此等振盪器可如何用作奈米級磁場偵測器。進一步之細節可見於T. Nagasawa、H. Suto、K. Kudo、T. Yang、K. Mizushima及R. Sato，「Delay detection of frequency modulation signal from a spin-torque oscillator under a nanosecond-pulsed magnetic field」，Journal of Applied Physics，第111卷，07C908 (2012)中。

【0109】

光學感測器

一些核酸定序方法使用螢光標記。在此種方法中，將所定序的核酸分子固定於固體支撐物上，且監測螢光標記之靶分子(例如核苷酸)與分子之結合。光學儀器(例如用於螢光之激發及讀取裝置)提供某一波長的光以激發螢光標記且自標記偵測以稍微不同的波長發射之螢光。因為激發光之光束路徑(光路徑)必須至少部分不同於螢光之光束路徑(光路徑)，故可使用激發及發射過濾器(其光譜沒有明顯重疊)來達成光譜分離，及/或可使用垂直或側面照明。

【0110】 此項技術中熟知使用螢光標記(例如螢光團)之光學感測器及定序裝置及方法。

【0111】

擴增/複製

核酸定序裝置一般仰賴於擴增(或複製)過程以自單個核酸股產生大量核酸例項(例如來自一個DNA分子之單側DNA股(ssDNA)之例項)。聚合酶鏈反應(PCR)係一種用於擴增雙股DNA之熟知方法，該方法使得能夠自少量初始量複製大量DNA。

【0112】

簇定序裝置

一些定序裝置(在本文中稱為簇(CLUS)裝置)使用擴增技術來形成許多DNA股之局部簇。例如，一條DNA股用作模板，且PCR擴增在局部區域中產生數千或數百萬個DNA序列例項。將PCR引物之至少一部分固定至固體支撐物，該固體支撐物允許產生的DNA分子固定至局部簇以便形成可區分之「純系」。產生的DNA簇可包含ssDNA。選殖擴增技術之實例包括橋式PCR及乳液PCR，包括基於微珠之乳液PCR。對於橋式擴增，使

用附接至固體表面(諸如載玻片)之引物藉由原位PCR擴增單個DNA分子以形成DNA簇。各DNA簇係由DNA股之例項組成之物理分離之「純系」。對於基於乳液PCR之選殖擴增，單個DNA分子係在乳液液滴中選殖擴增。在一些方法中，將DNA股附接至液滴內部的微珠。單分子之選殖擴增亦可在單獨微孔中進行。

【0113】 如本文所用，術語「簇」係指理想地具有相同序列之核酸股之局部簇，其係藉由選殖擴增產生。當核酸為DNA時，簇包含(理想地)附接至固體支持物之相同DNA股(或片段)。例如，簇可在載玻片之斑點上產生或附接至微珠、微孔或其他微粒。

【0114】 CLUS裝置用於基於螢光之DNA定序之用途係熟知的。

【0115】 使用磁感測器陣列之用於使用簇之核酸定序之定序裝置描述於例如2021年3月7日申請且題為「MAGNETIC SENSOR ARRAYS FOR NUCLEIC ACID SEQUENCING AND METHODS OF MAKING AND USING THEM」 (代理人檔案編號ROA-1001-WO/P35967-WO)之PCT申請案第PCT/US2021/021274號中。

【0116】 圖4A說明CLUS裝置之單個感測器15，該CLUS裝置用於感測在其附近的一些*N*個選殖擴增之DNA股101。感測器15可為例如磁感測器以感測附接至併入的核苷酸之磁標記。為了方便起見，圖4A顯示與感測器15接觸之股101，但應明瞭在感測器15與股100之間可存在障壁(例如絕緣層)。感測器15可為例如如上文引用的PCT申請案第PCT/US2021/021274號中所述的磁感測器。

【0117】 當前最先進技術商業CLUS裝置(諸如彼等感測螢光標記者)可使用數億個感測器15，各感測器感測各別擴增之DNA股101之許多例

項。一些CLUS裝置之一個缺點係達成最佳簇密度對於高品質定序可為至關重要。具體而言，使用大簇傾向於提供較高資料品質，但降低資料輸出，而使用小簇可導致運行失敗，運行性能差，Q30分數較低，引入定序工件，及降低總資料輸出。為減輕此等問題，較新的CLUS裝置使用圖案化流動池，該流動池具有不同奈米孔以用於簇產生。此等奈米孔經組構成六邊形配置以可更有效地使用流動池表面積。

【0118】

單分子陣列定序裝置

單分子陣列定序裝置(本文稱為「SMAS裝置」)係CLUS裝置之替代品。與感測且定序單個核酸股之多個例項之局部簇之CLUS裝置相反，SMAS裝置使用個別地感測及定序核酸之個別股之感測器。一般而言，在SMAS裝置中，沒有感測器會感測超過一個物理核酸股，但不同感測器感測相同股之例項。換言之，存在核酸股之多個例項，但各感測到的股藉由不同各別感測器來感測。取決於所使用的擴增技術，該等個別股可隨機分佈於SMAS裝置之流體腔室中，或其可位於更多局部區域中。如以下進步論述，可識別特定股之例項之位置，且可在識別鹼基之前將錯誤校正程序應用於對應於該等例項之偵測結果以相對於CLUS裝置改良定序之準確度。除此之外，相對於CLUS裝置，對於合理的化學失敗率，SMAS裝置需要待定序的各核酸股之更少例項以達成準確定序結果。

【0119】 圖4B說明示例性的複數個*S*個單分子感測器105，各感測器藉由SMAS裝置用於監測各別單股DNA (ssDNA) 101。該複數個S個感測器105中之各者可為例如磁感測器、光學感測器等。圖4B說明五個單分子感測器105A、105B、105C、105D及105E，其各者感測各別DNA股101

(其可為相同DNA股之例項，或不同DNA股之例項)。各感測器105可為例如奈米級感測器，其係如此之小以致於僅單個DNA股101可結合至與感測器105相關之結合位點。(為了方便起見，圖4B顯示與感測器105接觸之股101，但如下文進一步說明，在一些實施例中，將股100附接至個別結合位點，該等結合位點各者與各別感測器105相關聯。)

【0120】 考慮結合至包含密集封裝之感測器105陣列之固體表面之選殖擴增之DNA，如圖4B中所顯示。DNA可藉由固相擴增(SPA)複製以建立單株DNA簇，各股意欲藉由不同感測器105感測，或DNA可大量擴增且然後固定於SMAS裝置之表面上。若DNA在SMAS裝置之流體腔室之表面上擴增(例如藉由SPA)，則感測器105A、105B、105C、105D、105E可感測選殖DNA之例項。或者，若DNA在裝置外大量擴增且添加至SMAS裝置的流體腔室，則經擴增之DNA股101可更隨機地分佈於感測器105中。

【0121】 圖5A為顯示根據一些實施例之用於核酸定序之示例性SMAS裝置100之組件之方塊圖。如所說明，裝置100包括感測器陣列110，其係耦合至電路120，該電路係耦合至至少一個處理器130。感測器陣列110包括複數個感測器105 (例如磁感測器、光學感測器等)，其可以任何適宜方式配置，如以下進一步描述。感測器陣列110中之感測器105之特性及性質取決於用於定序之標記之類型。

【0122】 電路120可包括例如一或多條線，其允許感測器陣列110中之感測器105藉由至少一個處理器130訊問(例如藉助於此項技術中熟知的其他組件，諸如電流源等)。例如，在操作中，處理器130可導致電路120施加電流至此類線以偵測感測器陣列110中之複數個感測器105中之至少

一者之特性，其中該特性指示在感測器105之範圍內存在標記或不存在任何標記。換言之，特性(例如電阻、頻率、電壓、信號位準等)指示感測器105已偵測到至少一個標記或尚未偵測到任何標記。例如，該至少一個處理器130可評定特性(例如頻率、波長、磁場、電阻、雜訊位準、強度、光之顏色等)之值且基於特性值與臨限值之比較(例如藉由判定感測器105之特性值是否滿足或超過臨限值)或基線值來判定偵測到(或未偵測到)標記。作為另一個實例，該至少一個處理器130可比較所獲得的感測器105之特性與先前偵測到的特性值(例如感測器105之基線值)且基於特性值之變化(例如磁場、電阻、雜訊位準、頻率、波長、強度、光之顏色等之變化)偵測到或未偵測到標記之判定。例如，如以下在圖19之論述中進一步描述，該至少一個處理器130可評估自感測器105獲得之特性以偵測在定序程序之第一查詢步驟期間偵測到標記之感測器105在應已移除標記之切割步驟之後是否仍偵測到該標記。類似地，該至少一個處理器130可評估自一個查詢步驟至下一個查詢步驟之特性變化以判定感測器105 (a)在任一查詢步驟期間未偵測到標記，(b)在兩個查詢步驟期間偵測到標記，(c)在第一查詢步驟期間未偵測到標記但在隨後查詢步驟期間偵測到標記，及/或(d)在第一查詢步驟期間偵測到標記但在隨後查詢步驟期間未偵測到標記。

【0123】 偵測到的特性取決於用於定序程序中之標記之類型。該等標記可為例如螢光，在該情況下，感測器105可為可偵測例如由螢光標記發出的光之波長、頻率、調變頻率、顏色或強度之光學偵測器。適用於偵測螢光標記之光學感測器係此項技術中熟知。在用於核酸定序程序中之標記為螢光之情況下，在一些實施例中，電路120允許至少一個處理器130

偵測藉由感測器陣列110中之一些或全部感測器105偵測到的光(或電磁能)之偏差或波動。

【0124】該等標記可為例如磁性(例如磁性奈米粒子、有機金屬化合物、帶電荷分子等)，在該情況下，感測器105可為可偵測磁性特性之磁感測器。磁感測器已描述於申請人的先前申請之專利申請案中，包括例如2020年4月8日申請之題為「NUCLEIC ACID SEQUENCING BY SYNTHESIS USING MAGNETIC SENSOR ARRAYS」(代理人檔案編號ROA-1000-WO / P35097-WO)且於2020年10月15日以WO 2020/210370公開之PCT申請案第PCT/US20/27290號。在其中標記係磁性之一些實施例中，感測器105為可偵測例如磁場或電阻、磁場之變化或電阻之變化、或雜訊位準之磁阻(MR)感測器。在一些實施例中，感測器陣列110之感測器105中之各者為薄膜裝置，其使用MR效應以偵測附接至併入與各別結合位點結合之核酸之單股中之核苷酸之磁標記。感測器105可用作電阻隨感測到的磁場之強度及/或方向變化而變化之電位計。在使用磁標記之一些實施例中，感測器105包含磁性振盪器(例如自旋轉矩振盪器(STO))，且指示是否偵測到至少一個標記之特性係與磁性振盪器相關聯或藉由磁性振盪器產生之信號之頻率或信號之頻率之變化。

【0125】在用於核酸定序程序中之標記係磁性之情況下，在一些實施例中，該至少一個處理器130在電路120之幫助下偵測感測器陣列110中之一些或全部感測器105之磁性環境中之偏差或波動。例如，與存在磁標記之感測器105相比，不存在磁標記之MR類型之感測器105應在一定頻率以上具有相對小的雜訊，因為來自磁標記之場波動將導致感測鐵磁體之磁矩之波動。此等波動可使用外差偵測(例如藉由測定雜訊功率密度)或藉由

直接測定感測器105之電壓測定且使用比較器電路以比較與不感測結合位點之另一個感測器元件加以評定。在感測器105包括STO元件之情況下，由於頻率之瞬時變化，來自磁標記之波動磁場將導致感測器105之相位躍變，此可使用相位偵測電路來偵測。另一個選項係將STO設計成使得其僅在小磁場範圍內振盪，因此磁標記之存在將關閉振盪。

【0126】 應瞭解，以上提供的標記及感測器105之實例僅係示例性。一般而言，可標記核苷酸前驅物之任何類型之標記可與可偵測該類型之標記之任何類型之感測器105之陣列110一起使用。

【0127】 圖5B、5C及5D說明根據一些實施例之用於核酸定序之例示性SMAS裝置100之部分。示例性SMAS裝置100使用磁標記及磁感測器105。圖5B係裝置100之俯視圖。圖5C係在由圖5B中標記為「5C」之長虛線指示的位置處之橫截面視圖，及圖5D係在由圖5B中標記為「5D」之長虛線指示的位置處之橫截面視圖。

【0128】 顯示於圖5B、5C及5D中之示例性裝置100包含用於感測流體腔室115內的磁標記之感測器陣列110。感測器陣列110包括複數個磁感測器105，其中在圖5B之陣列110中顯示十六個感測器105。應明瞭，SMAS裝置100之實施案可包括許多感測器105 (例如數百、數千或數百萬個感測器105)。為了避免混淆附圖，在圖5B中僅標記感測器105中的七個，亦即感測器105A、105B、105C、105D、105E、105F及105G。如上文所說明，磁感測器105偵測磁標記之存在或不存在。換言之，磁感測器105中之各者偵測在其附近是否存在至少一個磁標記。

【0129】 現參考圖5C及5D結合圖5B，各感測器105繪示於裝置100之示例性實施例中，其具有圓柱體形狀。然而，應明瞭，一般而言，感測

器105可具有任何適宜形狀。例如，感測器105在三個維度上可為長方體。此外，不同感測器105可具有不同形狀(例如一些可為長方體及其他可為圓柱體等)。應明瞭，附圖僅係示例性。

【0130】 如圖5C及5D中所顯示，裝置100包括流體腔室115。流體腔室115包含複數個結合位點116 (例如S個結合位點116)。在一些實施例中，流體腔室115容納在核酸定序程序期間使用的流體(例如核苷酸前驅物及其他流體)。然而，應明瞭，其中流體腔室115不容納流體之實施例係經考慮且在本文揭示內容之範疇內。例如，結合位點116可配置於可移除之(或可移動之)部分(例如面板、板、載玻片(slide)等)上，可在已將核酸股附接至結合位點116之後將該部分浸入至試劑及其他流體中且然後放置成使得感測器105可偵測標記。因此，儘管流體腔室115之名稱顯示其容納流體，但並不需要流體腔室115容納流體。

【0131】 如圖5B、5C及5D中所顯示，感測器105中之各者係與各別結合位點116相關聯。(為了簡單起見，本文件一般藉由參考數字116指代結合位點。對個別結合位點給予參考數字116，後跟一個字母。)換言之，感測器105及結合位點116係處於一對一的關係。如圖15B中所顯示，感測器105A係與結合位點116A相關聯，感測器105B係與結合位點116B相關聯，感測器105C係與結合位點116C相關聯，感測器105D係與結合位點116D相關聯，感測器105E係與結合位點116E相關聯，感測器105F係與結合位點116F相關聯，及感測器105G係與結合位點116G相關聯。顯示於圖5B中之其他未經標記之感測器105中之各者亦與各別結合位點116相關聯。在圖5B、5C及5D之實例實施例中，顯示各感測器105配置於其各別結合位點116下方，但應明瞭結合位點116可相對於其各別感測器105處於

其他位置。例如，結合位點116可位於其各別感測器105的側面。

【0132】 結合位點116中之各者係經結構設計成使不超過一個核酸股(例如ssDNA)結合至SMAS裝置100之流體腔室115內。換言之，各結合位點116具有允許核酸之一個且僅一個股結合至其以用於藉由各別感測器105感測(且用於定序)之特性及/或特徵。此後，各別感測器105可在核酸定序程序期間偵測附接至併入與結合位點116結合之核酸股中之核苷酸之標記，如下文進一步論述。在一些實施例中，結合位點116具有經結構設計成錨定核酸至結合位點116之結構(或多個結構)。例如，該結構(或該等結構)可包括空腔(cavity)或脊。圖5C及5D將結合位點116說明為自流體腔室115之表面延伸，但應明瞭結合位點116可與流體腔室115之表面齊平或經蝕刻至流體腔室115之表面中。

【0133】 結合位點116可具有有利於將核酸之一個且僅一個股附接至各結合位點116之任何適宜尺寸及形狀。例如，結合位點之形狀可與感測器105之形狀類似或相同(例如若感測器105在三個維度上係圓柱體，則結合位點116亦可為圓柱體，自流體腔室115之表面突出或形成流體容器於流體腔室115之表面內，其半徑可為與各別感測器105之半徑相比更大、更小之尺寸或與之相同之尺寸；若感測器105在三個維度上係長方體，則結合位點116亦可為表面116與感測器105之最接近部分相比更大、更小之尺寸或與之相同之尺寸之長方體等)。一般而言，結合位點116及流體腔室115之表面可具有有利於單個核酸股附接至各結合位點116且允許感測器105偵測在其各別結合位點116處附接至併入的核苷酸之標記之任何形狀及特性。

【0134】 圖5C及5D說明具有在x-y平面中延伸之頂部部分之經封閉

流體腔室115，但不需要封閉流體腔室115。在一些實施例中，流體腔室115之表面具有保護感測器105不受流體腔室115中之任何流體影響，同時仍允許核酸股與結合位點116結合及允許感測器105以偵測附接至併入與結合位點116附接之核酸股中之核苷酸之標記之性質及特性。流體腔室115之材料(且可能係結合位點116之材料)可為絕緣體或包含絕緣體。在一些實施例中，流體腔室115之表面包含有機聚合物、金屬或矽酸鹽。流體腔室115可包括例如金屬氧化物、二氧化矽、聚丙烯、金、玻璃或矽。流體腔室115之表面之厚度可經選擇使得感測器105可偵測附接至併入與流體腔室115內的結合位點116結合之核酸股中之核苷酸之磁標記。在一些實施例中，該表面為約3至20 nm厚使得各感測器105介於距附接至併入與感測器105的相應結合位點116結合之核酸股中之核苷酸之任何標記約5 nm與約50 nm之間。應瞭解，此等值僅係示例性。應瞭解，實施案可具有具有較厚或較薄表面之流體腔室115。

【0135】 裝置100之電路120可包括一或多條線125。在一些實施例中，複數個感測器105中之各者係耦合至至少一條線125。在顯示於圖5B、5C及5D中之實例中，裝置100包括八條線125A、125B、125C、125D、125E、125F、125G及125H。(為了簡單起見，本文件一般藉由參考數字125指代線。對個別線給予參考數字125，後跟一個字母。)線125對可用於訪問(例如訊問)個別感測器105。在顯示於圖5B、5C及5D中之示例性實施例中，感測器陣列110之各感測器105係耦合至兩條線125。例如，感測器105A係耦合至線125A及125H；感測器105B係耦合至線125B及125H；感測器105C係耦合至線125C及125H；感測器105D係耦合至線125D及125H；感測器105E係耦合至線125D及125E；感測器105F係耦合

至線125D及125F；及感測器105G係耦合至線125D及125G。在圖5B、5C及5D之示例性實施例中，顯示線125A、125B、125C及125D位於磁感測器105下方，及顯示線125E、125F、125G及125H位於磁感測器105上方。圖5C顯示關於線125D及125E之感測器105E、關於線125D及125F之感測器105F、關於線125D及125G、及關於線125D及125H之感測器105D。圖5D顯示關於線125D及125H之感測器105D、關於線125C及125H之感測器105C、關於線125B及125H之感測器105B、及關於線125A及125H之感測器105A。

【0136】圖5B、5C及5D之示例性SMAS裝置100之感測器105係經配置在矩形圖案感測器陣列110中。(應明瞭，正方形圖案係矩形圖案之特殊情況。)線125中之各者標識感測器陣列110之行或列。例如，線125A、125B、125C及125D中之各者標識感測器陣列110之不同行，及線125E、125F、125G及125H中之各者標識感測器陣列110之不同列。如圖5C中所顯示，線125E、125F、125G及125H中之各者係順著橫截面與感測器105中之一者接觸(亦即線125E係與感測器105E之頂部接觸，線125F係與感測器105F之頂部接觸，線125G係與感測器105G之頂部接觸，及線125H係與感測器105D之頂部接觸)，及線125D係與感測器105E、105F、105G及105D中之各者之底部接觸。類似地，且如圖5D中所顯示，線125A、125B、125C及125D中之各者係順著橫截面與感測器105中之一者之底部接觸(亦即線125A係與感測器105A之底部接觸，線125B係與感測器105B之底部接觸，線125C係與感測器105C之底部接觸，及線125D係與感測器105D之底部接觸)，及線125H與感測器105D、105C、105B及105A中之各者之頂部接觸。

【0137】 圖5B中使用虛線繪示感測器105及連接至感測器陣列110之線125之部分以指示可將其嵌入裝置100內。如上文所說明，感測器105可經保護(例如藉由絕緣體)不受流體腔室115之內含物影響，該流體腔室115本身可進行封閉。因此，應瞭解，各種所說明的組件(例如線125、感測器105、結合位點116等)在裝置100之物理實例化中不一定係可見的(例如其可經嵌入保護材料諸如絕緣體中或藉由保護材料諸如絕緣體覆蓋)。

【0138】 在一些實施例中，一些或全部結合位點116駐留在穿過感測器105之線125中之奈米孔或溝槽中。例如，如圖5D之實例中所顯示，線125H可於感測器105上比其在感測器105之間更細。例如，線125H具有於感測器105D上方之第一厚度、在感測器105D及105C之間的第二較大厚度及在感測器105C上方之第一厚度。此一組態可使用習知薄膜製造方法(例如藉由沉積材料，施加遮罩至所沉積的材料，及根據遮罩移除(例如藉由蝕刻)一些所沉積的材料)來有利地製造。結合位點116及(若存在的話)奈米孔均可使用習知技術來製造。

【0139】 為了簡化說明，圖5B、5C及5D說明示例性裝置100，其具有在感測器陣列110中之僅十六個感測器105、僅十六個各別結合位點116及八條線125。應明瞭，裝置100可具有在感測器陣列110中之更少或更多個感測器105，且因此，且可具有更多個或更少個結合位點116。類似地，包括線125之實施例可具有更多或更少條線125。一般而言，可使用感測器105及結合位點116之任何組態，其允許感測器105偵測附接至併入與結合位點116附接之單個核酸股中之核苷酸之標記。類似地，可使用一或多條線125或允許判定感測器105是否已感測到一或多個標記之某種其他機制之任何組態。本文呈現的實例無意為限制性的。

【0140】 如以上所說明，顯示於圖5B、5C及5D中之感測器105可為磁感測器105。因此，感測器105緊鄰結合位點116，且因此，其亦緊鄰結合至結合位點116之核酸股。應瞭解，感測器陣列110相對於結合位點116之適宜位置部分取決於所使用的標記之類型，及因此取決於所使用的感測器105之類型。例如，若標記為螢光團，且感測器105為光學感測器，則感測器陣列110遠離結合位點116 (例如位於結合位點116上方)可能係適宜的。

【0141】 儘管圖5B、5C及5D(及本文其他附圖)以一對一的關係說明感測器105及結合位點116，但應明瞭，各結合位點116可藉由超過一個感測器105感測。將SMAS裝置100與CLUS裝置區分開之特性係SMAS裝置100之感測器105沒有感測到超過一個核酸股例項。若SMAS裝置100具有比結合位點116更多的感測器105，則藉由多個感測器105感測至少一些核酸股 (例如以改良標記偵測之準確度)可為可行的。

【0142】 顯示且描述於圖5B、5C及5D之內文中之示例性感測器陣列110為矩形陣列，其中感測器105以行及列配置。換言之，感測器陣列110之該複數個感測器105以矩形格網圖案配置。在一些實施例中，矩形格網圖案之相鄰行及列係彼此等距，此導致感測器105以正方形格網(或格子)圖案配置，如圖5E中所說明。在其中感測器105以正方形網格圖案配置之實施例中，各感測器105具有多達四個最近相鄰者。例如，如圖5E中所顯示，感測器105A具有四個最近相鄰者，標記為105B、105C、105D及105E。如圖5E中所顯示，最接近的感測器105相距最近相鄰距離112。因此，感測器105B、105C、105D及105E中之各者與感測器105A相距距離112。

【0143】 商業上可行的SMAS裝置100可使用能夠識別個別標記之密集封裝之奈米級感測器105之高精度奈米級製造。官能化結合位點116之尺寸可類似於例如附接標記之DNA之尺寸，使得多個股不能結合至相同結合位點116或不能被相同感測器105感測到。用於評估定序儀的商業競爭力之公認指標係將DNA股一起封裝於流體腔室115中之密度程度。

【0144】 可基於感測器105之性質、裝置100意欲定序的核酸股之長度及所使用的標記之性質來確定最近相鄰距離112之適宜值，然後可將該適宜值用於確定SMAS裝置100之尺寸及/或可裝配於選定尺寸的SMAS裝置100內的感測器105之最大數量。例如，核酸股之組合長度及待使用的標記之尺寸可針對SMAS裝置100中之兩個感測器105可進行定位的接近程度提供物理限制。在一些實施例中，感測器105之尺寸可受到用於製造SMAS裝置100之製程之奈米級圖案化能力的限制。例如，使用在寫入之時可用的技術，各磁感測器105 (例如假設是圓柱體感測器105，則是感測器105在x-y平面內的直徑)之尺寸可為約20 nm。假設待定序的核酸之類型為DNA，且期望定序長度多達150個鹼基對(bp)之片段，則在長形狀態下待定序的DNA股101之最大長度為約50 nm，儘管ssDNA構形可在長形及螺旋狀之間變化，如圖6A中所顯示，取決於緩衝液之離子強度。因為標記102參與單分子反應，故標記102應具有分子尺寸。對於使用磁感測器105之SMAS裝置100，標記102可為例如超順磁性奈米粒子、有機金屬化合物、或可藉由奈米級磁感測器105偵測之任何其他官能分子基團。因此，假設各標記102具有不大於約10 nm之尺寸。在此等假設下，圖6B顯示磁感測器105、處於長形狀態下之DNA股101及磁標記102之相對尺寸。

【0145】 使用磁感測器105以偵測用作標記102之磁奈米粒子之實際

SMAS裝置100可使用現有技術來實施。為了論證起見，假設僅偵測到感測器105之邊緣的20 nm以內的標記102。各感測器105之偵測範圍小，因為可針對於核酸定序應用而選擇的磁標記102 (例如超順磁性奈米粒子、有機金屬化合物等)不會對偵測到的磁場產生明顯擾動。儘管附接至併入與特定感測器105的結合位點116結合之ssDNA中之核苷酸之標記102可暫時駐留在各別感測器105之範圍之外，因為ssDNA在偵測過程期間假設各種構形狀態，但期望標記在ssDNA假設其完全長形狀態時不允許到達相鄰感測器105之敏感空間(偵測區域)。

【0146】 可得出實際SMAS裝置100之感測器封裝極限，例如假設標記為超順磁性奈米粒子(例如氧化鐵、鐵鉑等)，及SMAS裝置100之感測器陣列110為類似於彼等用於非易失性資料儲存應用中者之磁穿隧接面(MTJ)之矩形(例如正方形)陣列。在此種情況下，各奈米級感測器105之區域或其緊鄰處可經功能化以充當各別結合位點116。用於估計SMAS裝置100之感測器陣列封裝極限之簡單幾何配置顯示於圖7A中，其顯示兩個感測器105A、105B。假設各感測器105A、105B (僅為了方便起見而假設具有圓柱體形狀)具有約20 nm之直徑(如上文所說明)且假設能夠偵測自其邊緣20 nm以內的任何標記。感測區域邊界111以顯示於圖7A中之內部虛線表示。感測器105A感測與其結合位點結合之DNA股101A，及感測器105B感測與其結合位點結合之DNA股101B。當附著至併入股101A、101B中之核苷酸時，標記102A、102B之最大到達範圍(maximum reaches)(例如當具有150個鹼基之DNA股處於其完全非螺旋狀態下時)以外部點劃線圓圈103顯示。為了使定序結果準確，期望各感測器105僅偵測附接至併入與感測器105的各別結合位點116結合之DNA股101中之核苷

酸之標記102。因此，在上文描述的假設下，感測器105之間的避免串擾(例如附接至併入與另一個感測器105的結合位點116結合之核酸股101中之核苷酸之偵測標記102)之最小最近相鄰距離112為約100 nm。

【0147】 在SMAS裝置100之一些實施例中，感測器105 (例如MTJ)係以與現有交叉點MRAM感測器幾何形狀相容之正方形格子配置，如圖7B中所顯示。單位格子114之面積為10^4 nm^2，此允許各DNA股101延伸穿過約10^4 nm^2之面積，此產生SMAS裝置之DNA表面密度為約10^{10}個股/cm^2。假設在感測器陣列110中使用各個股101之至少十個例項，可同時定序約10^9個獨特股/cm^2，產生150 Gbase (10億×150 bp DNA股長度)資訊/平方厘米感測器陣列110。在理想情況下(例如，當化學失敗率很低時，僅需要三個DNA例項，如下文進一步論述)，可同時定序約3.3×10^9個不同股/cm^2，且每平方厘米感測器陣列110可產生約500 Gbase資料。

【0148】 作為一個特定實例，具有類似於2016年在國際電子裝置會議(the International Electron Devices Meeting；IEDM)首次引入的單個Toshiba 4 Gbit密度STT-MRAM晶片之構形之SMAS裝置100可潛在地產生約600 Gbase高品質資料。Toshiba平臺之感測器105之間的最小距離112為90 nm，該最小距離僅略低於以上得出的100 nm之估算的最小距離112。因此，使用類似於Toshiba平臺之組態之串擾甚至就150個鹼基長度的ssDNA而言仍可能很低，但可定序較短片段以甚至進一步減少串擾。

【0149】 應瞭解，感測器105以網格圖案(例如如顯示於圖7B中之正方形格子)之配置係許多可能配置中之一者。一般技術者應瞭解，感測器105之其他配置係可能的且在本文揭示內容之範疇內。例如，感測器105可以六邊形圖案配置，如圖8A中所顯示，其顯示SMAS裝置100之俯視

圖。顯示於圖8A中之示例性SMAS裝置100包含感測器陣列110，其用於感測流體腔室115內的標記102。感測器陣列110包括複數個感測器105，其中顯示十六個感測器105。應明瞭，裝置100之實施案可包括任何數量的感測器105 (例如數百、數千、數百萬等)。為了避免混淆附圖，在圖8A中僅標記感測器105中的兩個，亦即感測器105A及105B。如上文所說明，感測器105可為例如磁感測器(例如以偵測磁性或磁奈米粒子之效應)。如上文至少在圖5B、5C及5D之論述中所說明，一般而言，感測器105可具有任何適宜尺寸及形狀。

【0150】 如圖8A中所顯示，感測器105中之各者係與各別結合位點116相關聯。換言之，感測器105及結合位點116係處於一對一的關係。如圖8A中所顯示，感測器105A係與結合位點116A相關聯，感測器105B係與結合位點116B相關聯，及其他未標記之感測器105中之各者亦係與各別結合位點116相關聯。在圖8A之實例實施例中，顯示各感測器105配置於其各別結合位點116下方，但應明瞭結合位點116可相對於其各別感測器105處於其他位置。例如，結合位點116可位於其各別感測器105的側面。在至少圖5B、5C及5D之說明中對結合位點116之論述適用於圖8A及顯示結合位點116之其他圖式且在此不予以重複。

【0151】 圖8A之示例性SMAS裝置100亦包括上文描述於圖5B、5C及5D中之論述中之流體腔室115。彼等描述亦適用於圖8A且在此不予以重複。

【0152】 圖8A之裝置100之電路120可包括一或多條線125。圖8A之示例性實施例中之線125中之各者標識感測器陣列110之行或對角列。例如，線125A、125B、125C及125D中之各者標識感測器陣列110之不同

行，及線125E、125F、125G及125H中之各者標識感測器陣列110之不同對角列。在顯示於圖8A中之實例中，裝置100具有八條線125A、125B、125C、125D、125E、125F、125G及125H，且線125對可用於訪問個別感測器105。例如，線125A及125H可用於訪問感測器105A，及線125B及125H可用於訪問感測器105B。線125可定向於感測器105下方及/或之上，如圖5B、5C及5D等之論述中所描述。

【0153】 儘管圖8A說明示例性裝置100，其具有在感測器陣列110中僅十六個感測器105、僅十六個相應結合位點116及八條線125，但應瞭解該SMAS裝置100可在感測器陣列110中具有更少或更多個感測器105，且因此，其可具有更多或更少個結合位點116。此外，SMAS裝置100可具有更多或更少條線125。一般而言，可使用感測器105及結合位點116之任何組態，其允許感測器105偵測附接至併入與結合位點116附接之單個核酸股中之核苷酸之標記。類似地，可使用一或多條線125或允許判定感測器105是否已感測到一或多個標記之某種其他機制之任何組態。

【0154】 如圖8B中所顯示，當感測器105以六邊形圖案配置時，各感測器105具有至多六個最近相鄰者，全部在最近相鄰距離112處。換言之，各感測器105與最接近其的六個其他感測器105中之各者相距最近相鄰距離112。例如，如圖8B中所顯示，在附圖中間的未標記之感測器105具有六個最近相鄰感測器105，標記為105A、105B、105C、105D、105E及105F，其等均相距最近相鄰距離112。

【0155】 可得出使用光學感測器及螢光標記102 (例如螢光團)且具有結合位點116之六邊形圖案之SMAS裝置100之結合位點116封裝極限。假設標記102為螢光團，結合位點116呈六邊形圖案，且感測器陣列110遠

離結合位點116，則來自標記102之單分子螢光可投射至遠場中，在此處其可藉由包含光敏性感測器105之感測器陣列110偵測到。可使用單分子超解析成像技術(諸如彼等描述於C.G. Galbraith及J.A. Galbraith，「Super-resolution microscopy at a glance」，Journal of Cell Science，第124(10)卷，1607-11 (2011)中者)以解析個別螢光團標記102在SMAS裝置100中之位置。因為DNA封裝尺寸遠低於繞射極限，故可解析螢光團標記102之位置。儘管此種類型之偵測可能稍微複雜及/或昂貴，但最近已在商業定序系統中引入該技術以改良基於簇之定序儀之通量。此外，該技術可在不久的將來在大型單分子陣列之成像中實施。

【0156】 用於估計在使用螢光團標記102之SMAS裝置100中位於六邊形圖案中之結合位點116之封裝極限之簡單幾何配置顯示於圖9A中。DNA股101A係與結合位點116A結合，及DNA股101B係與結合位點116B結合。(感測器105未繪示於圖9A中，因為假設感測器陣列110遠離結合位點。)標記102A、102B之最大到達範圍(例如當具有150個鹼基之DNA股處於其完全非螺旋狀態時) (在附接至併入的核苷酸時)由點劃線圓圈103表示。為了避免串擾，在成像過程期間不允許附接至相鄰結合位點116之螢光團標記102佔據重疊空間，例如，應不允許附接至特定結合位點116A之螢光團標記102A到達ssDNA 101A探索其允許的構形狀態時附接至相鄰結合位點116B之螢光團標記102B可接近之空間。此種限制亦有助於避免螢光淬滅。假設使用螢光團標記102，則結合位點116可密集地封裝於六邊形格子中，如圖9B中所顯示。假設150 bp DNA股101之最大長度為50 nm，則螢光團標記102之大小為10 nm，自各結合位點116之中心至其邊緣之最小距離為20 nm，且各DNA股101結合至其各別結合位點116之中

心，該最小距離112為140 nm。因此，如圖9B中所顯示，允許每個DNA股101佔據具有1.7×10^4 nm^2之面積之單位格子114，此產生5.9×10^9個股/cm^2，或若SMAS裝置100中存在各DNA股之約10個例項，則為5.9×10^8個獨特股/cm^2之DNA表面密度。SMAS裝置100將自感測器陣列110的每平方厘米產生約90 Gbase資料。在最佳情境下，當僅需要3個DNA複製物時，感測器陣列110保持約2×10^9個獨特DNA股/cm^2，且SMAS裝置100能夠自感測器陣列110的每平方厘米產生約300 Gb資料。

【0157】 上文對六邊形陣列之論述係在螢光團標記102及光學感測器105之背景下進行。亦可使用磁感測器105之六邊形配置。可如上文在圖7A及7B之論述中所述得出具有結合位點116及磁感測器105之六邊形配置之SMAS裝置100之感測器封裝極限。對於磁感測器105，最近相鄰距離112為約100 nm，其意指(六邊形)單位格子面積114 (參見圖9B)為約8.7×10^3 nm^2。

【0158】 圖10比較描述於圖7A及7B (磁標記102及磁感測器105)及圖9A及9B (螢光標記102及光學感測器105)之內文中之SMAS實施案之密度與當前最先進技術CLUS定序儀之密度。為了論證起見，假設圖案化流通池之奈米孔陣列之間距為約500 nm。如圖10之左手側小圖中所顯示，CLUS定序儀之奈米孔係以具有500 nm格子常數之六邊形格子配置。各奈米孔保持在約50個至約200個相同DNA股(例如藉由固相橋擴增產生)。圖10之上右手側顯示使用螢光團標記及超解析成像之六邊形SMAS格子(例如如圖9A及9B之內文中所述)，及圖10之下右手側顯示使用超順磁性奈米粒子標記及MTJ之感測器陣列110之正方形SMAS格子(例如如圖7A及7B之內文中所述)。圖10中之三個表示按比例調整以顯示SMAS格子組態與

CLUS組態相較的程度。黑色六邊形(左側及右上側)及正方形(右下側)標記保持s識別核酸股之序列所需的最少數量之個別分子之單位格子。對於SMAS格子，說明其中僅需要三個DNA股以進行成功鹼基識別之理想情況，該情況進一步詳細地論述於下文。應注意的是，在SMAS情況(圖10之右手側)下，DNA例項隨機分佈在整個感測器陣列110中，且其位置可在第一定序循環期間進行識別，如下文進一步論述。

【0159】 如圖10中所顯示，CLUS裝置之單位格子之面積為2.2×10^5 nm^2，其對應於4.6×10^8個簇/cm^2之DNA簇密度。運用上文進行的假設，CLUS定序儀為每平方厘米感測區域產生約70 Gbase資料。相反地，在理想情況下，當僅使用股的三個例項時，SMAS裝置100產生約500 Gb/cm^2 (磁感測器105 (例如MTJ)及磁標記102 (例如超順磁性奈米粒子))及約300 Gb/cm^2 (光學感測器105 (超解析成像)及螢光標記102)資料。CLUS定序儀及SMAS裝置100之示例性實施案之結果概述於下表中，該表估算定序通量，假設各DNA股僅三個例項且假設SMAS實施案之各DNA股十個例項。

平臺	**簇/DNA股分離 [nm]**	**估算的通量(Gb/cm^2) (3個DNA例項)**	**估算的通量(Gb/cm^2) (10個DNA例項)**
CLUS	~500	~70	~70
螢光SMAS	140	~300	~90
磁性SMAS	100	~500	~150

【0160】 上表顯示當用於下文另外描述之算法錯誤校正之DNA例項之數量較小(例如<10)時，SMAS裝置100優於當前最先進技術CLUS裝置。由於錯誤校正程序仰賴於各ssDNA之更多例項，故SMAS裝置100開始表現得像CLUS裝置，且不同於感測簇，在感測個別分子方面幾乎沒有效益。螢光SMAS基本上代表將簇減少到單個分子之限制。降低定序成本

的一種方法係縮小簇大小且將DNA簇封裝成彼此更接近以便自經固定之感測區域獲得更多資訊。儘管此種方法減少運行定序化學所需的試劑的量，但藉由不斷推動商業光學儀器目前可能的極限，其亦顯著增加成像硬體之複雜性及成本。該策略係一項艱巨任務，因為沒有化學方面之並行改良，則無法進行按比例調整(in-scaling)。此係因為隨著簇變小，各反應變得越來越重要，且於單分子位準上隨機發生之化學失敗變得更加明顯且難以容忍。

【0161】 在CLUS裝置中實施超解析成像之成本係使得SMAS裝置100，且特別是使用磁感測器105及磁標記之SMAS裝置100成為可能破壞性定序替代者。此處揭示的SMAS裝置100，且特別是彼等使用磁感測器105者藉由利用藉由大規模半導體及資料儲存工業開發的技術及大量製造以顯著更低儀器成本保證優異通量。

【0162】

SMAS定序方案

如以上所說明，當SMAS裝置100用於核酸定序時，可在將核酸添加至SMAS裝置100之前或之後(例如使用橋式擴增)擴增核酸股。無論如何擴增核酸，該等股可藉由SBS (例如藉由自ssDNA合成dsDNA)一次一個鹼基地進行定序。描述SMAS定序方案，假設定序的核酸為DNA。應瞭解，所揭示的方案可經修改以用於其他核酸之定序。藉由對本文揭示內容之理解，此類修改將在一般技術者的能力範圍內。

【0163】 為了簡化分析且說明使用所揭示的SMAS裝置100而不是CLUS定序儀之益處，考慮DNA定序方案，其中將單一類型之標記(例如分子、螢光、磁性等)附接至所有四個核苷酸(**A**、**T**、**C**及**G**)。換言之，將

某種類型之相同標記附接至四個核苷酸中之各者(例如若所選擇的標記102為FePt粒子，則**A**、**T**、**C**及**G**中之各者係經FePt粒子標記)。然後使用終止化學將此等經標記之核苷酸一次一個鹼基地併入至DNA股中，例如一旦併入核苷酸，則在聚合酶移至下一鹼基上之前切割標記102。感測器105偵測附接至核苷酸之標記102。

【0164】使用SMAS裝置100定序複數個核苷酸股(例如ssDNA)之示例性方法200繪示於圖11中。在202處，該方法開始。在204處，一或多個核酸股可視需要在被添加至SMAS裝置100之前進行擴增。在206處，將複數個S個核酸股結合至SMAS裝置100之複數個S個結合位點116(其中該複數個包括SMAS裝置100之至少兩個但不一定是所有結合位點116)。視需要，在208處，擴增核酸股(例如經由橋式擴增，此可以在204處之擴增之補充或替代地進行)。在210處，進行定序程序。定序程序可為例如下文進一步描述之加性方法、減性方法或改進型加性方法。在210處進行的定序程序產生S個記錄，S個記錄中之各者捕獲複數個S個感測器中之一者之M個偵測結果(其中，再者，該複數個包括SMAS裝置100中之至少兩個但不一定是所有感測器105，且該M個偵測結果可包含少至一個偵測結果、在定序程序期間獲得的偵測結果總數之一些子組、或在定序程序期間獲得的所有偵測結果)。M個偵測結果中之各者指示在M個查詢步驟之各別步驟期間記錄所對應的感測器105是否偵測到至少一個標記。該M個偵測結果可儲存在記錄中，該記錄可儲存在記憶體中。在212處，進行錯誤校正程序，如下文進一步所述。錯誤校正程序可包含判定性及/或概率性錯誤校正技術。錯誤校正程序可例如藉由SMAS裝置100之至少一個處理器130來進行。或者，其可藉由SMAS裝置100外部的處理器(例如裝置外處理器，

諸如在外部電腦中)來進行。錯誤校正程序可在定序程序進行時(例如即時或近即時)進行，或其可在某個後來的時間進行。在214處，方法200結束。

【0165】 如上所述，在210處，可使用SMAS裝置100來實施多種方案以讀取核酸序列(例如DNA序列)。為了簡化分析，假設SMAS裝置100之複數個S個感測器105僅偵測標記102之存在或不存在且不基於所偵測到的信號位準來區分核苷酸。因此，在一些實施例中，各感測器105的偵測結果之記錄僅包含在特定查詢步驟期間感測器105偵測到標記或未偵測到標記之「是」或「否」 (或1/0或任何其他二進制指示符)指示。應明瞭，其他方法係可行的且在本文揭示內容之範疇內。例如，不同標記102可附接至不同核苷酸。作為另一個實例，不同於二進制「是」或「否」決策，可偵測(例如電阻、頻率、強度等)及/或記錄特性值，且基於該基礎上做出是否偵測到標記之決策。例如，替代僅將0及1(或「否」及「是」)作為定序程序之可能輸出，對於不同核苷酸使用不同標記可導致以下五個等級之一：0 (未偵測到標記)，等級1 (偵測到標記1)，等級2 (偵測到標記2)，等級3 (偵測到標記3)，及等級4 (偵測到標記4)。在此種情況下，可限定所偵測到的特性之範圍以區分是否完全偵測到標記且若偵測到，則偵測到哪個標記(例如若特性值在0與第一值之間，則判定沒有偵測到標記；若特性值在第一值與第二值之間，則判定偵測到第一標記；若特性值在第二值與第三值之間，則判定偵測到第二標記；等)。

【0166】 下文為DNA定序方案之三個實例之說明，各實例包含重複查詢循環，各查詢循環具有四個查詢步驟。在各查詢循環期間，對定序的各ssDNA回答四個二進制「是」或「否」問題。在一個查詢步驟中，回答

問題「偵測到的鹼基為腺嘌呤?」(「**A**?」)。在另一個查詢步驟中，回答問題「偵測到的鹼基為胸腺嘧啶?」(「**T**?」)。在另一個查詢步驟中，回答問題「偵測到的鹼基為胞嘧啶?」(「**C**?」)。且在另一個查詢步驟中，回答問題「偵測到的鹼基為鳥嘌呤?」(「**G**?」)。在定序程序期間獲得的偵測結果之記錄可建立為查詢循環，包括重複**A**?⇒**T**?⇒**C**?⇒**G**?查詢步驟。應明瞭，所描述的其中引入核苷酸及偵測鹼基之順序係任意的(意指查詢步驟之順序係任意的)，及其中在本文實例中測試鹼基之排序(**A**?⇒**T**?⇒**C**?⇒**G**?)僅係示例性的。

【0167】

加性方法

在加性方法中，感測器105偵測結合至具有可切割之連接子之核苷酸之奈米級標記102。所有四種類型之核苷酸攜載相同類型之標記102 (例如分子、螢光、磁性等)且使用相同類型之可切割之連接子。根據一個實施例，將產生四個偵測結果(該等偵測結果中之一者將(不存在錯誤之情況下)為複數個S個核酸股101中之各者之標記偵測)之查詢循環涉及以下步驟：

1. 獲得SMAS裝置100之複數個S個感測器105 (其可為感測器陣列110中之全部或少於全部感測器105)中之各者之基線特性(例如藉由測定複數個S個感測器105中之各者處之基線信號)。

2. 引入且併入經標記之**A**核苷酸。沖洗掉未結合的經標記之分子。

3. 查詢步驟1：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢

循環之查詢步驟1之記錄中之位置。

4. 引入且併入經標記之**T**核苷酸。沖洗掉未結合的經標記之分子。

5. 查詢步驟2：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟2之記錄中之位置。

6. 引入且併入經標記之**C**核苷酸。沖洗掉未結合的經標記之分子。

7. 查詢步驟3：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟3之記錄中之位置。

8. 引入且併入經標記之**G**核苷酸。沖洗掉未結合的經標記之分子。

9. 查詢步驟4：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟4之記錄中之位置。

10. 切割且沖洗掉**A**、**T**、**C**及**G**核苷酸之標記。

【0168】 然後可為下一個查詢循環重複步驟1至10。應明瞭，步驟1至10中之某些之排序係示例性的，且進一步地，步驟1至10之數量及編號係為了方便起見且可進行修改。作為一個實例，且如前面所說明，核苷酸引入之順序係任意的。作為另一個實例，步驟2、4、6及8包括引入及併入核苷酸，且以單個步驟沖洗掉未結合的核苷酸，但應明瞭步驟2、4、6及8中之各者可分為一系列較小的步驟。類似地，步驟3、5、7及9可進一

步分為一系列較小的步驟(例如獲得特性，判定是否偵測到標記，保存偵測結果)。相反地，步驟可組合(例如步驟2及3可組合，步驟4及5可組合等)。

【0169】 應明瞭，若加性方法之任何查詢循環期間可能沒有錯誤發生，則一旦偵測到標記即可識別(判定)個別股之各別鹼基。例如，參照以上步驟，若在涉及經標記之**A**核苷酸之查詢步驟1處，對於特定感測器105，所獲得的特性指示感測器105偵測到標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**A**互補之鹼基(**T**)。類似地，若在涉及經標記之**T**核苷酸之查詢步驟2處，對於特定感測器105，所獲得的特性指示感測器105偵測到標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**T**互補之鹼基(**A**)。同樣地，若在涉及經標記之**C**核苷酸之查詢步驟3處，對於特定感測器105，所獲得的特性指示感測器105偵測到標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與C互補之鹼基(**G**)。最後，若在涉及經標記之**G**核苷酸之查詢步驟4處，對於特定感測器105，所獲得的特性指示感測器105偵測到標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**G**互補之鹼基(**C**)。然而，如下文所進一步詳細說明，有幾種類型之錯誤可在定序程序期間(例如在加性方法期間)發生，且因此，在一些實施例中，在定序程序期間建立記錄以記錄在各查詢循環之各查詢步驟期間偵測到/未偵測到標記。然後可在識別鹼基之前將錯誤校正程序應用於一些或全部記錄。

【0170】 圖12為根據一些實施例之使用加性方法之定序程序220之流程圖。定序程序220可為例如在顯示且描述於圖11之論述中之使用SMAS裝置100定序複數個核酸股(例如ssDNA)之示例性方法200之步驟

210處進行之定序程序。在222處，定序程序220開始。在224處，獲得S個感測器105中之各者之基線特性(例如藉由SMAS裝置100之至少一個處理器130，藉助於電路120)。當查詢循環開始時，在226處，選擇第一經標記之核苷酸(例如參照以上步驟1至10，該第一經標記之核苷酸將為**A**)。在228處，將所選擇的經標記之核苷酸引入至流體腔室115中且將核苷酸潛在地併入與結合位點116結合之核酸股中。在230處，沖洗掉未結合的核苷酸。在232處，自該複數個S個感測器中之各者獲得特性，且判定該複數個S個感測器105中之各者之偵測結果(例如偵測到標記或未偵測到標記)。在234處，將S個偵測結果記錄在S個記錄中(例如以1指示偵測到標記或以0指示未偵測到標記)。在236處，判定最後測試的核苷酸是否為查詢循環之最後一個核苷酸。對於在以上步驟1至10中假設的核苷酸測試之實例排序，將在236處(例如藉由至少一個處理器130)判定**G**是否為最後測試的核苷酸。若不是，則在238處選擇在查詢循環中欲測試的下一個經標記之核苷酸，且重複步驟228至236直至在236處判定最後測試的核苷酸為查詢循環之最後一個核苷酸。在240處，將標記切割且沖洗掉。在242處，判定(例如藉由至少一個處理器130)最後完成的查詢循環是否為定序程序220之最後一個查詢循環。例如，該至少一個處理器130可判定是否已記錄足夠的偵測結果以使得至少一個處理器130 (或一些其他處理實體，諸如外部處理器)以判定目標數量之鹼基(例如150個鹼基)。若不是，則定序程序220返回至步驟224。若是，則定序程序220在244處結束。同樣地，如上文所說明，測試核苷酸之順序係任意的。

【0171】 加性定序方案(其在DNA定序之示例性情況下包含四次核苷酸併入及一次標記切割反應)概述於圖13中。圖13之最左側小圖說明具

有總共100個個別感測器105之感測器陣列110，其以正方形顯示。出於說明之目的，假設感測器陣列110中之100個結合位點116中之各者保持各別DNA股，且各DNA股藉由各別感測器105感測(換言之，結合位點116及感測器105處於一對一的關係)。一些DNA股可為其他DNA之複本。將經標記之核苷酸一次一種類型地添加至流體腔室115，且在併入核苷酸後同時切割標記。在不存在錯誤的情況下，可在五次反應(亦即，四次核苷酸併入及一次鹼基切割反應)之後完成鹼基識別。若發生錯誤，則可應用如下文所述的錯誤糾正程序。

【0172】

減性方法

在減性方法中，感測器105偵測與具有可切割之連接子之核苷酸結合之奈米級標記102。所有四種類型之核苷酸攜載相同類型之標記(例如分子、螢光、磁性等)，但各者具有不同類型之可切割之連接子。在一個實施例中，在不存在錯誤的情況下將產生四個偵測結果(該等偵測結果中之一者將(不存在錯誤的情況下)為複數個S個核酸股101中之各者之標記偵測)之查詢循環涉及以下步驟：

1. 同時引入經標記之**A**、**T**、**C**及**G**核苷酸，併入，且沖洗未結合的經標記之分子。獲得複數個S個感測器105中之各者之基線特性(例如藉由偵測該複數個S個感測器105中之各者處之信號)。在不存在錯誤的情況下，全部感測器105均將偵測標記。

2. 查詢步驟1：引入僅自第一核苷酸(例如**A**)切割標記之試劑(例如酵素)，沖洗，且獲得該複數個S個感測器105中之各者處之特性(例如測定信號)。判定(例如基於基線特性之變化)哪些感測器105不再偵測標記。將各

感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟1之記錄中之位置。

3. 查詢步驟2：引入僅自第二核苷酸(例如**T**)切割標記之試劑，沖洗，且獲得該複數個S個感測器105中之各者處之特性(例如測定信號)。判定(例如基於基線特性之變化)哪些感測器105不再偵測標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟2之記錄中之位置。

4. 查詢步驟3：引入僅自第三核苷酸(例如**C**)切割標記之試劑，沖洗，且獲得該複數個S個感測器105中之各者處之特性(例如測定信號)。判定(例如基於基線特性之變化)哪些感測器105不再偵測標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟3之記錄中之位置。

5. 查詢步驟4：引入僅自第四核苷酸(例如**G**)切割標記之試劑，沖洗，且獲得該複數個S個感測器105中之各者處之特性(例如測定信號)。判定(例如基於基線特性之變化)哪些感測器105不再偵測標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟4之記錄中之位置。

【0173】 對於下一查詢循環，可重複步驟1至5。應明瞭，步驟1至5中之某些之排序係示例性的，且進一步地，步驟1至5之數量及編號係為了方便起見且可進行修改。作為一個實例，且如前面所說明，切割核苷酸之順序係任意的。類似地，在步驟1中，可繼而引入(不必同時地)核苷酸。作為另一個實例，查詢步驟1、2、3及4包括引入試劑，沖洗，獲得特性，判定哪些感測器不再(或仍在)偵測標記，及以單個步驟將結果保存，

但應明瞭，各查詢步驟可分為一系列較小的步驟。

【0174】 應明瞭，若減性方法之任何查詢循環期間可能沒有錯誤發生，則一旦首先偵測到標記移除(標記不存在)即可識別(判定)個別股之各別鹼基。例如，參照以上步驟，若在涉及經標記之**A**核苷酸之查詢步驟1處，對於特定感測器105，所獲得的特性指示感測器105不再偵測標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**A**互補之鹼基(**T**)。類似地，若在涉及經標記之**T**核苷酸之查詢步驟2處，對於特定感測器105，所獲得的特性指示感測器105不再偵測標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**T**互補之鹼基(**A**)。同樣地，若在涉及經標記之**C**核苷酸之查詢步驟3處，對於特定感測器105，所獲得的特性指示感測器105不再偵測標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與C互補之鹼基(**G**)。最後，若在涉及經標記之**G**核苷酸之查詢步驟4處，對於特定感測器105，所獲得的特性指示感測器105不再偵測標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**G**互補之鹼基(**C**)。然而，如下文所進一步詳細說明，有幾種類型之錯誤可在定序程序期間(例如在減性方法期間)發生，且因此，在一些實施例中，在定序程序期間建立記錄以記錄在各查詢循環之各查詢步驟期間偵測到/未偵測到標記。然後可在識別鹼基之前將錯誤校正程序應用於一些或全部記錄。

【0175】 圖14為根據一些實施例之使用減性方法之定序程序250之流程圖。定序程序250可為例如在顯示且描述於圖11之論述中之使用SMAS裝置100定序複數個核酸股(例如ssDNA)之示例性方法200之步驟210處進行之定序程序。定序程序250在252處開始。在254處，將全部經

標記之核苷酸引入至流體腔室115中且將核苷酸併入與S個結合位點116結合之核酸股中。在256處，沖洗掉未結合的核苷酸。在258處，獲得S個感測器105中之各者之基線特性 (例如藉由SMAS裝置100之至少一個處理器130，藉助於電路120)。假設已將核苷酸引入至與S個結合位點中之各者結合之核酸股中，所獲得的特性表示感測器105在其正在偵測至少一個標記時之特性。在260處，選擇可切割之連接子中之一者以用於切割(或，等效地，選擇核苷酸中之一者)。在262處，切割且沖洗掉附接至所選擇的核苷酸之標記。假設沒有錯誤，在步驟262之後，感測彼等併入所測試的核苷酸之核酸股(例如藉由所選擇的可切割之連接子附接標記之核酸股)之感測器105將展現特性之變化(例如與感測器105相關聯或由其產生之信號之變化)。在264處，自該複數個S個感測器中之各者獲得特性，且判定該複數個S個感測器105中之各者之偵測結果(例如偵測到標記或未偵測到標記)。在266處，將S個偵測結果記錄在S個記錄中(例如以1指示偵測到標記或以0指示未偵測到標記)。在268處，判定最後測試的核苷酸是否為查詢循環之最後一個核苷酸。對於在以上步驟1至5中假設的核苷酸測試之實例排序，將在268處(例如藉由至少一個處理器130)判定**G**是否為最後測試的核苷酸。若不是，則在270處選擇在查詢循環中欲切割的下一個可切割之連接子(或等效地，欲測試的下一個核苷酸)，且重複步驟262至268直至在268處判定最後切割之連接子(或等效地，最後測試的核苷酸)係查詢循環之最後一個連接子(或核苷酸)。在272處，判定(例如藉由至少一個處理器130)最後完成的查詢循環是否為定序程序250之最後一個查詢循環。例如，該至少一個處理器130可判定是否已記錄足夠的偵測結果以使得至少一個處理器130 (或一些其他處理實體，諸如外部處理器)以識別目標數量

之鹼基(例如150個鹼基)。若不是，則定序程序250返回至步驟254。若是，則定序程序250在274處結束。同樣地，如上文所說明，測試核苷酸之順序係任意的。

【0176】 減性定序方案(其在DNA定序之示例性情況下包含一次核苷酸併入及四次鹼基切割反應)概述於圖15中。圖15之最左側小圖說明具有總共100個個別感測器105之感測器陣列110，其以正方形顯示。出於說明之目的，假設感測器陣列110中之100個結合位點116中之各者保持各別DNA股，且各DNA股藉由各別感測器105感測(換言之，結合位點116及感測器105處於一對一的關係)。一些DNA股可為其他DNA之複本。將所有四種類型之經標記之核苷酸同時添加至流體腔室115，且在併入之後一次一種類型之核苷酸(例如可切割之連接子)地移除標記。在不存在錯誤的情況下，可在五次反應(亦即，一次核苷酸併入及四次鹼基切割反應)之後完成鹼基識別。若發生錯誤，則可應用如下文所述的錯誤糾正程序。

【0177】

改進型加性方法

在改進型加性方法中，感測器105偵測與具有可切割之連接子之核苷酸結合之奈米級標記102。所有四種類型之核苷酸攜載相同類型之標記102 (例如分子、螢光、磁性等)且使用相同類型之可切割之連接子。分別添加經標記之核苷酸，且在添加各核苷酸之後，偵測到標記102之存在。在一個實施例中，在不存在錯誤的情況下將產生四個偵測結果(該等偵測結果中之至少一者將為複數個S個核酸股101中之各者之標記偵測)之查詢循環涉及以下步驟：

1. 獲得SMAS裝置100之複數個S個感測器105 (其可為感測器陣列

110中之全部或少於全部感測器105)中之各者之基線特性(例如藉由測定複數個S個感測器105中之各者處之基線信號)。

2. 引入且併入第一經標記之核苷酸，例如經標記之**A**核苷酸。沖洗掉未結合的經標記之分子。

3. 查詢步驟1：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟1之記錄中之位置。

4. 切割且沖洗掉標記。

5. 引入且併入第二經標記之核苷酸，例如經標記之**T**核苷酸。沖洗掉未結合的經標記之分子。

6. 查詢步驟2：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟2之記錄中之位置。

7. 切割且沖洗掉標記。

8. 引入且併入第三經標記之核苷酸，例如經標記之**C**核苷酸。沖洗掉未結合的經標記之分子。

9. 查詢步驟3：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟3之記錄中之位置。

10. 切割且沖洗掉標記。

11. 引入且併入第四經標記之核苷酸，例如經標記之**G**核苷酸。沖洗掉未結合的經標記之分子。

12. 查詢步驟4：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟4之記錄中之位置。

13. 切割且沖洗掉標記。

【0178】 然後，對於下一查詢循環，可重複步驟1至13。應明瞭，步驟1至13中之某些之排序係示例性的，且進一步地，步驟1至13之數量及編號係為了方便起見且可進行修改。作為一個實例，且如前面所說明，核苷酸引入之順序係任意的。作為另一個實例，步驟2、5、8及11包括引入及併入核苷酸，且以單個步驟沖洗掉未結合的核苷酸，但應明瞭步驟2、5、8及11中之各者可分為一系列較小的步驟。類似地，步驟3、6、9及12 (分別係查詢步驟1、2、3及4)可進一步分為一系列較小的步驟(例如獲得特性，判定是否偵測到標記，保存偵測結果)。相反地，步驟可組合(例如步驟2及3可組合，步驟3及4可組合，步驟2至4可組合，步驟5及6可組合，步驟6及7可組合，步驟5至7可組合等)。

【0179】 應明瞭，若改進型加性方法之任何查詢循環期間可能沒有錯誤發生，則一旦偵測到標記即可識別(判定)各個股之各別鹼基。例如，參照以上步驟，若在涉及經標記之**A**核苷酸之查詢步驟1處，對於特定感測器105，所獲得的特性指示感測器105偵測到標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**A**互補之鹼基(**T**)。類似地，若在涉及經標記之**T**核苷酸之查詢步驟2處，對於特定感測器105，所

獲得的特性指示感測器105偵測到標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**T**互補之鹼基(**A**)。同樣地，若在涉及經標記之**C**核苷酸之查詢步驟3處，對於特定感測器105，所獲得的特性指示感測器105偵測到標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與C互補之鹼基(**G**)。最後，若在涉及經標記之**G**核苷酸之查詢步驟4處，對於特定感測器105，所獲得的特性指示感測器105偵測到標記，則保存偵測結果可等同於識別該偵測器105 (及結合位點116)之與**G**互補之鹼基(**C**)。然而，如下文所進一步詳細說明，有幾種類型之錯誤可在定序程序期間(例如在加性方法期間)發生，且因此，在一些實施例中，在定序程序期間建立記錄以記錄在各查詢循環之各查詢步驟期間偵測到/未偵測到標記。然後可在識別鹼基之前將錯誤校正程序應用於一些或全部記錄。

【0180】 圖16為根據一些實施例之使用改進型加性方法之定序程序350之流程圖。定序程序350可為例如在顯示且描述於圖11之論述中之使用SMAS裝置100定序複數個核酸股(例如ssDNA)之示例性方法200之步驟210處進行之定序程序。在352處，定序程序350開始。在354處，獲得S個感測器105中之各者之基線特性 (例如藉由SMAS裝置100之至少一個處理器130，藉助於電路120)。當查詢循環開始時，在356處，選擇第一經標記之核苷酸(例如參照以上步驟1至13，該第一經標記之核苷酸將為**A**)。在358處，將所選擇的經標記之核苷酸引入至流體腔室115中且將核苷酸潛在地併入與結合位點116結合之核酸股中。在360處，沖洗掉未結合的核苷酸。在362處，自該複數個S個感測器中之各者獲得特性，且判定該複數個S個感測器105中之各者之偵測結果(例如偵測到標記或未偵測到標

記)。在364處，將S個偵測結果記錄在S個記錄中(例如以1指示偵測到標記或以0指示未偵測到標記)。在366處，將標記切割且沖洗掉。在368處，判定最後測試的核苷酸是否為查詢循環之最後一個核苷酸。對於在以上步驟1至13中假設的核苷酸測試之實例排序，將在368處(例如藉由至少一個處理器130)判定**G**是否為最後測試的核苷酸。若不是，則在370處選擇在查詢循環中欲測試的下一個經標記之核苷酸，且重複步驟358至368直至在368處判定最後測試的核苷酸為查詢循環之最後一個核苷酸。在372處，判定(例如藉由至少一個處理器130)最後完成的查詢循環是否係定序程序350之最後一個查詢循環。例如，該至少一個處理器130可判定是否已記錄足夠的偵測結果以使得至少一個處理器130 (或一些其他處理實體，諸如外部處理器)以識別目標數量之鹼基(例如150個鹼基)。若不是，則定序程序350返回至步驟354。若是，則定序程序350在374處結束。同樣地，如上文所說明，測試核苷酸之順序係任意的。

【0181】 改進型加性定序方案(其在DNA定序之示例性情況下包含四次核苷酸併入及四次鹼基切割反應)說明於圖17中。圖17之最左側小圖說明具有總共100個個別感測器105之感測器陣列110，其以正方形顯示。出於說明之目的，假設感測器陣列110中之100個結合位點116中之各者保持各別DNA股，且各DNA股藉由各別感測器105感測(換言之，結合位點116及感測器105處於一對一的關係)。一些DNA股可為其他DNA之複本。如所顯示及所描述，將經標記之核苷酸一次一種類型地添加至流體腔室115，且在併入及標記偵測之後切割標記。在不存在錯誤的情況下，可在平均5次反應(亦即，2.5次核苷酸併入及2.5次鹼基切割反應)之後完成鹼基識別。

【0182】 因此，在不存在錯誤的情況下，對於DNA定序，改進型加性方法在8次反應(4次核苷酸併入及4次鹼基切割)以測試所有鹼基之後每個ssDNA產生至少一次鹼基識別。然而，平均而言，僅在5次反應(2.5次核苷酸併入及2.5次鹼基切割)之後即可進行鹼基識別。因為標記係在每個核苷酸之引入之後被移除，故可在單個**A?⇒T?⇒C?⇒G?**查詢循環期間併入且識別多個核苷酸。具體而言，在未知ssDNA序列中，未知鹼基是**T**的機率為四分之一。若鹼基恰好是**T**，則將在第三步驟處在引入**A**核苷酸時在一次併入及一次鹼基切割反應之後偵測到其。未知鹼基是**A**的機率為四分之一。若鹼基恰好是**A**，則將在查詢循環**A?⇒T?**之第五步驟處在已引入**T**核苷酸且已進行兩次引入及兩次切割時偵測到其。未知鹼基是**G**的機率為四分之一。若鹼基恰好是**G**，則將在查詢循環**A?⇒T?⇒C?**之第七步驟處在已引入**C**核苷酸且已進行三次引入及三次切割時偵測到其。最後，未知鹼基是**C**的機率為四分之一。若鹼基恰好是**C**，則將在查詢循環**A?⇒T?⇒C?⇒G?**之第十一步驟處在已引入**C**核苷酸且已進行四次引入及四次切割時偵測到其。因此需要平均2.5次查詢(5次反應)($\frac{1}{4}\times1+\frac{1}{4}\times2+\frac{1}{4}\times3+\frac{1}{4}\times4=2.5$)以識別單個未知鹼基。或者，若特定ssDNA之未知4-鹼基序列恰好是最佳情境**ATCG** (對於針對本實例假設的所引入的核苷酸之選定順序)，則僅需要進行一個查詢循環**A?⇒T?⇒C?⇒G?**：總共8次反應(4次核苷酸併入及4次鹼基切割)，或每次鹼基識別2次反應。然而，若未知序列恰好是例如**GCTA**、**GGCT**、**GCTT**、**GGGG**等，則需要進行四個查詢循環，各者包括所有**A?⇒T?⇒C?⇒G?**，導致總共32次反應(16次核苷酸併入及16次鹼基切割)，或每次鹼基識別8次反應。然而，平均而言，對於隨機DNA序列，需要2.5次查詢或5次反應(2.5次核苷酸併入及2.5次鹼基

切割)以進行一次鹼基識別。

【0183】

定序錯誤之來源

理想地，不論在CLUS裝置或SMAS裝置100中，定序程序均將係無錯誤的。換言之，例如，核苷酸將始終經正確標記，核苷酸將始終經正確地併入DNA中，將在切割步驟期間成功切割所有標記，將成功沖洗掉所有切割的標記等。然而，事實上，錯誤可在任何定序程序期間發生。本部分探討CLUS裝置及SMAS裝置100兩者之定序錯誤之來源且描述SMAS裝置100之錯誤減輕策略。如下文另外所說明，錯誤校正方法可用於改良SMAS裝置100之定序準確度。

【0184】 因為上文描述的改進型加性方法係概念上簡單(且對稱，因為各核苷酸以相同方式處理)定序程序，故其係用於說明錯誤如何在CLUS裝置及SMAS裝置100兩者中傳播之良好模型。考慮到錯誤之四個來源，假設奈米級標記係經由可切割之連接子附接至核苷酸。各錯誤以表示為r之比率發生，其具有0至1之值。錯誤之四個來源係：

失敗之核苷酸併入(FNI)：失敗之核苷酸併入(FNI)發生在經正確標記之核苷酸分子尚未到達ssDNA結合位點或聚合酶未能將其併入之時。圖18A說明定序ssDNA之五個例項之CLUS裝置之FNI。在互補核苷酸之流動之後，五個ssDNA中僅三個已併入經標記之核苷酸(說明為具有磁標記)。因此，五分之二的核苷酸($r = 0.4$)不能併入。圖18B說明SMAS裝置100之FNI。五個結合位點116中之各者保持ssDNA之例項。在互補核苷酸之流動之後，五個ssDNA (彼等與結合位點116A、116B及116C結合之ssDNA)中僅三者已併入經標記之核苷酸(僅出於實例之目的說明為具有磁

標記)。再者，五個ssDNA例項中的兩個($r = 0.4$)無法併入核苷酸。

【0185】 **失敗之標記移除(FLR)**：在併入經標記之核苷酸分子，但由於切割試劑尚未到達連接子或未能切割其而在標記偵測後未移除標記時，導致失敗之標記移除(FLR)。圖18C說明上文在圖18A之論述中描述之CLUS裝置之FLR。在併入互補核苷酸且沖洗以移除未結合的核苷酸，偵測標記，及切割且沖洗標記之後，一個標記仍附接至ssDNA例項中之一者 ($r = 0.2$)。類似地，在圖18D中，其說明上文在圖18B之論述中描述之SMAS裝置100之FLR，在併入互補核苷酸且沖洗以移除未結合的核苷酸，偵測標記，及切割且沖洗標記(例如上文描述的步驟2至4、5至7、8至10、及/或11至13)之後，標記仍附接至結合位點116A處之ssDNA ($r = 0.2$)。

【0186】 **失敗之核苷酸移除(FNR)**：當經標記之核苷酸(不論是互補還是非互補)非特異性結合至結合位點116及/或感測器105之表面時，導致失敗之核苷酸移除(FNR)。圖18E說明上文在圖18A中描述的CLUS裝置之FNR之一個實例。在核苷酸之流動且沖洗以移除未結合的核苷酸之後，兩個不良核苷酸及其標記保留在結合位點之表面上。類似地，在圖18F中，該圖說明上文在圖18B之論述中描述之SMAS裝置100之FNR，在核苷酸之流動且沖洗以移除未結合的核苷酸之後，一個不良核苷酸保留在結合位點116A之表面上，及另一個不良核苷酸保留在結合位點116D之表面上。在本實例中，對於兩個CLUS裝置及SMAS裝置100兩者，$r = 0.4$。

【0187】 **失敗之標記偵測(FLD)**：在併入正確互補核苷酸，，但由於標記缺失或感測器未能識別其而未偵測到標記時，導致失敗之標記偵測(FLD)。圖18G說明上文在圖18A之論述中描述的CLUS裝置之FLD。在併

入互補核苷酸且沖洗以移除未結合的核苷酸之後，ssDNA例項中的兩者已併入互補核苷酸，但標記缺失($r = 0.4$)。類似地，在圖18H中，該圖說明上文在圖18B之論述中描述的SMAS裝置100之FLD，在併入互補核苷酸且沖洗以移除未結合的核苷酸(例如上文描述的步驟2、5、8或11)之後，應附接至在結合位點116C及116D處併入ssDNA中之核苷酸之標記缺失($r = 0.4$)。

【0188】 圖18A至18H將標記說明為磁體，藉此表明磁標記及磁感測器，但應明瞭，如上文所說明，標記可為任何類型之可偵測之標記(例如螢光、磁性等)及感測器可為能夠偵測選定類型之標記(例如光學、磁性、有機金屬、帶電荷分子等)之任何類型之感測器。

【0189】 假設四個錯誤類型(FNI、FLR、FNR及FLD)以相同比率r發生，其中$0 < r < 1$；例如，若$r = 0.01$，則平均100例中有1例失敗。亦假設SMAS裝置100之感測器105 (例如奈米級感測器105)可幾乎每次偵測單個標記，及用於CLUS裝置中之大型簇感測器之反應係線性的，例如，CLUS裝置之感測器可針對於N之所有值來區分N及$N + 1$個經標記之股。

【0190】

簇定序儀與單分子陣列定序儀：定性比較及錯誤校正

本文揭示兩種類型之錯誤校正，稱為判定性錯誤校正及概率性錯誤校正。SMAS裝置100可使用一種或兩種類型之錯誤校正，如下文所進一步說明。

【0191】 如上文所說明，改進型加性方法係用於說明錯誤如何傳播及如何可實施所揭示的錯誤校正算法之良好模型。應明瞭，當使用其他定序方法(諸如加性方法或減性方法)時，亦可應用所揭示的錯誤減輕算法。

【0192】 考慮使用改進型加性方法定序程序之CLUS裝置及SMAS裝置100，其具有$r = 0.1$ (例如10次反應中1次失敗)之大錯誤率及(理想上相同)股之少數例項，例如，$N = K = 3$，其中變數N表示用於CLUS裝置中之簇大小，及變數K表示感測相同DNA股之例項之SMAS裝置100之感測器105之數量。(如前面所說明，K感測器可彼此接近，或其可散佈在感測器陣列110內)。為了描述判定性錯誤校正之實施例，最初僅考慮FNI及FLR錯誤。然後考慮FNI、FLR及FLD錯誤，且描述錯誤減輕策略。最後，考慮所有四種類型之錯誤，且描述解決所有四種類型之錯誤之錯誤校正程序。

【0193】 當使用SMAS裝置100時，可偵測到且移除FLR錯誤，不論在定序程序期間或在隨後的某個時間以即時方式。FLR錯誤可藉由在切割且沖洗標記之後獲得S個感測器105中之各者之特性來偵測到。FNI錯誤可藉由檢查各感測器105的記錄且識別感測器105未能偵測到任何標記之查詢循環來偵測到。因此，可根據一個實施例如下調整改進型加性方法以添加此等偵測步驟：

1. 獲得SMAS裝置100之複數個S個感測器105 (其可為感測器陣列110中之全部或少於全部感測器105)中之各者之基線特性(例如藉由測定複數個S個感測器105中之各者處之基線信號)。

2. 引入且併入第一經標記之核苷酸，例如經標記之**A**核苷酸。沖洗掉未結合的經標記之分子。

3. 查詢步驟1：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢

循環之查詢步驟1之記錄中之位置。

4. 切割且沖洗掉標記。

5. 獲得在步驟3中偵測到標記之該複數個S個感測器105中之各者之特性。若彼等感測器105中之任何者之所獲得的特性指示感測器105仍在偵測標記，則化學無法切割標記(例如，對於該感測器，存在FLR錯誤)。

6. 引入且併入第二經標記之核苷酸，例如經標記之**T**核苷酸。沖洗掉未結合的經標記之分子。

7. 查詢步驟2：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟2之記錄中之位置。

8. 切割且沖洗掉標記。

9. 獲得在步驟7中偵測到標記之該複數個S個感測器105中之各者之特性。若彼等感測器105中之任何者之所獲得的特性指示感測器105仍在偵測標記，則化學無法切割標記(例如，對於該感測器，存在FLR錯誤)。

10. 引入且併入第三經標記之核苷酸，例如經標記之**C**核苷酸。沖洗掉未結合的經標記之分子。

11. 查詢步驟3：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟3之記錄中之位置。

12. 切割且沖洗掉標記。

13. 獲得在步驟11中偵測到標記之該複數個S個感測器105中之各者

之特性。若彼等感測器105中之任何者之所獲得的特性指示感測器105仍在偵測標記，則化學無法切割標記(例如，對於該感測器，存在FLR錯誤)。

14. 引入且併入第四經標記之核苷酸，例如經標記之**G**核苷酸。沖洗掉未結合的經標記之分子。

15. 查詢步驟4：獲得該複數個S個感測器105中之各者之特性(例如藉由偵測複數個S個感測器105中之各者處之信號)且判定各感測器105是否偵測到至少一個標記。將各感測器105之偵測結果保存在對應於當前查詢循環之查詢步驟4之記錄中之位置。若存在沒有為查詢循環分配鹼基之感測器105 (例如在查詢循環期間無法偵測到**A**、**T**、**C**或**G**之感測器105)，則化學無法將核苷酸併入(例如對於此等感測器105，存在FNI)。

16. 切割且沖洗掉標記。

17. 獲得在步驟15中偵測到標記之該複數個S個感測器105中之各者之特性。若彼等感測器105中之任何者之所獲得的特性指示感測器105仍在偵測標記，則化學無法切割標記(例如，對於該感測器，存在FLR錯誤)。

【0194】 可然後針對下一個查詢循環重複步驟1至17 (例如以估計下一個鹼基或若先前查詢循環無法讀取當前鹼基則再讀取該當前鹼基)。應明瞭，步驟1至17中之某些之排序係示例性的，且進一步地，步驟1至17之數量及編號係為了方便起見且可進行修改。作為一個實例，且如前面所說明，核苷酸引入之順序係任意的。作為另一個實例，步驟2、6、10及14包括引入及併入核苷酸，且以單個步驟沖洗掉未結合的核苷酸，但應明瞭步驟2、6、10及14中之各者可分為一系列較小的步驟。類似地，步驟

3、7、11及15 (分別係查詢步驟1、2、3及4)可進一步分為一系列較小的步驟(例如獲得特性，判定是否偵測到標記，保存偵測結果)。同樣地，儘管步驟15包括識別FNI錯誤，但該任務可以單獨步驟進行。相反地，步驟可組合(例如一些或全部步驟2至5、一些或全部步驟6至9、一些或全部步驟10至13、一些或全部步驟14至17等)。

【0195】 圖19為根據一些實施例之使用具有FLR及FNI錯誤偵測之改進型加性方法之示例性定序程序400之流程圖。定序程序400可為例如在顯示且描述於圖11之論述中之使用SMAS裝置100定序複數個核酸股(例如ssDNA)之示例性方法200之步驟210處進行之定序程序。在402處，定序程序400開始。在404處，獲得S個感測器105中之各者之基線特性 (例如藉由SMAS裝置100之至少一個處理器130，藉助於電路120)。當查詢循環開始時，在406處，選擇第一經標記之核苷酸(例如參照以上步驟1至17，該第一經標記之核苷酸將係**A**)。在408處，將所選擇的經標記之核苷酸引入至流體腔室115中且將核苷酸潛在地併入與結合位點116結合之核酸股中。在410處，沖洗掉未結合的核苷酸。在412處，自該複數個S個感測器中之各者獲得特性，且判定該複數個S個感測器105中之各者之偵測結果(例如偵測到標記或未偵測到標記)。在414處，將S個偵測結果記錄在S個記錄中(例如以1指示偵測到標記或以0指示未偵測到標記)。在416處，將標記切割且沖洗掉。在418處，獲得在步驟412/414期間偵測到標記之彼等感測器105之特性。在420處，判定在步驟412/414期間偵測到標記之感測器105中之任何者是否仍在偵測標記。若是，則在422處判定已偵測到對於仍在偵測至少一個標記之感測器105之FLR錯誤，即使在416處切割且沖洗掉標記。定序程序400然後繼續至424。若在420處判定(例如藉由至少

一個處理器130)在步驟412/414期間偵測到標記之感測器105中無一者仍在偵測標記，則定序程序亦繼續至424。在424處，判定最後測試的核苷酸是否為查詢循環之最後一個核苷酸。對於在以上步驟1至17中假設的核苷酸測試之實例排序，將在368處(例如藉由至少一個處理器130)判定**G**是否為最後測試的核苷酸。若不是，則在426處選擇在查詢循環中欲測試的下一個經標記之核苷酸，且重複步驟408至420 (且若適用，則至422)直至在424處判定最後測試的核苷酸為查詢循環之最後一個核苷酸。在428處，偵測在最後完成的查詢循環期間無法偵測到任何標記之S個感測器105中之彼等感測器之FNI錯誤。在430處，判定(例如藉由至少一個處理器130)最後完成的查詢循環是否為定序程序400之最後一個查詢循環。例如，該至少一個處理器130可判定是否已記錄足夠的偵測結果以使得至少一個處理器130 (或一些其他處理實體，諸如外部處理器)以識別目標數量之鹼基(例如150個鹼基)。若不是，則定序程序400返回至步驟404。若是，則定序程序400在432處結束。同樣地，如上文所說明，測試核苷酸之順序係任意的。

【0196】

減輕FNI及FLR錯誤

為了說明FNI及FLR錯誤於CLUS裝置及SMAS裝置100之效應，使用各類型之定序儀以識別示例性DNA序列，其中在使用上文描述的SBS之改進型加性方法讀取序列時隨機發生FNI及FLR錯誤。假設FNI及FLR錯誤之錯誤率均為$r \cong 0.1$。示例性序列為：**TAG CAA GGT CCG CTA CTG GCA GAC TGG**。圖20顯示在整個18個**A?⇒T?⇒C?⇒G?**查詢步驟之查詢循環中以$r \cong 0.1$產生的兩種類型之錯誤。如20中所顯示，10次反應中約1

次失敗，且對於所定序的三個ssDNA例項，錯誤均勻分佈於FNI錯誤與FLR錯誤之間。模型情況代表集體行為之許多可能情境中之一者。針對當將三個DNA股放在CLUS裝置之單個感測器上時且當將其放在SMAS裝置100之三個離散奈米級感測器105上時之情況分析FNI及FLR錯誤於鹼基判定精度之後果。

【0197】 圖21說明藉由CLUS裝置感測器偵測到的預期信號位準，該感測器捕捉定序程序期間分子集體之行為。在各查詢步驟處，CLUS裝置感測器可偵測分子集體(由三個ssDNA組成)之四個信號強度位準：亦即偵測到0個標記、1個標記、2個標記或3個標記。CLUS裝置之定序程序會考慮集體之組合信號且無法區分何時對個別股之反應失敗。每當CLUS裝置感測器感測至少兩個標記時，在特定查詢步驟處識別一個鹼基。該臨限值可由決策標準表示：當CLUS感測器信號位準大於1.5時，識別一個鹼基。如圖21指示，化學失敗率高會導致顯著鹼基識別錯誤及極低鹼基識別精度。CLUS裝置方法僅導致21個中的6個(約29%)識別的鹼基符合真實序列。此準確度位準僅比具有25%準確度之隨機猜測略佳(由於具有4個鹼基，故正確猜測一個鹼基為四分之一機率)。此外，CLUS裝置不能分辨出成功及失敗化學反應之間的差異，CLUS裝置亦不知曉顯示於圖20中之FNI (虛線圓)或FLR (帶有反斜杠填充之圓圈)錯誤之位置。對於CLUS裝置，集體平均化會掩蓋FLR錯誤之確切位置。藉由基本上進行關於鹼基插入、刪除及取代位點之位置之有根據猜測，僅可實施概率性錯誤校正算法以略微提高CLUS裝置之鹼基識別之品質。示例性算法描述於例如A. Cacho等人，「A Comparison of Base-calling Algorithms for Illumina Sequencing Technology」，Briefings in Bioinformatics，第17(5)卷，786

至795，2016；W.C. Kao等人，「BayesCall: A model-based base-calling algorithm for high-throughput short-read sequencing」，Genome Res.，第19(10)卷，1884至1895，2009；及C. Ledergerber及C. Dessimoz，「Base-calling for next-generation sequencing platforms」，Brief Bioinform.，第12卷，489–97，2011中。

【0198】圖22說明當使用本文描述的錯誤校正技術時，SMAS裝置100如何可提供更佳準確度。如上文所說明，可在定序程序期間偵測到發生於定序程序期間的FLR錯誤。具體而言，SMAS裝置100知曉(或可找到)FLR之位置，因為獲得各感測器105之特性(例如信號位準)且在切割並沖洗掉標記之後且在引入下一個核苷酸之前進行記錄。FLR錯誤可藉由在進行鹼基識別時將其視為「未偵測到標記」來校正。換言之，若定序程序之記錄包含各查詢步驟之二進制(例如0/1)條目，則FLR可藉由將在彼等查詢步驟處之值自「偵測到的」值更改為「未偵測到的」值來校正。作為一個特定實例，若0表示未偵測到標記及1表示偵測到標記，則在錯誤校正之前，在第m個查詢步驟處的FLR將以記錄中第m個位置中的1表示。該錯誤可藉由將記錄中第m個位置處的值1更改為值0來校正。圖22之頂部說明在錯誤校正以移除FLR錯誤之前SMAS裝置100之三個感測器105中之各者之偵測結果。圖22之下部顯示在識別鹼基之前校正FLR錯誤之結果。

【0199】當超過一半的K個感測器105 (在$K = 3$ (兩個或三個感測器105)之實例中)在特定查詢步驟期間偵測到標記時，使用SMAS裝置100之改進型加性定序程序允許該查詢步驟識別鹼基。然而，不同於CLUS裝置，SMAS裝置100收集相當多的資訊，因為其在複數個(在該實例中假設為3個)結合位點116中之每個結合位點116處且在定序程序之每個查詢步驟

偵測標記之存在或不存在。因此，使用SMAS裝置100可導致進行更少鹼基識別，但彼等識別導致比藉由CLUS裝置識別者顯著準確得多的估計的序列。具體而言，對於示例性序列，一旦已移除FLR錯誤(如圖22之下部所顯示)，使用SMAS裝置100導致16個中的11個(約69%)識別的鹼基符合真實序列。因此，圖21及22說明對於兩種類型之定序裝置，化學失敗於鹼基識別準確度之後果係顯著不同的，且SMAS裝置100提供更佳準確度。

【0200】 當使用SMAS裝置100時，亦可校正FNI錯誤，因為失敗之併入在SMAS感測器105偵測結果中(例如在由在定序程序期間藉由感測器105偵測到/未偵測到標記組成之記錄中)建立特性記號。特別地，改進型加性方法中之FNI錯誤導致四個或更多個連續查詢步驟之一串(連續序列)零(或其他「未偵測到標記」偵測結果)。如圖19之論述中所說明，一些FNI錯誤可藉由識別特定感測器105在查詢循環期間未偵測到任何標記來偵測。應瞭解，FNI錯誤亦可「跨越」多個查詢循環。例如，假設在具有**A**?⇒**T**?⇒**C**?⇒**G**?查詢步驟之第一查詢循環期間，特定感測器105在**A**?查詢步驟期間偵測到標記，且然後其沒有偵測到任何標記直至下一個查詢循環之**C**?查詢步驟。因為**C**?查詢步驟在示例性查詢循環中之**A**?查詢步驟之後，且改進型加性方法用作定序循環，故第一查詢循環之**C**?查詢步驟應已導致偵測到標記。應注意，圖19之步驟428在第一查詢循環或第二查詢循環期間將不會導致偵測到任何FNI錯誤，因為任何一個查詢循環均不會導致特定感測器105未偵測到標記。但對偵測結果記錄之檢查將顯示存在FNI錯誤。FNI錯誤可藉由刪除若干串(在DNA定序之情況下，四個)零以將不良股與不受FNI錯誤影響之股比對來判定性地校正。圖23說明藉由在

定序程序之偵測結果之記錄中刪除若干串四個「未偵測到標記」條目來校正FNI錯誤。如圖23中所顯示，FNI錯誤校正導致識別的序列與真實序列之間完全比對。

【0201】 具有有限組錯誤之簡化模型系統之定性分析表明至少在所定序的DNA股之例項數K很小且化學失敗率很高時使用SMAS裝置100進行核酸定序大大地優異於使用CLUS裝置。為了設置用於兩個平臺之定量比較之框架，下文探討簇大小(對於CLUS裝置)及所定序的例項數(對於SMAS裝置100)如何影響鹼基識別精度。對於FNI及FLR兩種錯誤，考慮其中$N = K = 11$及$r = 0.1$之情況。假設感測器正在讀取上文所考慮的相同實例序列(**TAG CAA GGT CCG CTA CTG GCA GAC TGG**)及18個**A**?⇒**T**?⇒**C**?⇒**G**?查詢步驟之查詢循環隨機出現導致FNI及FLR之化學錯誤。圖24說明基於具有大化學失敗率($r \cong 0.1$或10%)之DNA股之11個例項之示例性SBS反應之結果。如圖24中所顯示，10次反應中有約1次失敗。

【0202】 圖25說明較大簇大小N於CLUS裝置之鹼基識別精度之影響。圖25顯示藉由CLUS裝置感測器偵測到的預期信號位準，該感測器捕捉定序程序期間分子集體之行為。在各查詢步驟處，CLUS裝置感測器可偵測到分子集體(十一個ssDNA)之十二個信號強度位準中之任何一者，亦即偵測到0至11個標記。當藉由CLUS感測器偵測到的信號位準為大於5.5時，在特定查詢步驟處識別鹼基。如圖25顯示，失敗之化學導致鹼基識別錯誤：18個中僅11個(約61%)識別的鹼基符合真實序列。

【0203】 圖25與圖21之比較指示CLUS裝置在$N = 11$下之精度比當$N = 3$時更佳。具體而言，增加簇大小N導致鹼基判定錯誤大大減少。而

在$N = 3$情況下，僅約29%的識別的鹼基與真實序列一致，增加簇大小至$N = 11$使得一致率為約61%，因為該CLUS裝置得益於較大集體之集合行為。當前最先進技術商業CLUS型定序儀與容納約100個DNA股例項之簇陣列一起工作。

【0204】 圖26說明根據一些實施例之在$K = 11$下使用SMAS裝置100之情況下之結果(換言之，ssDNA之11個例項，各藉由不同感測器105偵測到)及FLR及FNI錯誤之判定性錯誤校正。當超過一半(例如，對於$K = 11$，至少6個)的感測器105偵測到標記時，在特定查詢步驟處識別鹼基。如藉由圖26所顯示，實施如上文所述的判定性FLR錯誤校正(中間)及FNI錯誤校正(下部)導致識別的序列與真實序列之間的完全比對。應注意，若不進行錯誤偵測/校正，則基於來自SMAS裝置100之資料之識別的序列將與使用來自CLUS裝置之資料之該識別的序列相同，因為沒有錯誤校正之SMAS裝置100藉由將所有個別感測器結果加總簡單地再建立集體結果。偵測及校正定序資料中之錯誤之能力使得SMAS裝置100相對於CLUS裝置具有優勢。

【0205】 因此，若僅發生FNI及FLR錯誤，則將SMAS裝置100連同判定性錯誤校正一起使用可導致真實序列與識別的序列之間的完全一致。此外，若僅發生FNI及FLR錯誤，則實際上可僅使用讀取單個ssDNA之單個感測器105以及上文論述的判定性錯誤校正技術(例如將FLR更改為「未偵測到標記」及/或自偵測結果之記錄刪除若干串指定長度(例如4)「未偵測到標記」)來識別無錯誤序列。

【0206】 然而，當引入FNR及/或FDL錯誤時，僅使用判定性錯誤校正一般不大可能消除偵測結果記錄中之所有錯誤。為了解決FNR及/或

FDL錯誤，除了判定性錯誤校正之外或替代判定性錯誤校正，可包括概率性錯誤校正。

【0207】

減輕FNI、FLR及FNR錯誤

本部分進一步包括分析中之FNR錯誤。此類錯誤於CLUS裝置的鹼基識別準確度之影響因為CLUS裝置偵測核酸例項簇中之標記時固有的平均化而等效於FNI及FLR之影響。FNR錯誤明顯更不利於使用SMAS裝置100之定序方法之性能，因為不能判定性地校正FNR錯誤。(應注意的是，CLUS裝置本身根本無法校正FNR錯誤。相反地，CLUS裝置仰賴於集體行為以減輕FLR及其他類型之錯誤之影響。)

【0208】 圖27說明由示例性序列(**TAG CAA GGT CCG CTA CTG GCA GAC TGG**)中之FNR錯誤引入之問題，假設FNI、FLR及現在亦有FNR錯誤在18個**A**?⇒**T**?⇒**C**?⇒**G**?查詢步驟之查詢循環期間隨機發生。出於實例之目的，假設$K = 3$ (亦即三個結合位點116中之各者保持特定ssDNA之例項，且三個各別感測器105中之各者感測三個ssDNA例項中之各別一者)，平均100次反應中有15次失敗($r = 0.15$，此係很大的化學失敗率)，及該等錯誤平均分配於FNI錯誤、FLR錯誤及FNR錯誤之間。在此處做出的實例條件及假設下，僅給定藉由SBS使用SMAS裝置100建立的資料記錄，就不可能在資料記錄中區分正確地偵測事件(圖27中之實心圓)與FNR (帶有正斜杠填充之圓圈)。圖28說明當在感測器S1、S2、S3中超過一半(3個中有至少2個)偵測到標記下識別鹼基時之結果。儘管可判定性地校正FLR錯誤(如上文所述，藉由將其視為「未偵測到標記」)，但無法識別FNR錯誤，因為其與正確標記偵測事件無法區分。因此，在本實例中，

17個中僅8個(約47%)識別的鹼基符合真實序列。因此，引入FNR錯誤使得判定性FNI錯誤校正更具挑戰性，因為FNR錯誤破壞該串四個或更多個「未偵測到標記」偵測結果，否則該等偵測結果可能已被移除。若未經處理地藉由刪除若干串四個零以嘗試將不良股與不受錯誤影響之股比對來實施FNI錯誤校正，則定序精度不會提高。實際上，如圖29中所顯示，對於本實例，鹼基識別精度似乎變差，因為在移除該等若干串四個「未偵測到標記」偵測結果後，20個中僅9個(45%)鹼基識別與真實序列一致。

【0209】 錯誤校正可藉由應用概率性錯誤校正來改良以減輕FLR及FNI錯誤之外的FNR錯誤。例如，應注意在位置2處之胸腺嘧啶查詢步驟(查詢循環1之查詢步驟2)。感測器S1及S3偵測到標記，但S2不能偵測到。由於在感測器S1及S3處同時發生FNR錯誤，或由於在感測器S2處發生FNI錯誤，故S2無法偵測到標記。假設各錯誤之概率為r，在感測器S1及S3處同時發生FNR錯誤之概率為r^2，及在感測器S2處FNI錯誤之概率為r。錯誤校正算法(例如藉由至少一個處理器130或另一個處理器進行)假設發生更可能的事件(在感測器S2處存在FNI錯誤)且自捕獲來自感測器S2之偵測結果之資料記錄刪除位置2至5中將S2偵測結果移位於S2記錄中之所有條目。因此，將S2記錄中之偵測結果與藉由感測器S1及S3產生的偵測結果再比對，如圖30之標記為「A」之上部中所顯示。先前(刪除前)在位置4 (在圖30之標記為「A」之部分中)處之**G**標記偵測現可歸因於FNR，此乃因感測器S1及S3在位置4中未偵測到標記(查詢循環1之查詢步驟4)。

【0210】 可在位置13 (如圖30之標記為「B」之部分中所顯示)、32(標記為「C」)及46 (標記為「D」)處自左至右進行相同錯誤校正程序以顯示偵測結果之S1、S2及S3記錄之間的比對之逐漸改良，如圖30之標記

為「E」之部分中所說明。圖30之標記「E」之部分指示儘管實施多個概率性錯誤校正步驟將所有感測器S1、S2及S3之輸出比對，但似乎並未改良識別的序列與真實序列之間的比對。甚至在錯誤校正後，20個中僅9個(45%)鹼基得以正確識別。換言之，仍發生鹼基識別錯誤。具體而言，錯誤校正程序之後，所有三個感測器S1、S2及S3均報告在應偵測到標記之查詢步驟已偵測到標記，但該等感測器中之一些亦偵測到在位置10、22、40及50(顯示於圖30之連續視圖中)處藉由FNR不正確併入的標記。

【0211】 當超過一半的感測器105之偵測結果一致時(錯誤校正之後)識別鹼基導致在序列位置8 (查詢步驟22)處之胸腺嘧啶插入錯誤，其中感測器S1及S3均偵測到在相同查詢步驟期間結合至非互補核苷酸之標記。(應瞭解，可知曉在位置8處存在胸腺嘧啶插入錯誤之原因是因為錯誤化資料係出於說明之目的而建立且係已知的。在一個實施案中，感測器105僅指示在查詢步驟期間是否偵測到標記，而不指示該偵測(或缺少偵測)是正確的還是錯誤的。因此，在一個實施案中，在查詢步驟22處之錯誤將基本上無法與正確地偵測結果區分。) 清楚地展現單個錯誤鹼基插入之位置之正確比對的真實序列及識別的序列可呈現為：

錯誤： ***\|插入***
真實序列： TAG CAA G*G TCC GCT ACT GGC
識別的序列： TAG CAA GTG TCC GCT ACT GGC
*插入位置

【0212】 若鹼基識別規則經修改以要求所有三個感測器S1、S2及S3均一致，則可校正此種插入錯誤。就此一規則而言，所有三個感測器S1、S2及S3必須同時遭遇FNR錯誤以導致錯誤的鹼基識別。此一事件之概率僅為r^3。假設$r = 0.05$，在相同查詢步驟期間所有三個感測器S1、S2及S3均遭遇FNR事件，平均100,000次查詢中僅125次(或0.000125之概率)，甚

至對於用於當前實例中之極高錯誤率，此係極低的。然而，若亦發生FLD錯誤，則實施此一規則可導致不正確的識別，如下文進一步論述。

減輕FNI、FLR、FNR及FLD錯誤

【0213】用於一些實施例中之一般錯誤校正策略解決且減輕導致FNI、FLR、FNR及FLD錯誤之所有四種類型之化學失敗。圖31說明示例性序列(**TAG CAA GGT CCG CTA CTG GCA GAC TGG**)，假設FNI、FLR、FNR及現在亦有FLD錯誤在18個**A**?⇒**T**?⇒**C**?⇒**G**?查詢步驟之查詢循環期間隨機發生。出於在定序資料中建立許多錯誤以提供媒體來說明示例性錯誤校正程序之目的，假設極高平均錯誤率為5次中有1次失敗之反應($r \cong 0.2$、或20%錯誤率)，且亦假設錯誤在FNI錯誤、FLR錯誤、FNR錯誤及FLD錯誤之間平均分配。因此，100次反應中有約20次失敗，且該等失敗在四種錯誤類型之間相等分配。應瞭解，此一高錯誤率實務上不太可能發生，且因此此處所考慮的實例之難度可能比將在真實世界實施案中遭遇到的難度高得多。

【0214】在此處做出的實例條件及假設下，僅給定藉由SBS使用SMAS裝置100建立的資料記錄，就不可能區分正確核苷酸併入與FNR，亦不能區分正確核苷酸非併入與FNI。儘管可如前面所描述判定性地偵測及校正FLR錯誤(藉由在切割且沖洗掉標記後檢查感測器105，且將FLR視為「未偵測到標記」)，但無法識別FNR錯誤，因為其無法與正確地偵測事件區分，且無法識別FNI及FLD錯誤，因為其無法與未併入正確核苷酸區分。然而，仍可使用概率性錯誤校正技術來完成錯誤減輕。例如，如上文所說明，當在一個特定查詢步驟期間少於所有感測器S1、S2及S3偵測到或未偵測到標記時，可計算兩個(或更多個)事件之概率，該具有最高概

率之事件可假設為是正確事件，且可採用適宜錯誤校正步驟。

【0215】 圖32說明在上文描述的條件及假設下將錯誤校正程序應用於在SBS期間捕獲之資料。圖32之標記為「A」之部分係移除FLR錯誤之前的原始資料。假設如上文所述，在切割且沖洗掉標記之後檢查感測器105信號位準，已知FLR錯誤之位置。FLR錯誤可使用判定性錯誤校正來完全移除，亦即藉由將對應於偵測到FLR錯誤之查詢步驟之位置中之資料記錄中之「偵測到標記」值(例如1或「是」)更改為「未偵測到標記」(例如0或「否」)值。應注意，在顯示於圖31中之查詢循環15期間，在感測器S2之資料中之FLD錯誤之後係FLR錯誤。換言之，在第15個查詢循環之第一查詢步驟期間感測器S2無法偵測到併入的核苷酸之標記。當在第15個循環之第一查詢步驟之後，且在第15個查詢循環之第二查詢步驟之前切割標記時，檢查感測器S2之信號位準。此種檢查顯示在感測器S2處存在標記，此將被認為是FLR錯誤，因為在最後一個查詢步驟之後應已切割且沖洗掉所有標記。因此，甚至在FLR錯誤跟隨另一個錯誤時，其亦係可偵測的且可被移除。

【0216】 圖32之標記為「B」之部分顯示經由判定性錯誤校正移除FLR錯誤後之偵測結果之記錄，如前面所述進行應用。顯示於「B」中之資料記錄現僅包含藉由感測器S1、S2、S3中之各者在所顯示的(4×18)查詢步驟中之各者處偵測到或未偵測到標記之指示。(應瞭解，該等記錄可比圖32中所顯示短或長。)如上文所說明，自此等記錄並不知曉哪些「偵測到標記」指示是正確的及哪些是FNR錯誤，且並不知曉哪些「未偵測到標記」指示是正確的及哪些是FNI或FLD錯誤。因此，概率性錯誤校正可用於估計序列。

【0217】 為了說明如何可應用概率性錯誤校正，下表顯示圖32之在已移除FLR錯誤(例如自圖32中之標記為「B」之記錄)後三個感測器S1、S2及S3之前五個查詢循環(查詢步驟1至20)之資料記錄。換言之，下表顯示判定性錯誤校正移除FLR錯誤後之前20個偵測結果。對於感測器偵測到標記之查詢步驟，該表包含值1，及對於感測器未偵測到標記之查詢循環，該表包含值0：

步驟	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
	A	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**
S1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0
S2	0	0	0	1	0	1	0	0	0	0	0	1	0	0	1	0	1	0	0	0
S3	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0

【0218】 如上文所說明，在移除FLR錯誤之後的簡單大多數表決將導致正確識別17個鹼基中僅8個，如圖32之標記為「B」之部分中所顯示。概率性錯誤校正如下文所述可提供顯著改良。

【0219】 考慮到作為一個實例之查詢步驟2，感測器S1及S3均偵測到標記(上表中之條目為1)，但感測器S2未偵測到標記(表條目為0)。因此，感測器S1及S3係錯誤的，或感測器S2係錯誤的。藉由考慮可導致此等結果中之各者之各種事件之概率，錯誤校正算法可減輕定序資料中之錯誤。具體而言，因為已自資料記錄移除FLR，故感測器S1及S3在查詢步驟2期間均不正確地偵測到標記之唯一方法係兩者在該查詢步驟期間是否均遭遇FNR錯誤。若FNR錯誤之概率為r，則感測器S1及S3在單個查詢步驟期間均遭遇FNR錯誤之概率為r^2。出於本實例之目的，假設$r = 0.2$之高錯誤率，且因此感測器S1及S3在查詢步驟2期間均不正確地偵測到標記之概率為0.04。

【0220】 若感測器S2係錯誤的，則是因為感測器S2由於FLD錯誤或FNI錯誤而無法偵測到標記。回想一下，當併入正確互補核苷酸，但缺失

標記或感測器無法偵測到其標記時發生FLD錯誤，及當在定序循環期間根本沒有併入正確互補核苷酸時發生FNI錯誤。FLD及FNI錯誤係相互排他的(亦即，感測器一次僅可遭遇其中之一者，而從不會是兩者)。因此，假設各類型之錯誤之概率為r，感測器S2遭遇FLD錯誤或FNI錯誤之概率為$2r$。對於此處的實例，已假設$r = 0.2$之高錯誤率，因此在查詢步驟2期間感測器S2係錯誤之概率為0.4。將感測器S2在查詢步驟2期間係錯誤之概率與感測器S1及S3均係錯誤之概率進行比較，因為$0.4 \gg 0.04$，故感測器S2係錯誤之可能性更大。在一些實施例中，錯誤校正算法假設發生更可能之事件，意指假設感測器S2係錯誤的，且丟棄感測器S1及S3均係錯誤之概率且不做進一步考慮。

【0221】 如上文所說明，感測器S2由於FLD錯誤或FNI錯誤中任一者而可能係錯誤的。在FLD錯誤之後，藉由感測器S2感測到的DNA股將與藉由感測器S1及S3感測到的DNA股保持「同步」或「比對」。換言之，若查詢步驟m定序藉由感測器S1、S2及S3中之各者感測到的DNA股之第40個鹼基，則查詢步驟$m + 1$將定序各股之第41個鹼基，即使感測器中之一者(例如感測器S2)在查詢步驟m期間遭遇FLD錯誤。另一方面，FNI錯誤之後果係藉由遭遇FNI錯誤之感測器感測到的DNA股與藉由未遭遇FNI錯誤之感測器感測到的DNA股「不同步」或變成「錯誤比對」。在當前實例中，若在查詢步驟2處之錯誤係由於FNI所致(例如，其將「位於」藉由感測器S1及S3以四個查詢步驟感測到的DNA股「後面」，此將係下一次互補核苷酸之併入)，則藉由感測器S2感測到的DNA股將與藉由感測器S1及S3感測到的DNA股變成不同步。

【0222】 在一些實施例中，藉由錯誤校正算法採取的動作部分取決

於候選錯誤經校正資料之檢查，該檢查分別假設兩種類型之錯誤中之各者已發生。換言之，可修改偵測結果之記錄以校正錯誤，假設錯誤係由於FLD錯誤引起，以產生第一候選經校正資料記錄，且可分別修改該偵測結果之記錄以校正錯誤，假設錯誤係由FNI錯誤引起，以產生第二候選經校正資料記錄。可然後檢查及/或分析及/或比較兩個候選經校正資料記錄以判定哪個更可能是正確的。為了校正FLD錯誤，將「未偵測到標記」指示翻轉為「偵測到標記」指示。為了校正FNI錯誤，將資料條目移位四個位置(例如至左側作為資料記錄呈現於本文實例中)。

【0223】 為了說明實例資料記錄中查詢步驟2之特定實例，第一候選經校正資料記錄選項A假設影響感測器S2的輸出之(假定的)錯誤係FLD錯誤。藉由將感測器S2的記錄中之查詢步驟2之位元自0翻轉為1來校正假定的錯誤，如以下表選項A中以粗體、加底線值「1」所顯示：

步驟	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
	A	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**
S1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0
S2	0	**1**	0	1	0	1	0	0	0	0	0	1	0	0	1	0	1	0	0	0
S3	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0

【0224】 第二候選經校正資料記錄選項B假設影響感測器S2的輸出之錯誤係FNI錯誤。該假定的錯誤藉由自感測器S2資料條目刪除在查詢步驟2、3、4及5期間記錄的資料以使對應於感測器S2之資料記錄與感測器S1及S3之資料記錄「再同步」或「再比對」來校正，此得到下表(原來在位置21至24處之值移位至位置17至20中)。藉由錯誤校正算法修改之選項B表條目以粗體、加底線字體顯示：

步驟	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
	A	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**
S1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0
S2	0	**1**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**1**	**0**	**0**	**1**
S3	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0

【0225】 可然後比較及/或分析選項A及B以判定哪個更可能是正確的，且可丟棄該等選項中之一者。例如，處理器(例如至少一個處理器130或另一個處理器)可判定各候選經校正資料記錄之度量值且至少部分地基於度量之比較來決定選項A及B中哪個更可能是正確的。度量之一個實例係自經現在校正之當前查詢步驟之後的查詢步驟開始之查詢步驟數及在所有三個(或更一般而言，*K*)感測器的標記偵測結果一致之資料記錄中更遠的查詢步驟*J*位置。例如，使用此度量，且將值*J*設置為8，選項A之度量值為3，及選項B之度量值為6。在一些實施例中，僅基於此結果，假設因為選項B之度量值顯著大於選項A之度量值，故選項B更可能是正確的，且丟棄選項A。在一些實施例中，丟棄兩個選項中之一者，唯若其度量值超過另一選項的度量值某個臨限值(例如百分比、量(例如至少兩倍、至少1.5倍大等)等)。在一些實施例中，保留選項A，且直至稍後才丟棄選項。

【0226】 在一些實施例中，基於自經現在校正之當前查詢步驟考慮的資料之距離來加權對度量值的貢獻。例如，因為已引入資料記錄中之另外錯誤之可能性隨著更多鹼基被定序而增加(例如在查詢步驟3與查詢步驟40之間*K*個感測器中之一者發生某種類型之錯誤之可能性大於在查詢步驟3與查詢步驟6之間*K*個感測器中之一者發生某種類型之錯誤之可能性)，故度量可假設更近資料條目比更遠資料條目更可能是正確的，且因此較彼等更遠資料條目，對更接近經現在校正資料之資料條目提供更多權重。加權可為例如線性或非線性的。僅作為一個實例，對於資料貢獻多達12個查詢步驟遠之度量，經現在校正之資料之四個查詢步驟內的查詢步驟貢獻可賦予權重1，經現在校正之資料之五個至八個查詢步驟之查詢步驟貢獻可賦予權重0.5，及經現在校正之資料之九個至十二個查詢步驟之查詢步驟貢

獻可賦予權重0.2。應明瞭，可使用許多可能的度量，無論具有或不具有加權，及彼等上文提供之度量僅係示例性的且無意為限制性的。

【0227】 亦應明瞭，儘管上文描述的度量使用自經現在校正之當前查詢步驟之後的查詢步驟開始之查詢步驟數及在所有三個(或更一般而言，*K*個)感測器的標記偵測結果一致之資料記錄中更遠之查詢步驟*J*位置，但其可等效地使用自經現在校正之當前查詢步驟之後的查詢步驟開始之查詢步驟數及在所有三個(或更一般而言，*K*個)感測器的標記偵測結果不一致之資料記錄中更遠之查詢步驟*J*之位置。在此種情況下，大的度量值將指示感測器資料條目之間更多的不匹配，且因此候選經校正資料記錄對於較低度量值將更可能是正確的。如一般技術者所可明瞭，可對欲應用的任何加權進行調整。

【0228】 亦應明瞭，在資料記錄中之假定的錯誤之校正之後，不必丟棄可能選項中之一者。例如，在感測器S2的記錄中之查詢步驟2處之(假定的)錯誤之(假定的)校正之後，選項A及B兩者皆可保留，且於兩者上並行進行進一步之錯誤偵測及校正。同樣地，每次校正假定的錯誤，可判定及/或評定/比較候選序列之多個選項。可在錯誤校正程序之各步驟處維持各可能選項/候選序列之運行度量值，且可在某個點(例如在已判定且評估所有候選選項(例如相對於彼此)之後，或在一些另外數量之查詢步驟之後等)判定最可能之候選序列。

【0229】 此外，儘管在上文實例中，立即丟棄感測器S1及S3兩者錯誤偵測標記之概率，因為該事件之概率(給定本文的假設)顯著低於該感測器S2係錯誤之概率，可替代地遵循與針對感測器S2相同的程序。換言之，可判定在查詢步驟2處之選項C，假設感測器S1及S3兩者均遭遇FNR

錯誤，且感測器S2是正確的。在此種情況下，可調整度量以說明各種可能結果之可能性(例如藉由基於感測器S1及S3同時遭遇FNR錯誤之概率「懲罰」選項C之度量(例如將度量乘以兩個感測器S1及S3均係錯誤之概率與感測器S2係錯誤之概率之比率))。

【0230】 應明瞭，本文描述的錯誤校正方法可以多種方式利用以改良使用SMAS裝置100之核酸定序之準確度。假設足夠的計算能力，實施案(例如使用至少一個處理器130或另一個處理器或處理器)可判定且評估應用錯誤校正之詳盡候選序列組，且然後從其當中選擇最可能是正確之候選序列。為了降低計算複雜度，實施案亦可在錯誤校正過程期間作出決策以消除被認為不太可能是正確之候選錯誤經校正序列(或潛在錯誤來源)(例如上文實例中之選項C)且僅保留彼等更可能是正確之候選錯誤經校正序列。應明瞭，所揭示的原理之靈活性使得其適於具有多種計算能力之系統中之錯誤減輕。

【0231】 返回上文實例，假設選項B係在將錯誤校正應用至來自查詢步驟2之資料之後保留的唯一選項，經校正資料顯示如下：

步驟	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
	A	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**	**A**	**T**	**C**	**G**
S1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0
S2	0	1	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1	0	0	1
S3	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0

【0232】 三個感測器S1、S2及S3不一致之下一個查詢步驟係在查詢步驟5處。再一次，感測器S2與感測器S1及S3以與查詢步驟2中相同之方式不一致。在一些實施例中，錯誤校正算法判定(a)感測器S2係錯誤之概率係大於感測器S1及S3兩者均係錯誤之概率，及(b)感測器S2在查詢步驟5處遭遇FNI錯誤或FLD錯誤。再一次，可建立兩個選項，一個選項假設錯誤係FLD錯誤(藉由翻轉位元來校正)，及另一個選項假設錯誤係FNI (藉

由將資料移位四個位置來校正)。經校正之資料記錄顯示如下：

選項A (假定FLD錯誤經校正)：

步驟	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G
S1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0
S2	0	1	0	0	**1**	0	0	1	0	0	1	0	1	0	0	0	1	0	0	1
S3	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0

選項B (假定FNI錯誤經校正)：

步驟	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G
S1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0
S2	0	1	0	0	**0**	**0**	**1**	**0**	**1**	**0**	**0**	**0**	**1**	**0**	**0**	**1**	**0**	**0**	**0**	**1**
S3	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0

【0233】 再一次，可計算選項A及B之度量，且可丟棄該等選項中之一者，或可保留兩者。為了實例起見，假設保留選項A，產生以下錯誤經校正資料：

步驟	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G
S1	0	1	0	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0
S2	0	1	0	0	1	0	0	1	0	0	1	0	1	0	0	0	1	0	0	1
S3	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0

【0234】 感測器的資料不一致之下一個查詢步驟係查詢步驟10。此處，感測器S1偵測到標記，但感測器S2及感測器S3均未偵測到標記。因為已自資料記錄移除FLR錯誤，故感測器S1在查詢步驟10期間錯誤地偵測到標記之唯一方法係其在該查詢步驟期間是否遭遇FNR錯誤。FNR錯誤之概率為r。若感測器S2及S3均為錯誤的，則是因為(a)兩者均遭遇FNI錯誤，(b)兩者均遭遇FLD錯誤，或(c)其中之一者遭遇FNI錯誤及另一者遭遇FLD錯誤。相互排他之事件(a)、(b)或(c)中之任何者之概率為$4r^2$。因此，在一些實施例中，假設發生更可能之事件，亦即，該感測器S1遭遇FNR錯誤(因為對於假設的r值，$r \gg 4r^2$)。如上文所說明，FNR錯誤可藉由

將資料條目自「偵測到標記」值翻轉為「未偵測到標記」值來校正，此得到下表：

步驟	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G	A	T	C	G
S1	0	1	0	0	1	0	0	1	0	**0**	1	0	1	0	0	0	1	0	0	0
S2	0	1	0	0	1	0	0	1	0	0	1	0	1	0	0	0	1	0	0	1
S3	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0

【0235】 該錯誤校正程序可如所述繼續在整個其餘資料記錄中進行。圖32之標記為「C」之部分顯示該實例之結果。如所示，在應用如上文所述的概率性錯誤校正之後，正確地識別20個中的16個(80%)鹼基。

【0236】 圖33為說明根據一些實施例之錯誤校正程序450之流程圖。錯誤校正程序450可為例如說明於圖11中之錯誤校正程序212，且其可藉由處理器(例如說明於下文論述的圖5A或圖50中之至少一個處理器130)進行。在452處，錯誤校正程序450開始。在454處，在由於使用SMAS裝置100之核酸定序程序而產生之定序資料中識別複數個記錄。識別的複數個記錄中之各者包含複數個條目，該複數個條目中之各者捕獲核酸之特定股之一個例項之偵測結果。因此，若識別的記錄數為K，則該K個記錄中之各者包含一個條目/偵測結果/定序程序查詢步驟。各偵測結果指示，在查詢步驟期間，(a)藉由相應感測器105偵測到標記，或(b)藉由相應感測器105未偵測到標記。該複數個記錄可以多種方式來識別。例如，如下文所進一步說明，可將不同獨特條碼拼接至核酸股之引物端使得在定序程序之循環期間讀取已知序列。因此，該複數個記錄可藉由搜索與核酸之特定股相關之條碼之定序資料來識別。作為另一個實例，可在定序資料中(例如在記錄定序程序之前約35個查詢步驟之偵測結果之條目內)識別條目之共同序列。

【0237】 在456處，基於該複數個記錄，判定核酸之特定股之複數

個候選序列。該複數個候選序列中之各者估計核酸之特定股之核酸序列之至少一部分(例如，少至一個鹼基)。在一些實施例中，判定該複數個候選序列包括在該複數個記錄內識別特定查詢步驟，在該特定查詢步驟處，第一感測器偵測到各別標記及第二感測器未偵測到任何標記；及確立兩個候選序列，該兩個候選序列中之一者假設該第一感測器正確地偵測到各別標記及該兩個候選序列中之第二者假設該第一感測器不正確地偵測到各別標記。在一些實施例中，判定該複數個候選序列包括在該複數個記錄內識別特定查詢步驟，在該特定查詢步驟處，第一感測器偵測到各別標記及第二感測器未偵測到任何標記；及確立兩個候選序列，該兩個候選序列中之一者假設該第二感測器不正確地未偵測到任何標記及該兩個候選序列中之第二者假設該第二感測器正確地未偵測到任何標記。在一些實施例中，判定該複數個候選序列包括在該複數個記錄中之至少一者中識別指示未偵測到標記之一組連續條目(例如四個條目)，及自該複數個記錄中之至少一者刪除指示未偵測到標記之該組連續條目。在一些實施例中，該複數個條目中之各者係第一二進制值(指示偵測到標記)或第二二進制值(指示未偵測到標記)，及判定該複數個候選序列包括在該複數個記錄中之至少一者中識別一串(例如四個)第二二進制值，及自該複數個記錄中之至少一者刪除該串第二二進制值。

【0238】 在458處，該複數個候選核酸序列中之特定候選序列經識別為自該複數個候選序列當中最可能是正確之序列。在一些實施例中，識別該複數個候選序列中最可能是正確之特定候選序列包括判定或估計該複數個候選序列中哪個具有是正確之最高概率。在一些實施例中，識別該複數個候選序列中最可能是正確之特定候選序列包括判定該等候選序列中之

各者之各別度量，且至少部分地基於各別度量及標準(例如最小發生可能性、臨限發生可能性)，選擇特定候選序列作為最可能是正確者。在一些實施例中，識別該複數個候選序列中最可能是正確之特定候選序列包括識別由複數個記錄表示之特定查詢步驟之大多數結果(例如感測器105中超過一半偵測到標記或感測器105中超過一半未偵測到標記)。在一些實施例中，識別該複數個候選序列中最可能是正確之特定候選序列包括判定該複數個候選序列中之各者之各別發生可能性，及基於其滿足約束之各別發生可能性(例如最小概率)來選擇特定候選序列。在一些實施例中，候選序列中具有最高發生可能性之特定候選序列經識別為最可能是正確者。在一些實施例中，該等候選序列中之一者或多者係基於已知約束諸如鹼基之特定序列係不可能之知識來消除。例如，自核酸之起源或來源(例如人類)可知曉鹼基之特定序列係不可能的，且因此可藉由進一步考慮消除具有此種不可能之序列之候選序列。

【0239】 在460處，錯誤校正程序450結束。

【0240】 應瞭解，僅當識別的最可能之情境(例如在圖33之458處之識別)實際上是正確者時，概率性錯誤校正才成功。若化學失敗率很高，如在本文描述的實例中，則可存在同樣可能發生之多個情境(或其發生概率彼此接近)，在該情況下，可採用更複雜的生物資訊工具。例如，候選序列可基於所定序核酸之來源之知識(例如基於給定核酸之來源/起源情況下，鹼基之特定序列係不可能之知識)來消除。然而，若如本文所述正確地實施，則該錯誤校正過程導致感測器105輸出之正確比對。在顯示於圖32中之實例中，在移除FNI及FLR之後，所有三個感測器S1、S2及S3均在應偵測到標記之正確偵測查詢步驟報告標記，但該等感測器在許多查詢位

置(5、10、13、20、22、27、32、40、41、48及50)不一致，其中感測器偵測到藉由FNR不正確併入之標記或由於FLD而無法偵測到標記。當在比對的序列中感測器105中超過一半一致時識別鹼基導致在序列位置8處之胸腺嘧啶插入(查詢步驟22)及在位置13處之鳥嘌呤刪除(查詢步驟32)。清楚地展現鹼基插入及刪除位置之正確比對的真實序列及識別的序列可呈現為：

錯誤：				***插入***		***刪除***			
真實序列	TAG	CAA	G	*	G	TCC	**G**	CT	ACT	GGC
識別的序列：	TAG	CAA	G	**T**	G	TCC	*	CT	ACT	GGC

【0241】 如有鑑於本文揭示內容所瞭解，巧合的FNR及FLD導致插入及刪除錯誤不能以算法方式校正且若不知曉真實序列則將保持不被發現。換言之，當比對的序列中單分子感測器105中超過一半給出錯誤答案時，不正確地識別鹼基。此類事件之概率取決於化學失敗發生之比率(r值)。如上文所說明，本文呈現的實例使用高錯誤率以便說明錯誤校正技術之應用。實際實施案中之錯誤率應顯著降低，藉此減少錯誤校正程序不能夠校正錯誤之可能性。所揭示的錯誤校正技術可用於在查詢步驟處正確地比對多個感測器105輸出。此可使用對可能錯誤類型之物理起源(例如某些序列對於源核酸係不可能之知識)、其平均發生率及其在感測器序列輸出中之記號之深刻理解來達成。若化學錯誤率很高且錯誤之記號被遮蓋，則錯誤校正算法可為計算密集且難以實施。下文論述描述不正確鹼基識別之概率如何取決於讀取長度、簇大小N (對於CLUS裝置)、感測相同核酸股之例項之感測器之數量K(對於SMAS裝置100)及失敗之化學錯誤率。

簇定序儀之一般定量結果

【0242】 本文開發一種簡單定量模型，其用於估計採用上文引入的改進型加性定序方案之簇定序儀中不正確鹼基識別之概率。假設各種類型

之錯誤(FNI、FLR、FNR及FLD)在整個簇中以比率r隨機發生，其中$0 < r < 1$。最初，簇股彼此同相(in-phase) (例如同步、比對、不同步)，且偵測到的信號與簇大小(N)成比例。當引入且成功併入互補標記之核苷酸時，偵測到信號。當在具有**A?⇒T?⇒C?⇒G?**查詢步驟之查詢循環期間引入非互補核苷酸時，應偵測不到信號。錯誤以比率r發生，此導致逐漸增加之股數與集體平均異相(不同步)。此在併入互補核苷酸時降低集體信號之強度(或幅度)且在引入非互補核苷酸時增加背景信號之強度或幅度。在因為引入且成功併入匹配之核苷酸(ON-State)而應偵測到標記之查詢步驟處之平均信號強度由以下給出：

$$\langle 1\rangle = \frac{N}{2}\left(1 + e^{-r\times C}\right). \qquad \text{(方程式1(a))}$$

其中C係偵測查詢步驟(或數量)。類似地，在因為引入非互補核苷酸(OFF-State)而不應偵測到標記之查詢步驟處之強度由以下給出：

$$\langle 0\rangle = \frac{N}{2}\left(1 - e^{-r\times C}\right) \qquad \text{(方程式1(b))}$$

【0243】 此背景信號係由異相核酸股產生，該異相核酸股併入與集體平均之同相位置不互補之核苷酸。方程式1(a)及(b)之函數繪製於圖34A中，對於$N = 11$及$r = 0.1$。圖34B說明函數如何擬合先前描述的簇模型實例之強度測量值。如所說明，正確識別鹼基直至$C \cong 15$，但在較大C值處發生頻繁錯誤。

【0244】 如藉由圖34A及34B所說明，在早期定序查詢(C小)期間，⟨1⟩及⟨0⟩狀態完全分開，但其遵循由方程式1(a)及(b)表示之函數形式快速接近平均值$N/2$。再者，因為錯誤發生係隨機無關事件，故兩種狀態之信號測量值圍繞其集體平均值⟨1⟩及⟨0⟩離散分佈。具體而言，當集體平均為⟨1⟩時簇大小N之ON-State強度測量值為k之概率藉由泊松分佈(Poisson

distribution)來給出：

$$P_{\langle 1\rangle}(k) \cong e^{-\langle 1\rangle} \times \frac{\langle 1\rangle^{N-k}}{(N-k)!} \qquad \text{(方程式2(a))}$$

類似地，當集體平均為⟨0⟩時相同簇之OFF-State強度記錄值為k之概率為：

$$P_{\langle 0\rangle}(k) \cong e^{-\langle 0\rangle} \times \frac{\langle 0\rangle^{k}}{k!} \qquad \text{(方程式2(b))}$$

57

【0245】 概率函數$P_{\langle 1\rangle}(k)$及$P_{\langle 0\rangle}(k)$，$N = 11$, $r = 0.1$及$C = 0, 5, 10, 15$及20，繪製於圖35中。該圖顯示兩種泊松分佈，且尾部隨著C增加而越來越多地重疊。在兩種離散分佈下$P_{\langle 0或1\rangle}(k)$之所有可能值之總和等於1：

$$\sum_{k=0}^{k=N} P_{\langle 0\ 或\ 1\rangle}(k) \cong \sum_{k=0}^{k=\infty} P_{\langle 0\ 或\ 1\rangle}(k) = 1 \qquad \text{(方程式3)}$$

【0246】 當將ON-State誤認為OFF-State或反之亦然時，發生鹼基識別錯誤。圖36說明在不同定序查詢步驟$C = 0, 5, 10, 15$及20處ON-State $P_{\langle 1\rangle}(k)$及OFF-State $P_{\langle 0\rangle}(k)$ ($N = 11$及$r = 0.1$)之離散概率函數。不正確鹼基識別之來源在$P_{\langle 0\rangle}(k)$之尾部延伸高於$N/2$中間值時以帶圖案的點顯示或在$P_{\langle 1\rangle}(k)$延伸低於$N/2$中間值時以虛線圓圈顯示。在ON-State分佈之尾部顯著延伸低於$k \cong N/2$(在圖36中，不正確⟨1⟩)或OFF-State分佈之尾部(不正確⟨0⟩)延伸高於$k \cong N/2$時，進行錯誤鹼基識別之概率變得很大。

【0247】 圖37A顯示平均ON-State及OFF-State強度圖與C ($r = 0.1$)及$N = 11$ (頂部)及$N = 101$(底部)之簇大小之函數關係。圖37B說明在$C = 1, 10, 20, 30$及40 ($r = 0.1$)及$N = 11$ (頂部)及$N = 101$ (底部)之簇大小下之OFF-State概率分佈函數$P_{\langle 0\rangle}(k)$。增加簇大小藉由減小$P_{\langle 0\rangle}(k)$分佈之相對寬度(此增加距$P_{\langle 1\rangle}(k)$之距離)而延遲鹼基判定錯誤之發生。

【0248】 一般而言，在定序查詢數量C處不正確鹼基識別之概率(對

於簇大小N及化學失敗率r)(表示為$P_{C,N,r}$)係不正確識別OFF-State之概率之總和，亦即，對於超過$k=(N+1)/2$之k值，其係$P_{\langle 0\rangle}(k)$值之總和。此等係圖36及37B中之帶圖案的點。增加簇大小N增加兩個離散分佈峰之間的初始間隔且延遲鹼基識別錯誤之發生。為了簡化進一步論述，僅考慮簇大小N為奇數之情況以避免當偵測到的信號為$N/2$(其既不是ON-State亦不是OFF-State)時引入之不判定性。對於N之奇數值，$P_{C,N,r}$由以下給出：

$$P_{C,N,r} \cong \sum_{k=\frac{N+1}{2}}^{k=N} P_{\langle 0\rangle}(k) = \sum_{k=\frac{N+1}{2}}^{k=N} \left\{ e^{\frac{N}{2}(e^{-r\times C}-1)} \times \frac{\left[\frac{N}{2}(e^{-r\times C}-1)\right]^{k}}{k!} \right\} \quad \text{(方程式4(a))}$$

【0249】 替代地，$P_{C,N,r}$係不正確識別ON-State之概率之總和，亦即對於k低於$k=(N-1)/2$之值，其為$P_{\langle 1\rangle}(k)$值之總和(圖36中的帶有反斜杠填充之圓圈36)，其由以下給出：

$$P_{C,N,r} \cong \sum_{k=0}^{k=\frac{N-1}{2}} P_{\langle 1\rangle}(k) = \sum_{k=0}^{k=\frac{N-1}{2}} \left\{ e^{\frac{N}{2}(e^{-r\times C}+1)} \times \frac{\left[\frac{N}{2}(e^{-r\times C}+1)\right]^{N-k}}{(N-k)!} \right\} \quad \text{(方程式4(b))}$$

【0250】 圖38A及38B繪製方程式4(a)及4(b)與N及r之各種組合之C之函數關係。圖38A繪製計算$P_{C,N,r}(C)$函數，$r=0.1$及$N=11,51,101$及151，及圖38B繪製計算$P_{C,N,r}(C)$函數$P_{C,N,r}$，$N=101$及$r=0.1,0.05$及0.01。該等圖顯示在各種臨限值C_{th}下不正確鹼基識別之概率之顯著增加率。如圖38A及38B指示，隨著C趨於無窮大，$P_{C,N,r}$接近0.5 。圖38A及38B中之圖顯示分析分子集體之定序儀(例如CLUS裝置)之行為特性。當C很小時，不正確鹼基識別之概率($P_{C,N,r}$)仍然很低，但其在特定臨限值(C_{th})處顯著增加，該臨限值係由N及r參數之量值決定。隨著C趨於無窮大，$P_{C,N,r}$接近0.5 ，在此點時，ON-State之強度等於OFF-State之強度，且有二分之一的機率進行不正確鹼基識別。$P_{C,N,r}$ 很大程度上取決於此三個參數

C,N,r。依賴於C特別重要，因為C_{th}限制在出現錯誤之概率變得過於大之前可識別連續鹼基之個數。

【0251】 圖39說明N-r參數空間，其中在位置150處之不正確鹼基識別之概率($P_{C=375,N,r}$)低於100分之一(Q20)、1,000分之一(Q30)、10,000分之一 (Q40)及100,000分之一(Q50)。增加簇大小N，或降低化學失敗率r，將臨限值C_{th}推至更高C值，但如圖39中定量顯示，簇大小相當大且允許的化學錯誤率必須很小以使DNA定序儀適合診斷應用。

【0252】 當前，定序行業之基準係讀取150個連續鹼基且在位置150處出現不正確鹼基識別之概率為1,000分之一之能力。此一般稱為Q30，但需要Q40且甚至Q50之顯著更大定序品質因子與更長讀取長度以偵測高精度診斷中之罕見突變。方程式3(a)及(b)中$P_{C,N,r}$之一般表示充分探索C-N-r參數空間且可用於估計任何定序度量之錯誤容限及簇大小要求。圖39顯示N-r參數空間之區域，其中在位置150 ($C \cong 375$)處不正確鹼基識別之概率低於100分之一(Q20)、1,000分之一(Q30)、10,000分之一(Q40)及100,000分之一(Q50)。例如，若定序陣列中之平均簇大小N為100個分子，且所需的定序精度為Q30，具有150 bp長讀段($C \cong 375$)，則允許的化學失敗率為$r \leq 0.002641$，亦即，在任何定序查詢步驟處允許於定序儀陣列上10,000次個別單分子反應中僅26次或更少次失敗。若所需的精度為Q50，則允許每10,000次反應中僅19次或更少次錯誤。若平均簇大小N減少至10個分子，則數量降至每10,000次反應中約6次 (Q30)及約1次(Q50)。

【0253】 圖40A顯示順著Q30等高線針對各種N-r組合之計算$P_{C,N,r}(C)$，在插圖中以交叉(「+」號)標記，所有交叉點均在$P_{C,N,r}(C =$

375) = 0.001處。該等圖顯示增加簇大小N不僅提高化學失敗容限，而且其藉由將臨限值C_{th}推至更高C值來延遲鹼基識別錯誤之發生，此導致累積錯誤降低。若在查詢循環C處進行不正確鹼基識別之概率為$P_{C,N,r}$，則進行正確識別之概率為$(1-P_{C,N,r})$。進行C連續正確識別之概率則為：

$$(1-P_{1,N,r})\times(1-P_{2,N,r})\times(1-P_{3,N,r})\times\ldots\times(1-P_{C,N,r}) \equiv \prod_{j=1}^{j=C}(1-P_{j,N,r}) \quad \text{(方程式5(a))}$$

【0254】 不進行以行形式之C正確鹼基識別之概率(其與在任何查詢循環C處出現至少一個錯誤之概率相同或更小) (或累積錯誤概率$\tilde{P}_{C,N,r}$)由以下給出：

$$\tilde{P}_{C,N,r} = 1-\prod_{j=1}^{j=C}(1-P_{j,N,r}) \quad \text{(方程式5(b))}$$

其中$P_{j,N,r}$由方程式4(a)或(b)給出。圖40B順著相同等高線繪製計算累積錯誤概率$\tilde{P}_{C,N,r}(C)$，且說明較大簇產生較低累積錯誤。

【0255】 最後，指示計算且繪製標記其中在位置150處不正確鹼基識別之累積概率(在一些實施例中，目標讀取長度)小於或等於100分之一($\tilde{Q}20$)、1,000分之一($\tilde{Q}30$)、10,000分之一($\tilde{Q}40$)及100,000分之一($\tilde{Q}50$)之區域之N-r參數空間。圖41說明在位置150處不正確鹼基識別之累積概率($\tilde{P}_{C=375,N,r}$)小於或等於100分之一($\tilde{Q}20$)、1,000分之一($\tilde{Q}30$)、10,000分之一($\tilde{Q}40$)及100,000分之一($\tilde{Q}50$)之 N-r參數空間。圖41中之圖定量地顯示CLUS定序儀可包括大DNA簇大小N以受益於集體行為，且其可能需要極可靠之化學(每10,000次反應中僅幾十次失敗)以用於高精度診斷應用。更具體言之，若定序陣列中之平均簇平均保持例如100個分子，且特定定序應用程式耐受1,000分之一($\tilde{Q}30$)之累積鹼基識別錯誤概率，則在任何定序查詢步驟處允許於定序儀陣列上10,000次個別單分子反應中僅約22次或更

少次失敗。圖41中之圖說明藉由減小簇大小N且將更多簇封裝至感測區域中來增加定序通量僅可以定序化學之並行改良來達成。所需改良率隨著簇大小N變小而加速，且CLUS裝置可不再受益於大集體行為。

單分子陣列定序儀之一般定量結果

【0256】 為比較CLUS及SMAS平臺，開發簡單定量模型以估計SMAS裝置100中不正確鹼基識別之概率。不同於適用於(上文描述的)CLUS裝置之集體情況(在該情況下幾乎不能至無法實施錯誤校正)，SMAS裝置100個別地定序且記錄對應於個別核酸分子之偵測結果之能力允許開發且實施識別且消除所得資料記錄中之至少一些錯誤之強力技術。如本文所揭示之一或多種錯誤校正技術可應用於在進行鹼基識別之前自定序程序(例如SBS)產生之資料以識別且校正偵測結果中之錯誤以改良識別的序列之準確度。具體言之，可改良在定序程序之一些或所有查詢步驟處來自多個感測器105之偵測結果之比對。即使錯誤校正算法成功地正確比對多個感測器偵測結果，仍可進行不正確鹼基識別。如上文所說明，巧合的FNR錯誤及FLD錯誤可導致可能無法校正之插入及刪除錯誤。取決於資料記錄中之錯誤數量(該數量部分由化學失敗率決定)，錯誤校正過程可為複雜且計算密集，但應瞭解現代化處理器具有足夠計算能力以進行甚至最計算密集的所揭示技術。

【0257】 下文，考慮SMAS裝置100之K個單分子感測器105之一般情況，各感測器能夠監測選殖DNA之單個例項。如在上文裝置CLUS裝置之分析中，假設四種類型之錯誤(FNI、FLR、FNR及FLD)在定序程序期間隨機發生且分佈於整個查詢步驟。

【0258】 如上文所說明，在一些實施例中，實施概率性錯誤校正算

法(例如藉由至少一個處理器130，其可包括在SMAS裝置100中或在SMAS裝置100外部)。在一些實施例中，概率性錯誤校正算法改良資料記錄中之至少一些感測器105偵測結果之比對。在一些實施例中，一些或所有錯誤校正算法係在一些或所有查詢步驟已完成且已捕獲一些或所有資料之後實施。如前面所述，錯誤校正程序基本上消除FNI及FLR、以及一些FLD。感測器105偵測結果之算法再比對亦使得進行不正確鹼基識別之概率與查詢步驟數C無關。再者，因為錯誤校正算法將資料記錄中之至少一些感測器105偵測結果再比對，藉此校正至少一些錯誤，有效錯誤率r小於在CLUS情況中。在應用示例性錯誤校正算法之後，在一些實施例中，僅當在算法上比對的序列中K個感測器105中超過一半給出不正確結果時，不正確地識別鹼基。

【0259】 進行不正確鹼基識別之概率($P_{K,r}$)僅係(a)定序相同核酸分子之例項之感測器105之數量K(其可為小於感測器陣列110中之所有感測器105)及(b)化學失敗率r之函數。類似於上文CLUS裝置之分析所採用的方法，將K值限制為奇數值以避免其中感測器105中恰好一半與另一半不一致之情況。進行不正確鹼基識別之概率由以下給出：

$$P_{K,r} = \binom{K}{(K+1)/2} r^{(K+1)/2}(1-r)^{(K-1)/2} + \binom{K}{(K+3)/2} r^{(K+3)/2}(1-r)^{(K-3)/2} + \binom{K}{(K+5)/2} r^{(K+5)/2}(1-r)^{(K-5)/2} + \binom{K}{(K+7)/2} r^{(K+7)/2}(1-r)^{(K-7)/2} + \ldots + \binom{K}{K} r^{K} \quad \text{(方程式6)}$$

其中$\binom{a}{b} = \frac{a!}{b!(a-b)!}$。例如，若$K = 3$，

$$P_{K=3,r} = \binom{3}{2} r^2(1-r) + \binom{3}{3} r^3 = 3r^2(1-r) + r^3 = 3r^2 - 2r^3 \quad \text{(方程式7)}$$

【0260】 在$K=3$之實例中，相乘$\binom{3}{2}=3$項解釋其中3個感測器105中有2個同時在特定查詢步驟處遭遇錯誤(例如其不正確地偵測到標記(FLR、FNR)或不正確地未偵測到標記(FNI、FLD))，藉此迫使不正確鹼基識別之情況。將三個感測器105表示為S1、S2及S3，此情景發生在以下情況：(1) S1及S2同時遭遇錯誤，(2) S1及S3同時遭遇錯誤，或(3) S2及S3同時遭遇錯誤。$\binom{3}{3}=1$項解釋不可能之情況，即所有三個感測器S1、S2及S3同時遭遇錯誤，此亦導致不正確鹼基識別。因為多項式擴展中之最大項為r^{K-1}且$0<r<1$，故進行不正確鹼基識別之概率藉由增加單分子感測器105之數量(亦即增加K值)顯著降低。

【0261】 例如，若$r=0.1$，則$P_{K=3,r=0.1}=0.029$，此意指進行不正確鹼基識別的機率為約100分之三。換言之，150次鹼基識別中平均約4.35次將係不正確的，此對於一些診斷應用而言過大。為使用三個奈米級感測器105以Q30 ($P_{K,r}=0.001$)定序，將需要化學失敗率降低至$r=0.01837$，意指將允許1,000次查詢中僅約19次係錯誤的。然而，若將感測器105之數量(K值)增加至11，則可容忍一百次反應中超過12次失敗。

【0262】 如上文針對CLUS裝置所進行，下文針對SMAS裝置100探索K-r參數空間以識別其中在任何查詢位置處不正確鹼基識別之概率低於100分之一(Q20)、1,000分之一(Q30)、10,000分之一(Q40)及100,000分之一(Q50)之區域。圖42說明其中在每一查詢步驟處不正確鹼基識別之概率($P_{K,r}$)低於100分之一(Q20)、1,000分之一(Q30)、10,000分之一(Q40)及100,000分之一(Q50)之 K-r參數空間之計算結果。如圖42中所顯示，若感測相同核酸分子之例項之單分子感測器105之數量K為11，且所需的定序精度為Q30，則允許的化學失敗率為$r \lesssim 0.13$，意指允許彼等11個感測器

105當中100次個別單分子反應中多達約13次失敗。若所需的精度為Q50，則允許11個感測器105當中每100次反應中約6次或更少次錯誤。

【0263】 如與圖39之比較指示，SMAS裝置100之允許的錯誤率顯著大於針對於CLUS裝置而言允許的比率，然而單獨該結果並不能公平地比較兩個平臺，因為在CLUS裝置中進行不正確鹼基識別之概率($P_{C,N,r}$)在早期查詢步驟期間極低且在臨限查詢步驟C_{th}處突然增加。結合圖39論述此種現象。另一方面，對於SMAS裝置100，不正確鹼基識別之概率($P_{K,r}$)在整個查詢步驟中保持恆定且因此導致較大累積錯誤。

比較CLUS裝置及SMAS裝置100之性能之一種更公平的方式係比較兩種裝置類型之累積錯誤概率。上文方程式5(b)表示CLUS裝置之累積錯誤概率。亦可導出SMAS裝置100之累積錯誤概率。在各查詢步驟C處進行不正確鹼基識別之概率為$P_{K,r}$(方程式6)，且因此進行正確識別之概率為$(1-P_{K,r})$。進行以行形式之C正確識別之概率則為$\left(1-P_{K,r}\right)^{C}$，及累積錯誤概率($\tilde{P}_{K,r}$)為

$$\tilde{P}_{K,r}=1-\left(1-P_{K,r}\right)^{C} \qquad \text{(方程式8)}$$

【0264】 圖43A及43B顯示CLUS裝置及SMAS裝置100之在位置150處不正確鹼基識別之累積概率。方程式5(b)可例如用於計算CLUS裝置在小於或等於150之任何鹼基位置處進行不正確鹼基識別之概率。圖43A顯示CLUS裝置之K-r參數空間且標記其中對於CLUS裝置在位置150處不正確鹼基識別之累積概率小於或等於100分之一($\tilde{Q}20$)、1,000分之一($\tilde{Q}30$)、10,000分之一($\tilde{Q}40$)及100,000分之一($\tilde{Q}50$)之區域。圖43B評估方程式(8)且顯示K-r參數空間，其標記其中對於SMAS裝置100在位置150處不正確鹼基識別之累積概率小於或等於100分之一($\tilde{Q}20$)、1,000分之一($\tilde{Q}30$)、

10,000分之一($\tilde{Q}40$)及100,000分之一($\tilde{Q}50$)之區域。

【0265】 圖43A及43B之比較顯示SMAS裝置100係潛在優於CLUS裝置之定序平臺。SMAS裝置100可具有較小佔據空間(footprint) (如例如圖7A、7B、9A、9B及10之論述中所說明)且可比CLUS裝置更具錯誤容限。與CLUS裝置相比，使用SMAS裝置100允許更高通量，更低錯誤率，及更長讀段長度，該等CLUS裝置更大且仰賴於大分子集體。商業上可行之SMAS裝置100及/或系統之開發可使用以下中之一些或全部：(a) 密集封裝之能夠識別個別標記之感測器105之高精度奈米級製造，(b)降低錯誤率至可接受之程度之化學步驟之最佳化，及/或(c)可利用有效生物信息學工具，藉由概率性消除錯誤來調整資料記錄中來自至少一些奈米級感測器105之定序資料之比對。

示例性SMAS定序程序

【0266】 如上文所說明，若亦降低定序化學失敗率，則可藉由減小簇大小N(藉此將更多簇封裝至裝置中)來達成CLUS裝置之定序通量之改良，此可具有挑戰性。相比之下，下文呈現根據一些實施例之使用單分子結合位點116之大陣列之錯誤容限、超高通量SMAS裝置100之可行實現。出於實例之目的，假設SMAS裝置100定序DNA，但應明瞭，一般而言，可定序任何種類之核酸。

【0267】 圖44及45說明根據一些實施例之示例性實例製備及加載過程500。圖44係說明過程500之流程圖，及圖45說明過程500之各個步驟之結果。在一些實施例中，樣品製備及加載過程500開始於502。在504處，進行DNA提取及純化，此導致幾個提取的DNA片段505，如圖45中所顯示。在506處，將與引物互補之轉接子拼接至提取的DNA的一端(例如3')

以產生顯示於圖45中之股507。在508處，進行PCR(或一些其他複製技術)以產生所提取的股之多個(理想地，相同)例項，如圖45中的509所顯示。在510處，將能夠在SMAS裝置100之流體腔室115 (結合位點116)之化學官能化表面建立強鍵(例如藉由點擊化學(click chemistry))之分子連接子附接至ssDNA片段的另一端(例如5')，藉此產生顯示於圖45中之股511。在512處，將官能化股511加載至流體腔室115中且在結合位點116當中隨機散佈且結合至結合位點116。如圖45之最右側部分中所顯示，結合位點116中之各者支持不多於單個DNA股。(儘管各結合位點116可支持不多於一個股，但應瞭解，並不需要每個結合位點116必須支持DNA股。無論是有意還是偶然地，均可使用SMAS裝置100之少於所有結合位點116。)假設提取的DNA片段503彼此不同，由於樣品製備及加載過程500，流體腔室115內將存在所提取的DNA片段505中之各者之多個例項，但其位置係未知的。在514處，示例性樣品製備及加載過程500結束。

【0268】 示例性樣品製備及加載過程500之益處在於，其簡化DNA擴增，此可在將DNA股添加至SMAS裝置100之前使用(例如)習知PCR在裝置外大量進行。相比之下，當使用CLUS裝置時，僅在已將DNA片段添加至CLUS裝置中之後執行擴增(例如橋式擴增)以便建立經擴增之DNA之連續簇陣列。

【0269】 在已進行樣品製備及加載過程500之後，可使用例如以上介紹的加性方法、減性方法或改進型加性方法來進行鹼基識別。圖46A、46B及46C說明在藉由利用具有以四個行及五個列配置之20個感測器105(及20個結合位點116)之感測器陣列110之實例SMAS裝置100進行之三個示例性查詢循環(對於總共12個查詢步驟，各為**A**?⇒**T**?⇒**C**?⇒**G**?)期間使

用改進型加性方法之模擬偵測結果(感測器105偵測到標記)。四個不同DNA股之多個例項隨機分佈在整個感測器陣列110中，但其於該感測器陣列110中之特定位置及其序列最初係未知的。

【0270】 圖47說明如何可重新配置說明於圖46A、46B及46C中之偵測資料以識別鹼基且顯示不同DNA股之位置。圖47提供顯示在個別查詢步驟處示例性陣列中每個感測器105之輸出及導致識別的序列之所得鹼基識別之表。圖47之右手側部分將感測器105再排序以將感測相同DNA股之例項之感測器105之偵測結果分組。如圖47中所顯示，四個序列識別為：**GCT** (股#1)、**TAG** (股#2)、**ACG** (股#3)及**TTA** (股#4)。

【0271】 若在查詢步驟期間發生錯誤(FNI、FLR、FNR或FLD)，則一些偵測結果(偵測到標記或未偵測到標記)將係不正確，且可實施以上描述的判定性及/或概率性錯誤偵測及/或校正技術以偵測且消除至少一些錯誤，只要判定感測相同DNA股之例項之彼等感測器105之同一性即可。回想一下，可將特定DNA股之例項附接至散佈於整個流體腔室115中之結合位點116，且在定序過程開始時，其位置一般係未知的。一旦啟動該過程，則在各查詢步驟期間，複數個S個感測器105中之各者在其各別結合位點116處偵測到標記。為進行錯誤校正，識別定序相同核酸股之例項之S個感測器105之亞組。

【0272】 考慮具有4億個不同DNA股之極大感測器陣列110 (例如40億個結合位點116及40億個各別感測器105)，各DNA股為約150個鹼基長度。此意指各獨特DNA股有約10個例項隨機分佈於整個流體腔室115 (及結合位點116及感測器陣列110)中。為了實例起見，亦假設序列係隨機的。假設合理地低的錯誤率r，在第一查詢循環之後，將識別保持(感測)

以**A**開始之DNA例項之幾乎所有結合位點116 (及感測器105)，將識別保持(感測) **T**之彼等、及保持(感測) **C**之彼等、及保持(感測) **G**之彼等。約10^9個感測器105將偵測到指示第一鹼基為**A**之標記，約10^9個感測器105將偵測到指示第一鹼基為**T**之標記，約10^9個感測器將偵測到指示第一鹼基為**C**之標記，及約10^9個感測器將偵測到指示第一鹼基為**G**之標記。在第二查詢循環之後，將識別保持(感測)以所有16種可能之組合(**AA**、**AT**、**AC**、**AG**、**TA**、**TT**、**TC**、**TG**、**CA**、**CT**、**CC**、**CG**、**GA**、**GT**、**GC**及**GG**)開始之DNA例項之幾乎所有結合位點116 (及感測器105)。約2.5×10^8個感測器將偵測到指示第一及第二鹼基為**AA**之標記，約2.5×10^8個感測器將偵測到指示第一及第二鹼基為**AT**之標記，約2.5×10^8個感測器將偵測到指示第一及第二鹼基為**AC**之標記。一般而言，在一些數量D個標記偵測(或假設將改進型加性方法用於定序之$C \cong 2.5 \times D$個查詢步驟)之後，將識別保持以一些D-鹼基長度之序列開始之DNA股之所有$4^D=4^{2C/5}$個結合位點116。此意指具有40億個感測器陣列110之SMAS裝置100中感測相同DNA股之例項之感測器105群組之平均大小為$4 \times 10^9/(4^{2C/5})$。因為吾等實例每個獨特股平均具有約10個例項，故將進行約$C \cong 35$個查詢循環以識別保持特定股之例項之結合位點116之位置。假設使用改進型加性方法，在該過程期間將識別約14個鹼基。因為人類基因組不是隨機的，且並非所有數學上可能之序列均被顯示，故對於診斷應用而言實際上可能需要顯著更少查詢步驟。若在DNA提取期間靶向特定組基因，則可以甚至更少步驟來判定保持相同DNA股之例項之結合位點116之同一性(位置)，此進一步減少鹼基之可能序列之數量且有利於結合位點116識別。

【0273】 已識別該正確組之結合位點116組之置信度隨著查詢步驟

之數量而增加，但因此出現偵測錯誤(例如不正確地偵測到標記或不正確地未偵測到標記)之概率亦增加。在最初的查詢循環期間可出現多個錯誤同時識別保持相同股之例項之結合位點116。CLUS裝置之獲得的結果表明此可能不是問題。例如，圖38A及38B顯示，在早期查詢步驟期間，CLUS裝置的進行不正確鹼基識別之概率極小，且僅在達到臨限值C_{th}時，錯誤概率才急劇增加。亦回想一下，若不應用錯誤校正，則SMAS裝置100之鹼基識別精度與CLUS裝置相同，因為SMAS裝置100將藉由加總個別感測器105結果來簡單地報告集體結果。

【0274】考慮例如上文的40億個感測器陣列實例且考慮監測特定DNA股之例項之一組11個感測器105($K = 11$)隨機分佈於整個結合位點116中。現在，將其視為集體($K = N = 11$)，就好像結合位點116正在形成簇且僅測量其各別感測器105之組合特性(例如信號)。圖48A及48B繪製進行不正確鹼基識別之計算概率$P_{C,N,r}$，其由方程式4(a)及(b)以查詢步驟數量C及化學失敗率r之函數關係給出。圖48A中之曲線標記C-r空間中$P_{C,N,r}$突然增加之臨限值之近似位置。圖48B係顯示於圖48A中之等高線圖之俯視圖且清楚地指示包含各DNA股平均約10個例項之40億個感測器SMAS裝置100之化學失敗容限。可可靠地判定保持(感測)各獨特DNA股之例項之約10個結合位點116 (及感測器105)之位置(同一性)，只要約35個查詢步驟中錯誤概率保持很低即可。此將最大允許化學失敗率限制為0.013，亦即，將容忍1,000個偵測事件中的13個。圖48A及4848B中之計算結果指示若化學失敗率保持低於每1,000個中約13個不正確地偵測事件，則40億個感測器SMAS裝置100應能夠於流體腔室115內(及結合位點116中及感測器105中)建立所有十億個不同DNA股之所有例項之位置。一

旦建立彼等位置，可立刻實施本文描述的錯誤校正技術以消除在剩餘約340個查詢步驟(假設使用改進型加性方法)期間發生的錯誤。

【0275】 若預期或已知化學錯誤率過高，使得錯誤可能困擾前約35個查詢步驟，則可使用替代方法以幫助識別攜帶相同DNA股之例項之結合位點116。例如，可將不同獨特條碼拼接至所提取的DNA子組中之引物末端使得在早期定序循環期間讀取已知序列。圖49說明根據一些實施例之條碼於樣品製備及DNA加載之用途。如圖49中所顯示，將獨特條碼拼接至所提取的DNA以有利於在存在定序錯誤下識別保持相同DNA之例項之位點。例如，圖49顯示四個獨特DNA股，對該等股中之各者分配獨特條碼(例如對股1分配條碼119A，對股2分配條碼119B，對股3分配條碼119C，及對股4分配條碼119D)。若條碼彼此顯著不同，則即使化學失敗率極高，其亦應易於識別。如所瞭解，對於高通量診斷應用，適宜數量之獨特條碼可很高。

【0276】 本文描述的示例性40億個感測器SMAS裝置100依當前標準被認為是相當高通量定序儀。此種SMAS裝置100在單個運行期間提供約150吉鹼基(Giga-base (Gb))讀段，此與2020年引入的當前最先進技術高端定序系統之輸出相媲美。

【0277】 應明瞭，存在實施本文揭示的裝置、系統及方法之許多方法。例如，用於核酸定序之系統可由單個裝置(例如SMAS裝置100，其包括可進行所揭示的操作之所有硬體及軟體)組成，或其可包括SMAS裝置100及一起進行所揭示的操作之其他組件。例如，系統可包括SMAS裝置100及SMAS裝置100外部(例如在外部電腦中)的至少一個處理器，SMAS裝置100進行核酸定序程序且保存來自該定序程序之偵測結果，，該至少

一個處理器對保存的偵測結果進行錯誤偵測及校正且識別鹼基。

【0278】 圖50說明根據一些實施例之示例性系統160。系統160包括(亦即包括但不限於)流體腔室115、複數個S個感測器105及至少一個處理器130。視需要，系統160包括用於儲存記錄之記憶體170，該記憶體包含在定序程序期間獲得的偵測結果(例如一或多個具有二進制條目之文檔，該二進制條目記錄在複數個查詢循環中之各者期間，複數個S個感測器105中之各者偵測到或未偵測到至少一個標記)。如以圖50中之虛線所顯示，若系統160包括記憶體170，則該至少一個處理器130可通訊耦合至記憶體170使得該至少一個處理器130可將資料儲存在記憶體170中及/或自記憶體170擷取資料。

【0279】 流體腔室115包括複數個S個結合位點，該等S個結合位點中之各者經結構設計成結合不超過一個欲定序的核酸股。圖50顯示四個結合位點116，但應明瞭，系統160可包括更多或更少個結合位點116。S個感測器105中之各者經結構設計成偵測存在於流體腔室115中的標記。圖50顯示四個感測器105，但應明瞭，系統160可包括更多或更少個感測器105。當系統160在操作中時，S個感測器105中之各者偵測到附接至併入與該等S個結合位點116之各別結合位點116結合之核酸之各別股中之核苷酸之標記。如前面所說明，感測器105可為磁感測器、光學感測器、或可偵測用於標記核苷酸之標記之任何其他類型之感測器。流體腔室115、感測器105及結合位點116詳細描述於上文中。彼等描述適用於圖50且在此不予重複。

【0280】 該至少一個處理器130經結構設計成執行一或多個可機器執行之指令。該等指令在被執行時導致該至少一個處理器130進行包括複

數個查詢步驟之定序程序(例如，如圖11、12、14、16、44中之任何者之內文中所描述)。具體而言，在操作中，在定序程序之查詢步驟期間，該至少一個處理器130獲得該等S個感測器105中之各者之各別特性(由至少一個處理器130與感測器與感測器105A、105B、105C及105D之間的虛線表示)。各別特性指示感測器105偵測到或未偵測到標記(例如其指示至少一個標記之存在或不存在)。該至少一個處理器130可解釋所獲得的特性以判定感測器105偵測到或未偵測到標記之存在。至少部分地基於所獲得的各別特性，該至少一個處理器130記錄各別感測器在查詢步驟期間是否偵測到至少一個標記之存在或不存在。該至少一個處理器130亦經結構設計成對包含定序程序之結果之至少一個記錄進行錯誤校正程序。錯誤校正程序可對藉由定序程序產生的一些或全部記錄操作，且其可對來自定序程序之一些或全部查詢步驟之偵測結果操作。例如，如上文所述，為應用錯誤校正程序，該至少一個處理器可識別K個記錄之子組且對其應用判定性或概率性錯誤校正，其中該子組中該K個記錄中之各者對應於來自感測相同核酸股之例項之感測器105之偵測結果。定序程序及錯誤校正程序詳細描述於上文中。彼等描述適用於圖50之系統及至少一個處理器130，且在此不予重複。

【0281】 至少一個處理器130可藉由通用或專用處理器(或處理核組)來實施且因此可執行一系列程式化指令以實現與獲得感測器105特性，進行錯誤校正程序，及/或與使用者、系統操作者或其他系統組件之互動相關之各種操作。

【0282】 系統160之至少一個處理器130可為單個處理器(例如在SMAS裝置100中)，或其可包括多個處理器，其可為共位(co-located) (例

如在SMAS裝置100中)或物理上彼此分開。例如，該至少一個處理器130之第一部分可包括在SMAS裝置100中，及該至少一個處理器130之第二部分可在SMAS裝置100的外部。在其中該至少一個處理器130包括第一及第二部分之實施例中，該第一部分可負責獲得感測器105之特性，基於該等特性來判定感測器105在查詢循環期間是否偵測到標記，且記錄(例如在記憶體170中)S個感測器105中之各者在查詢循環期間是否偵測到至少一個標記之存在或不存在，且該第二部分負責獲得偵測結果之記錄及進行錯誤校正程序。或者，該第一部分可負責獲得感測器105之特性，基於該等特性來判定感測器105中之各者在查詢循環期間是否偵測到至少一個標記，及經通訊介面(例如無線或有線介面，諸如乙太網路(Ethernet)、Wi-Fi等)將感測器105是否偵測到標記之指示提供至另一實體。在此類實施案中，該至少一個處理器130之第二部分可負責獲得藉由至少一個處理器130之第一部分提供的偵測結果之記錄(例如具有二進制條目之文檔，該等二進制條目記錄在各查詢循環期間複數個S個感測器105中之各者偵測到或未偵測到至少一個標記)，進行錯誤校正程序，及識別鹼基。

【0283】 在前述描述中及在附圖中，已闡述特定術語以提供對所揭示的實施例之透徹理解。在一些例項下，術語或附圖可意指實施本發明不需要的特定細節。

【0284】 為避免不必要地混淆本發明，熟知的組件以方塊圖形式顯示及/或在一些情況下根本不進行詳細論述。

【0285】 提供於實施方式中之部分標題僅係為了方便或參考而無意為限制性的。部分標題絕不定義，限制，解釋，或描述此等部分之範疇或程度。再者，儘管已揭示各種特定實施例，但顯而易見的是，在不脫離本

發明之更寬泛精神及範疇下，可對本發明進行各種修改及改變。例如，該等實施例中之任何者之特徵或態樣可與該等實施例中之任何其他者組合應用或替代其對應特徵或態樣應用。

【0286】 本文揭示的某些技術及方法(例如自感測器105獲得偵測結果，進行錯誤校正程序等)及/或用於結構設計及管理其之使用者介面可藉由機器執行一或多個序列指令(包括正確指令執行所需的相關資料)來實施。可將此種指令記錄在一或多個電腦可讀媒體上以供在專用或通用電腦系統或消費電子裝置或電器之一或多個處理器內之稍後擷取及執行。其中可體現此種指令及資料之電腦可讀媒體包括但不限於各種形式之非易失性儲存媒體(例如光學、磁或半導體儲存媒體)及可用於通過無線、光學或有線信號傳導媒體傳輸此種指令及資料之載波或其任何組合。藉由載波傳輸此種指令及資料之實例包括但不限於經網際網路及/或其他電腦網路經由一或多個資料傳輸協定(例如HTTP、FTP、SMTP等)之傳輸(上載、下載、電子郵件(e-mail)等)。

【0287】 除非本文另外明確定義，否則所有術語意欲給出其最寬廣之可能之解釋，包括本說明書及附圖涵蓋之含義及如熟習此項技術者所理解及/或如字典、專著等中所定義之含義。如本文明確陳述，一些術語可能與其尋常或慣常含義不符。

【0288】 如本說明書及隨附申請專利範圍中所使用，單數形式「一」、「一個」及「該」不排除複數個指示物，除非另有說明。除非另有說明，否則語詞「或」應解釋為包含性的。因此，片語「A或B」應解釋為以下所有含義：「A及B」、「A而不是B」及「B而不是A」。本文「及/或」之任何使用並不意指語詞「或」單獨表示排他性。

【0289】 如本說明書及隨附申請專利範圍中所使用，形式「A、B及C中之至少一者」、「A、B或C中之至少一者」、「A、B或C中之一者或多者」及「A、B及C中之一者或多者」之片語係可互換的，且各涵蓋以下所有含義：「僅A」、「僅B」、「僅C」、「A及B而不是C」、「A及C而不是B」、「B及C而不是A」及「A、B及C全部」。

【0290】 在實施方式或申請專利範圍中使用術語「包括(include(s))」、「具有(having)」、「具有(has)」、「具有(with)」及其變化形式之程度上，此類術語意欲以類似於術語「包含(comprising)」之方式為包含性的，亦即意指「包括但不限於」。

【0291】 術語「示例性」及「實施例」用於表示實例，而不是優選項或要求。

【0292】 術語「耦合」在本文中用於表示直接連接/附接以及透過一或多個中間元件或結構之連接/附接。

【0293】 術語「上方」、「下方」、「之間」及「之上」在本文中用於指一個特徵相對於其他特徵之相對位置。例如，配置於另一特徵「上方」或「下方」之一個特徵可與另一特徵直接接觸或可具有中間材料。此外，配置在兩個特徵「之間」的一個特徵可與兩個特徵直接接觸或可具有一或多個中間特徵或材料。相比之下，於第二特徵「之上」的第一特徵係與該第二特徵接觸。

【0294】 術語「實質上」用於描述大程度上或接近所述之結構、組態、尺寸等，但由於製造容限及類似物，可實務上導致其中結構、組態、尺寸等並非始終或不一定完全如所述之情境。例如，將兩個長度描述為「實質上相等」意指將兩個長度描述為基本相等意味著該兩個長度對於所

有實際目的均係相同的，但在足夠小的標度上其可能不(且不需要)精確相等。作為另一個實例，出於所有實際目的，「實質上垂直」之結構將被視為是垂直的，即使其相對於水平不是精確地成90度。

【0295】 附圖不一定按比例繪製，且特徵之尺寸、形狀及大小可實質上不同於其在附圖中之描繪方式。

【0296】 儘管已揭示特定實施例，但將顯而易見的是，在不脫離本發明之較寬廣精神及範疇下，可對其作出各種修改及改變。例如，該等實施例中之任何者之特徵或態樣可在至少可行之情況下與該等實施例中之任何其他實施例組合或替代其對應特徵或態樣加以應用。因此，本說明書及附圖應被認為係例示性而非限制性意義。

【符號說明】

【0297】

15: 感測器

100: SMAS裝置

101: DNA股

101A: DNA股

101B: DNA股

102: 標記

102A: 標記

102B: 標記

103: 點劃線圓圈

105: 磁感測器

105A: 感測器

105B: 感測器

105C: 感測器

105D: 感測器

105E: 感測器

105F: 感測器

105G: 感測器

106A: 鐵磁層

106B: 鐵磁層

107: 非磁性間隔層

108: 底表面

109: 頂表面

110: 感測器陣列

111: 感測區域邊界

112: 最近相鄰距離

114: 單位格子

115: 流體腔室

116A: 結合位點

116B: 結合位點

116C: 結合位點

116D: 結合位點

116E: 結合位點

116F: 結合位點

116G: 結合位點

119A: 條碼

119B: 條碼

119C: 條碼

119D: 條碼

120: 電路

125A: 線

125B: 線

125C: 線

125D: 線

125E: 線

125F: 線

125G: 線

125H: 線

130: 處理器

160: 系統

170: 記憶體

200: 方法

202: 步驟

204: 步驟

206: 步驟

208: 步驟

210: 步驟

212: 步驟

212: 錯誤校正程序

214: 步驟

220: 定序程序

222: 步驟

224: 步驟

226: 步驟

228: 步驟

230: 步驟

232: 步驟

234: 步驟

236: 步驟

238: 步驟

240: 步驟

242: 步驟

244: 步驟

250: 定序程序

252: 步驟

254: 步驟

256: 步驟

258: 步驟

260: 步驟

262: 步驟

264: 步驟

266: 步驟

268: 步驟

270: 步驟

272: 步驟

274: 步驟

350: 定序程序

352: 步驟

354: 步驟

356: 步驟

358: 步驟

360: 步驟

362: 步驟

364: 步驟

366: 步驟

368: 步驟

370: 步驟

372: 步驟

374: 步驟

400: 定序程序

402: 步驟

404: 步驟

406: 步驟

408: 步驟

410: 步驟

412: 步驟

414: 步驟

416: 步驟

418: 步驟

420: 步驟

422: 步驟

424: 步驟

426: 步驟

428: 步驟

430: 步驟

432: 步驟

450: 錯誤校正程序

452: 步驟

454: 步驟

456: 步驟

458: 步驟

460: 步驟

500: 樣品製備及加載過程

502: 步驟

504: 步驟

506: 步驟

508: 步驟

510: 步驟

512: 步驟

514: 步驟

【發明申請專利範圍】

【請求項1】

一種用於定序核酸之系統，該系統包含：

複數個S個結合位點，該等S個結合位點中之各者係經結構設計成結合不超過一個欲定序的核酸股；

複數個S個感測器，其經結構設計成偵測標記，該等S個感測器中之各者用於感測與該等S個結合位點之各別結合位點結合之核酸之各別股；及

至少一個處理器，其經結構設計成執行一或多個可機器執行之指令，該等指令在執行時導致該至少一個處理器：

(a)在定序程序之複數個M個查詢步驟之各查詢步驟處，且對於該等S個感測器中之各者：

獲得該各別感測器之各別特性，其中該各別特性指示至少一個標記之存在或不存在，且

至少部分地基於該所獲得的各別特性，記錄該各別感測器在該查詢步驟期間是否偵測到至少一個標記之存在或不存在，及

(b) 對至少一個記錄進行錯誤校正程序，該至少一個記錄包含在該等M個查詢步驟中之各者處針對該等S個感測器之至少一個子組之定序程序之結果，其中對該至少一個記錄進行該錯誤校正程序包括：

基於該至少一個記錄之至少一部分，識別與特定核酸股(strand)之例項(instances)相關之複數個候選序列，及

判定或估計該複數個候選序列中的哪個最可能是正確的。

【請求項2】

如請求項1之系統，其進一步包含儲存複數個記錄之記憶體，且其中

該至少一個記錄係該複數個記錄中之一者。

【請求項3】

如請求項1之系統，其中，當執行時，該一或多個可機器執行之指令進一步導致該至少一個處理器在該定序程序之該複數個M個查詢步驟之各查詢步驟處且針對該等S個感測器中之各者：

解釋該各別感測器之所獲得的各別特性以判定該各別感測器是否偵測到該至少一個標記之存在或偵測到該至少一個標記之不存在。

【請求項4】

如請求項1之系統，其中該複數個S個感測器中之各者係經結構設計成偵測螢光團。

【請求項5】

如請求項1之系統，其中該複數個S個感測器中之各者係經結構設計成偵測磁粒子。

【請求項6】

如請求項1之系統，其中該複數個S個感測器中之各者係經結構設計成偵測帶電荷分子。

【請求項7】

如請求項1之系統，其中該複數個S個感測器中之各者係經結構設計成偵測有機金屬錯合物。

【請求項8】

如請求項1之系統，其中該至少一個標記包含螢光團。

【請求項9】

如請求項8之系統，其中該等S個感測器中之各者包含光學感測器，

且其中該各別特性係由該螢光團發出的光之波長、頻率、強度或顏色。

【請求項10】

如請求項1之系統，其中該至少一個標記包含磁粒子。

【請求項11】

如請求項10之系統，其中該等S個感測器中之各者包含磁感測器。

【請求項12】

如請求項11之系統，其中該磁感測器包含磁振盪器，且其中該各別特性係與該磁振盪器相關聯或由該磁振盪器產生之信號之頻率。

【請求項13】

如請求項11之系統，其中該磁感測器包含磁穿隧接面。

【請求項14】

如請求項11之系統，其中該特性係磁場或電阻、磁場或電阻之變化、雜訊位準或雜訊位準之變化。

【請求項15】

如請求項1之系統，其中該至少一個標記包含帶電荷分子。

【請求項16】

如請求項1之系統，其中該至少一個標記包含有機金屬錯合物。

【請求項17】

如請求項1之系統，其中該S個感測器係以矩形陣列配置。

【請求項18】

如請求項17之系統，其中該至少一個標記係磁性的，且其中該矩形陣列中之相鄰感測器之間的最近相鄰距離為至少約70 nm。

【請求項19】

如請求項1之系統，其中該等S個感測器係以六邊形陣列配置。

【請求項20】

如請求項19之系統，其中該至少一個標記係螢光，且其中該六邊形陣列中之相鄰感測器之間的最近相鄰距離為至少約140 nm。

【請求項21】

如請求項1之系統，其中該複數個S個結合位點係位於流體腔室內，且其中該流體腔室之表面包含金屬氧化物、二氧化矽、聚丙烯、金、玻璃或矽。

【請求項22】

如請求項1之系統，其中該等S個結合位點中之各者包含經結構設計成錨定核酸之各別股之結構。

【請求項23】

如請求項22之系統，其中該結構包括空腔或脊。

【請求項24】

如請求項1之系統，其中對該至少一個記錄進行該錯誤校正程序包括以下中之至少一者：

進行判定性錯誤校正程序，或

進行概率性錯誤校正程序。

【請求項25】

如請求項1之系統，其中判定或估計該複數個候選序列中哪個具有是正確之最高概率包括：

判定該複數個候選序列中之各者之各別度量；及

至少部分地基於該等各別度量及標準，將特定候選序列選擇為最可

能是正確的。

【請求項26】

如請求項25之系統，其中該等各別度量係發生可能性，且其中該標準係最小發生可能性。

【請求項27】

如請求項25之系統，其中該等各別度量係發生可能性，且其中該標準係臨限發生可能性。

【請求項28】

如請求項1之系統，其中判定或估計該複數個候選序列中哪個具有是正確之最高概率包括基於對該特定核酸股之核酸序列之已知約束來消除該複數個候選序列中之至少一者。

【請求項29】

如請求項28之系統，其中該已知約束係鹼基之特定序列之不可能性。

【請求項30】

如請求項28之系統，其中判定或估計該複數個候選序列中哪個具有是正確之最高概率另外包括至少部分地基於該特定核酸股之來源來判定已知約束。

【請求項31】

如請求項1之系統，其中該至少一個記錄包含二進制值集合，其中第一二進制值指示偵測到該標記，及第二二進制值指示未偵測到標記，且其中進行該錯誤-校正程序包括：

在該至少一個記錄中識別一串第二二進制值，及

自該至少一個記錄刪除該串第二二進制值。

【請求項32】

如請求項31之系統，其中該串第二二進制值具有四之長度。

【請求項33】

如請求項1之系統，其中對該至少一個記錄進行錯誤校正程序包括：

在該至少一個記錄中識別該等S個感測器之第一感測器未偵測到標記之一組連續指示，及

自該至少一個記錄刪除該等S個感測器之該第一感測器未偵測到標記之該組連續指示。

【請求項34】

如請求項1之系統，其中對該至少一個記錄進行該錯誤校正程序包括：

基於特定查詢步驟之大多數結果來更改該至少一個記錄之至少一個條目。

【請求項35】

一種用於定序核酸之裝置，該裝置包含：

流體腔室，其包含複數個S個結合位點，該等S個結合位點中之各者經結構設計成結合不超過一個待定序的核酸股；

複數個S個磁感測器，其經結構設計成偵測存在於該流體腔室中之標記，該等S個磁感測器中之各者用於感測與該等S個結合位點之各別結合位點結合之核酸之各別股；及

至少一個處理器，其經結構設計成執行一或多個可機器執行之指令，該等指令在執行時導致該至少一個處理器在定序程序之複數個M個查

詢步驟之各查詢步驟處且針對該等S個磁感測器中之各者：

獲得該各別磁感測器之各別特性，其中該各別特性指示至少一個標記之存在或不存在，

至少部分地基於該所獲得的各別特性，判定該各別磁感測器在查詢步驟期間是否偵測到至少一個標記之存在或不存在，且

在與該各別磁感測器相關之各別記錄中記錄該各別磁感測器在該查詢步驟期間是否偵測到至少一個標記之存在或不存在。

【請求項36】

如請求項35之裝置，其進一步包含：

儲存複數個記錄之記憶體，

且其中與該各別磁感測器相關之該各別記錄係該複數個記錄中之一者。

【請求項37】

如請求項35之裝置，其中判定該各別磁感測器在該查詢步驟期間是否偵測到該至少一個標記之存在或不存在包括：

判定該各別磁感測器之所獲得的各別特性是否符合或超過臨限值。

【請求項38】

如請求項35之裝置，其中判定該各別磁感測器在該查詢步驟期間是否偵測到該至少一個標記之存在或不存在包括：

比較該各別磁感測器之所獲得的各別特性與先前偵測到的值。

【請求項39】

如請求項38之裝置，其中該先前偵測到的值係基線值。

【請求項40】

如請求項38之裝置，其中該先前偵測到的值係頻率、磁場或雜訊位準。

【請求項41】

如請求項35之裝置，其中該複數個S個磁感測器中之各者係經結構設計成偵測磁粒子。

【請求項42】

如請求項35之裝置，其中該複數個S個磁感測器中之各者係經結構設計成偵測帶電荷分子。

【請求項43】

如請求項35之裝置，其中該複數個S個磁感測器中之各者係經結構設計成偵測有機金屬錯合物。

【請求項44】

如請求項35之裝置，其中該至少一個標記包含磁粒子。

【請求項45】

如請求項35之裝置，其中該至少一個標記包含帶電荷分子。

【請求項46】

如請求項35之裝置，其中該至少一個標記包含有機金屬錯合物。

【請求項47】

如請求項35之裝置，其中該複數個S個磁感測器中之各者包含磁振盪器，且其中該各別特性係與該磁振盪器相關聯或由該磁振盪器產生之信號之頻率。

【請求項48】

如請求項35之裝置，其中該複數個S個磁感測器中之各者包含磁穿隧

接面。

【請求項49】

如請求項35之裝置，其中該特性係磁場或電阻、磁場或電阻之變化、雜訊位準或雜訊位準之變化。

【請求項50】

如請求項35之裝置，其中該等S個磁感測器係以矩形陣列配置。

【請求項51】

如請求項50之裝置，其中該至少一個標記係磁性的，且其中該矩形陣列中之相鄰磁感測器之間的最近相鄰距離為至少約70 nm。

【請求項52】

如請求項35之裝置，其中該流體腔室之表面包含金屬氧化物、二氧化矽、聚丙烯、金、玻璃或矽。

【請求項53】

如請求項35之裝置，其中該等S個結合位點中之各者包含經結構設計成錨定核酸之各別股之結構。

【請求項54】

如請求項53之裝置，其中該結構包括空腔或脊。

【請求項55】

如請求項35之裝置，其中，當由該至少一個處理器執行時，該一或多個可機器執行之指令進一步導致該至少一個處理器：

對至少一個記錄進行錯誤校正程序，該至少一個記錄包含在該等M個查詢步驟中之各者處針對該等S個磁感測器之至少一個子組之定序程序之結果。

【請求項56】

如請求項55之裝置，其中，當執行時，該一或多個可機器執行之指令導致該至少一個處理器藉由以下中之至少一者對該至少一個記錄進行該錯誤校正程序：

進行判定性錯誤校正程序，或

進行概率性錯誤校正程序。

【請求項57】

如請求項55之裝置，其中對該至少一個記錄進行該錯誤校正程序包括：

基於該至少一個記錄之至少一部分，識別與特定核酸股之例項相關之複數個候選序列，及

判定或估計該複數個候選序列中的哪個最可能是正確的。

【請求項58】

如請求項57之裝置，其中判定或估計該複數個候選序列中哪個最可能是正確的包括：

判定該複數個候選序列中之各者之各別度量；及

至少部分地基於該等各別度量及標準，將特定候選序列選擇為最可能是正確的。

【請求項59】

如請求項58之裝置，其中該等各別度量係發生可能性，且其中該標準係最小發生可能性。

【請求項60】

如請求項58之裝置，其中該等各別度量係發生可能性，且其中該標

準係臨限發生可能性。

【請求項61】

如請求項57之裝置，其中判定或估計該複數個候選序列中哪個最可能是正確的包括基於對該特定核酸股之核酸序列之已知約束來消除該複數個候選序列中之至少一者。

【請求項62】

如請求項61之裝置，其中該已知約束係鹼基之特定序列之不可能性。

【請求項63】

如請求項61之裝置，其中判定或估計該複數個候選序列中哪個具有是正確之最高概率另外包括至少部分地基於該特定核酸股之來源來判定已知約束。

【請求項64】

如請求項55之裝置，其中該至少一個記錄包含二進制值集合，其中第一二進制值指示偵測到該標記，及第二二進制值指示未偵測到標記，且其中進行該錯誤-校正程序包括：

在該至少一個記錄中識別一串第二二進制值，及

自該至少一個記錄刪除該串第二二進制值。

【請求項65】

如請求項64之裝置，其中該串第二二進制值具有四之長度。

【請求項66】

如請求項55之裝置，其中對該至少一個記錄進行該錯誤校正程序包括：

在該至少一個記錄中識別未偵測到標記之一組連續指示，及

自該至少一個記錄刪除未偵測到標記之該組連續指示。

【請求項67】

如請求項55之裝置，其中對該至少一個記錄進行該錯誤校正程序包括：

基於特定查詢步驟之大多數結果來更改該至少一個記錄之至少一個條目。

【請求項68】

一種使用定序裝置來定序複數個S個核酸股之方法，該定序裝置包含流體腔室及複數個S個感測器，該複數個S個感測器經結構設計成偵測存在於該流體腔室中之標記，該等S個感測器中之各者用於感測與該流體腔室內的複數個S個結合位點中之各別結合位點結合之各別核酸股，該等S個結合位點中之各者經結構設計成結合不超過一個用於定序之核酸股，該方法包括：

將該等S個核酸股結合至該等S個結合位點；

進行包括M個查詢步驟之定序程序以產生S個記錄，該等S個記錄中之各者捕獲該等S個感測器之各別感測器之M個偵測結果，該等M個偵測結果中之各者指示在該等M個查詢步驟之各別查詢步驟期間該S個感測器之各別感測器是否在該流體腔室中偵測到至少一個標記，其中該等S個記錄中之各者中之該等M個偵測結果中之各者係由二進制值表示，且其中進行該定序程序包括：

回應於偵測到該至少一個標記之該等S個感測器中之該各別感測器，在該等S個記錄之各別記錄中記錄第一二進制值，及

回應於未偵測到該至少一個標記之該等S個感測器中之該各別感測器，在該等S個記錄之該各別記錄中記錄第二二進制值；及

將錯誤校正程序應用至該等S個記錄之至少一個子組以估計該S個核酸股中之至少一者之核酸序列。

【請求項69】

如請求項68之方法，其中該等S個記錄之該子組捕獲針對特定核酸股之例項之定序程序之結果。

【請求項70】

如請求項69之方法，該方法進一步包括在使該等S個核酸股與該等S個結合位點結合之前擴增或複製該特定核酸股以建立該特定核酸股之例項。

【請求項71】

如請求項70之方法，其中擴增或複製包括進行至少一個促進聚合酶鏈反應(PCR)之程序。

【請求項72】

如請求項68之方法，其中該等S個記錄之至少一個子組之各記錄對應於特定核酸股之各別例項。

【請求項73】

如請求項72之方法，該方法進一步包括在應用該錯誤校正程序之前識別該等S個記錄之子組。

【請求項74】

如請求項73之方法，其中識別該等S個記錄之子組係基於與該特定核酸股相關之特定條碼之知識。

【請求項75】

如請求項73之方法，其中識別該等S個記錄之該子組包括在該等S個記錄之該子組之各記錄中識別與該特定核酸股相關之特定條碼。

【請求項76】

如請求項73之方法，其中識別該等S個記錄之該子組包括在該等S個記錄之該子組之各記錄中識別條目之共用序列。

【請求項77】

如請求項68之方法，其中該定序程序包括：

(a)將經標記之核苷酸引入至該流體腔室中；

(b)沖洗掉未結合之分子；

(c)自該複數個S個感測器之第一感測器獲得第一特性；

(d)自該複數個S個感測器之第二感測器獲得第二特性；

(e)基於該第一特性來判定該第一感測器是否在該流體腔室中偵測到至少一個標記；

(f)基於該第二特性來判定該第二感測器是否在該流體腔室中偵測到至少一個標記；

(g)在該等S個記錄之第一記錄中記錄第一指示，該第一指示指示該第一感測器是否在該流體腔室中偵測到至少一個標記；

(h)在該等S個記錄之第二記錄中記錄第二指示，該第二指示指示該第二感測器是否在該流體腔室中偵測到至少一個標記；

針對至少一個其他經標記之核苷酸重複(a)至(h)；及

在針對該至少一個其他經標記之核苷酸重複(a)至(h)之後，切割且沖洗掉標記。

【請求項78】

如請求項68之方法，其中該定序程序包括：

(a)將複數個經標記之核苷酸引入至該流體腔室中，該複數個經標記之核苷酸中之各者使用各別連接子；

(b)沖洗掉未結合之核苷酸；

(c)切割第一連接子；

(d)自第一感測器獲得第一特性；

(e)自第二感測器獲得第二特性；

(f)基於該第一特性來判定該第一感測器是否在該流體腔室中偵測到至少一個標記；

(g)基於該第二特性來判定該第二感測器是否在該流體腔室中偵測到至少一個標記；

(h)在該等S個記錄之第一記錄中記錄第一指示，該第一指示指示該第一感測器是否在該流體腔室中偵測到至少一個標記；

(i)在該等S個記錄之第二記錄中記錄第二指示，該第二指示指示該第二感測器是否在該流體腔室中偵測到至少一個標記；

切割第二連接子；及

在切割該第二連接子後，重複(d)至(i)。

【請求項79】

如請求項68之方法，其中該定序程序包括：

(a)將經標記之核苷酸引入至該流體腔室中；

(b)沖洗掉未結合之分子；

(c)自第一感測器獲得第一特性；

(d)自第二感測器獲得第二特性；

(e)基於該第一特性來判定該第一感測器是否在該流體腔室中偵測到至少一個標記；

(f)基於該第二特性來判定該第二感測器是否在該流體腔室中偵測到至少一個標記；

(g)在該等S個記錄之第一記錄中記錄第一指示，該第一指示指示該第一感測器是否在該流體腔室中偵測到至少一個標記；

(h)在該等S個記錄之第二記錄中記錄第二指示，該第二指示指示該第二感測器是否在該流體腔室中偵測到至少一個標記；

(i)切割且沖洗掉標記；及

在切割且沖洗掉標記後，針對至少一個其他經標記之核苷酸重複(a)至(i)。

【請求項80】

如請求項68之方法，其中該等S個記錄之至少一個子組中之記錄數係奇數。

【請求項81】

如請求項68之方法，其中應用該錯誤校正程序包括：

在該等S個記錄之該至少一個子組之至少一個記錄中識別一串第二二進制值，及

自該至少一個記錄刪除該串第二二進制值。

【請求項82】

如請求項81之方法，其中該串第二二進制值具有四之長度。

【請求項83】

如請求項68之方法，其中該定序程序包括(a)第一查詢步驟，(b)在該第一查詢步驟後移除存在於該流體腔室中之標記之標記移除步驟，(c)在該標記移除步驟後偵測存在於該流體腔室中之殘餘標記之感測步驟，及(d)在該感測步驟後之第二查詢步驟，且其中進行該錯誤校正程序包括：

回應於經由該感測步驟判定該等S個感測器中之特定感測器偵測到該流體腔室中之殘餘標記，將該第二二進制值記錄在該等S個記錄之特定記錄之特定位置中，該特定記錄捕獲該特定感測器之偵測結果，其中該特定位置捕獲該第二查詢步驟之結果。

【請求項84】

如請求項68之方法，其中應用該錯誤校正程序包括：

在該等S個記錄之該至少一個子組之至少一個記錄中識別未偵測到標記之一組連續指示，及

自該至少一個記錄刪除未偵測到標記之該組連續指示。

【請求項85】

如請求項68之方法，其中應用該錯誤校正程序包括修改該等S個記錄之該至少一個子組中之一者或多者。

【請求項86】

如請求項68之方法，其中該等S個記錄之該至少一個子組包含代表第一核酸股之例項之定序結果之奇數個至少三個記錄。

【請求項87】

如請求項86之方法，其中應用該錯誤校正程序包括：

在該等S個記錄之該至少一個子組中之各者中識別特定查詢步驟之大

多數偵測結果；及

至少部分地基於該大多數偵測結果來識別或不識別該第一核酸股之鹼基。

【請求項88】

如請求項86之方法，其中該等S個記錄之該至少一個子組由第一、第二及第三記錄組成，且其中對於該等M個偵測結果之選定偵測結果，應用該錯誤校正程序包括：

回應於該第一、第二及第三記錄中之至少兩者中之選定偵測結果係相同，至少部分地基於該相同的選定偵測結果來記錄該第一核酸股之鹼基。

【請求項89】

如請求項68之方法，其中對於該等M個偵測結果之選定偵測結果，應用該錯誤校正程序包括：

回應於在該等S個記錄之該至少一個子組中超過一半中之選定偵測結果係相同，至少部分地基於該相同的選定偵測結果來識別或不識別該等S個核酸股中之至少一者之鹼基。

【請求項90】

如請求項68之方法，其中對於該等M個偵測結果之選定偵測結果，應用該錯誤校正程序包括：

回應於該等S個記錄之至少一個子組中超過一半中之選定偵測結果指示在該流體腔室中偵測到該至少一個標記，識別該等S個核酸股中之至少一者之鹼基。

【請求項91】

如請求項68之方法，其中該至少一個標記係磁性的或螢光的。

【請求項92】

如請求項68之方法，其中該至少一個標記包含帶電荷分子或有機金屬錯合物。

【請求項93】

一種減輕由於使用單分子感測器陣列之核酸定序程序而產生之定序資料之錯誤之方法，該單分子感測器陣列具有複數個感測器，該複數個感測器中之各者與複數個結合位點中之各別結合位點相關聯，該複數個結合位點中之各者經結構設計成結合不超過一個欲定序的核酸股，該方法包括：

在該定序資料中識別複數個記錄，該複數個記錄中之各者捕獲核酸之第一股之各別例項之各別定序結果，該複數個記錄中之各者具有複數個條目，對於該核酸定序程序之複數個查詢步驟中之各別查詢步驟，該複數個條目中之各者指示(a)藉由與核酸之該第一股之該各別例項相關之各別感測器偵測到標記，或(b)藉由與核酸之該第一股之該各別例項相關之該各別感測器未偵測到標記；

基於該複數個記錄，判定核酸之該第一股之複數個候選序列，該複數個候選序列中之各者估計核酸之該第一股之核酸序列之至少一部分；及

識別該複數個候選序列之特定候選序列為核酸之該第一股之該核酸序列之至少一部分，該特定候選序列為該複數個候選序列中最可能是正確的。

【請求項94】

如請求項93之方法，其中識別該複數個記錄包括：

搜索與核酸之該第一股相關之條碼之定序資料。

【請求項95】

如請求項93之方法，其中識別該複數個記錄包括：

識別該複數個記錄中之各者中之條目之共同序列。

【請求項96】

如請求項93之方法，其中核酸之該第一股之該核酸序列之該至少一部分為單個鹼基。

【請求項97】

如請求項93之方法，其中判定核酸之該第一股之該複數個候選序列包括：

在該複數個記錄中識別特定查詢步驟，在該特定查詢步驟處，第一感測器偵測到各別標記及第二感測器未偵測到任何標記；

確立第一候選序列，該第一候選序列假設該第一感測器正確地偵測到該各別標記；及

確立第二候選序列，該第二候選序列假設該第一感測器不正確地偵測到該各別標記。

【請求項98】

如請求項93之方法，其中判定核酸之該第一股之該複數個候選序列包括：

在該複數個記錄中識別特定查詢步驟，在該特定查詢步驟處，第一感測器偵測到各別標記及第二感測器未偵測到任何標記；

確立第一候選序列，該第一候選序列假設該第二感測器不正確地未偵測到任何標記；及

確立第二候選序列，該第二候選序列假設該第二感測器正確地未偵測到任何標記。

【請求項99】

如請求項93之方法，其中該複數個條目中之各者係第一二進制值或第二二進制值，其中該第一二進制值指示藉由該各別感測器偵測到該標記，及該第二二進制值指示藉由該各別感測器未偵測到標記，且其中判定核酸之該第一股之該複數個候選序列包括：

在該複數個記錄中之至少一者中識別一串第二二進制值，及

自該複數個記錄中之該至少一者刪除該串第二二進制值。

【請求項100】

如請求項99之方法，其中該串第二二進制值具有四之長度。

【請求項101】

如請求項93之方法，其中判定核酸之該第一股之該複數個候選序列包括：

在該複數個記錄中之至少一者中識別指示未偵測到標記之一組連續條目，及

自該複數個記錄中之該至少一者刪除指示未偵測到標記之該組連續條目。

【請求項102】

如請求項93之方法，其中識別該複數個候選序列中最可能是正確之特定候選序列包括判定或估計該複數個候選序列中哪個具有是正確之最高概率。

【請求項103】

如請求項93之方法，其中核酸之該第一股之該核酸序列之該至少一部分係單個鹼基，且其中識別該複數個候選序列中最可能是正確之該特定候選序列包括識別由該複數個記錄表示之特定查詢步驟之大多數結果。

【請求項104】

如請求項93之方法，其中識別該複數個候選序列中最可能是正確之該特定候選序列包括：

判定該複數個候選序列中之各者之各別發生可能性；及

基於該特定候選序列滿足約束之各別發生可能性，選擇該特定候選序列。

【請求項105】

如請求項104之方法，其中該約束係最小概率。

【請求項106】

如請求項104之方法，其中該約束係該特定候選序列之各別發生可能性高於該複數個候選序列中之所有其他候選序列之各別發生可能性。

【請求項107】

如請求項93之方法，其中識別該複數個候選序列中最可能是正確之該特定候選序列包括基於核酸之該第一股之核酸序列之已知約束來消除該複數個候選序列中之至少一者。

【請求項108】

如請求項107之方法，其中該已知約束係鹼基之特定序列之不可能性。

【請求項109】

如請求項107之方法，該方法進一步包括至少部分地基於核酸之該第

一股之來源來判定該已知約束。

【請求項110】

一種使用定序裝置來定序複數個S個核酸股之方法，該定序裝置包含流體腔室及複數個S個感測器，該複數個S個感測器經結構設計成偵測存在於該流體腔室中之標記，該等S個感測器中之各者用於感測與該流體腔室內的複數個S個結合位點中之各別結合位點結合之各別核酸股，該等S個結合位點中之各者經結構設計成結合不超過一個用於定序之核酸股，該方法包括：

將該等S個核酸股結合至該等S個結合位點；

進行包括M個查詢步驟之定序程序以產生S個記錄，該等S個記錄中之各者捕獲該等S個感測器之各別感測器之M個偵測結果，該等M個偵測結果中之各者指示在該等M個查詢步驟之各別查詢步驟期間該S個感測器之該各別感測器是否在該流體腔室中偵測到至少一個標記；及

將錯誤校正程序應用至該等S個記錄之至少一個子組以估計該S個核酸股中之至少一者之核酸序列，其中應用該錯誤校正程序包括修改該等S個記錄之該至少一個子組中之一者或多者。

【請求項111】

如請求項110之方法，其中該等S個記錄之該子組捕獲針對特定核酸股之例項之定序程序之結果。

【請求項112】

如請求項111之方法，該方法進一步包括在使該等S個核酸股與該等S個結合位點結合之前擴增或複製該特定核酸股以建立該特定核酸股之該等例項。

【請求項113】

如請求項112之方法，其中擴增或複製包括進行至少一個促進聚合酶鏈反應(PCR)之程序。

【請求項114】

如請求項110之方法，其中該等S個記錄之該至少一個子組之各記錄對應於特定核酸股之各別例項。

【請求項115】

如請求項114之方法，該方法進一步包括在應用該錯誤校正程序之前識別該等S個記錄之該子組。

【請求項116】

如請求項115之方法，其中識別該等S個記錄之該子組係基於與該特定核酸股相關之特定條碼之知識。

【請求項117】

如請求項115之方法，其中識別該等S個記錄之該子組包括在該等S個記錄之該子組之各記錄中識別與該特定核酸股相關之特定條碼。

【請求項118】

如請求項115之方法，其中識別該等S個記錄之該子組包括在該等S個記錄之該子組之各記錄中識別條目之共用序列。

【請求項119】

如請求項110之方法，其中該等S個記錄之該至少一個子組中之記錄數係奇數。

【請求項120】

如請求項110之方法，其中該定序程序包括(a)第一查詢步驟，(b)在

該第一查詢步驟後移除存在於該流體腔室中之該等標記之標記移除步驟，(c)在該標記移除步驟後偵測存在於該流體腔室中之殘餘標記之感測步驟，及(d)在該感測步驟後之第二查詢步驟，且其中進行該錯誤校正程序包括：

回應於經由該感測步驟判定該等S個感測器中之特定感測器偵測到該流體腔室中之殘餘標記，將二進制值記錄在該等S個記錄之特定記錄之特定位置中，該特定記錄捕獲該特定感測器之該等偵測結果，其中該特定位置捕獲該第二查詢步驟之結果。

【請求項121】

如請求項110之方法，其中修改該等S個記錄之該至少一個子組中之該一者或多者包括：

在該等S個記錄之該至少一個子組之至少一個記錄中識別未偵測到標記之一組連續指示，及

自該至少一個記錄刪除未偵測到標記之該組連續指示。

【請求項122】

如請求項110之方法，其中該等S個記錄之該至少一個子組包含代表第一核酸股之例項之定序結果之奇數個至少三個記錄。

【請求項123】

如請求項122之方法，其中應用該錯誤校正程序進一步包括：

在該等S個記錄之該至少一個子組中之各者中識別特定查詢步驟之大多數偵測結果；及

至少部分地基於該大多數偵測結果來識別或不識別該第一核酸股之鹼基。

【請求項124】

如請求項122之方法，其中該等S個記錄之該至少一個子組由第一、第二及第三記錄組成，且其中對於該等M個偵測結果之選定偵測結果，修改該等S個記錄之該至少一個子組中之該一者或多者包括：

回應於該第一、第二及第三記錄中之至少兩者中之該選定偵測結果係相同，至少部分地基於該相同的選定偵測結果來記錄該第一核酸股之鹼基。

【請求項125】

如請求項110之方法，其中對於該等M個偵測結果之選定偵測結果，應用該錯誤校正程序進一步包括：

回應於在該等S個記錄之該至少一個子組中超過一半中之該選定偵測結果係相同，至少部分地基於該相同的選定偵測結果來識別或不識別該等S個核酸股中之該至少一者之鹼基。

【請求項126】

如請求項110之方法，其中對於該等M個偵測結果之選定偵測結果，應用該錯誤校正程序進一步包括：

回應於該等S個記錄之該至少一個子組中超過一半中之該選定偵測結果指示在該流體腔室中偵測到該至少一個標記，識別該等S個核酸股中之該至少一者之鹼基。

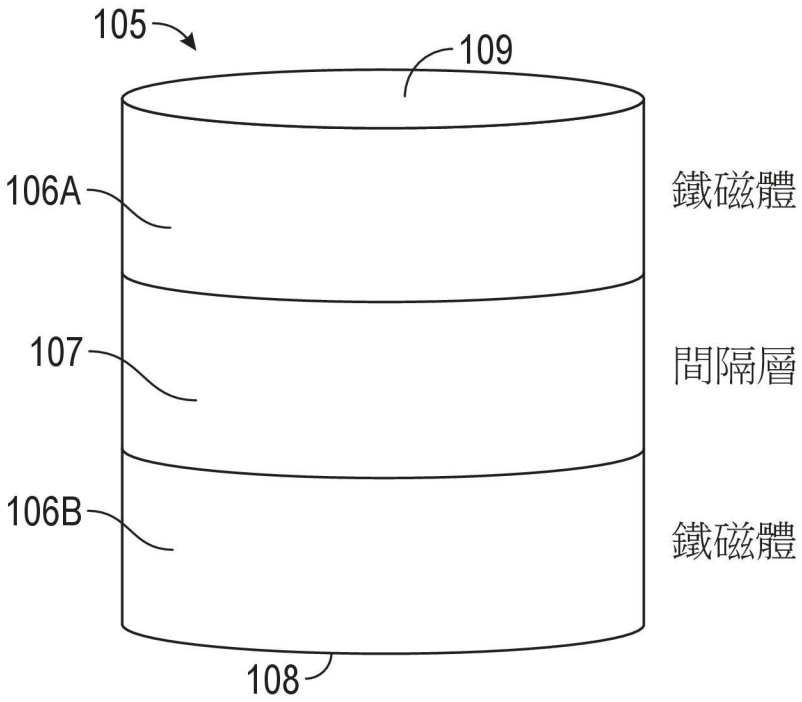
【請求項127】

如請求項110之方法，其中該至少一個標記係磁性的或螢光的。

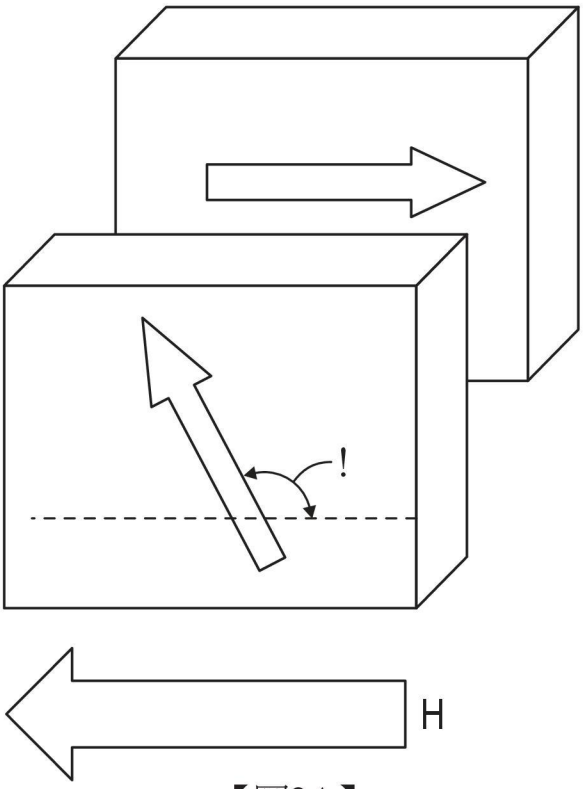
【請求項128】

如請求項110之方法，其中該至少一個標記包含帶電荷分子或有機金屬錯合物。

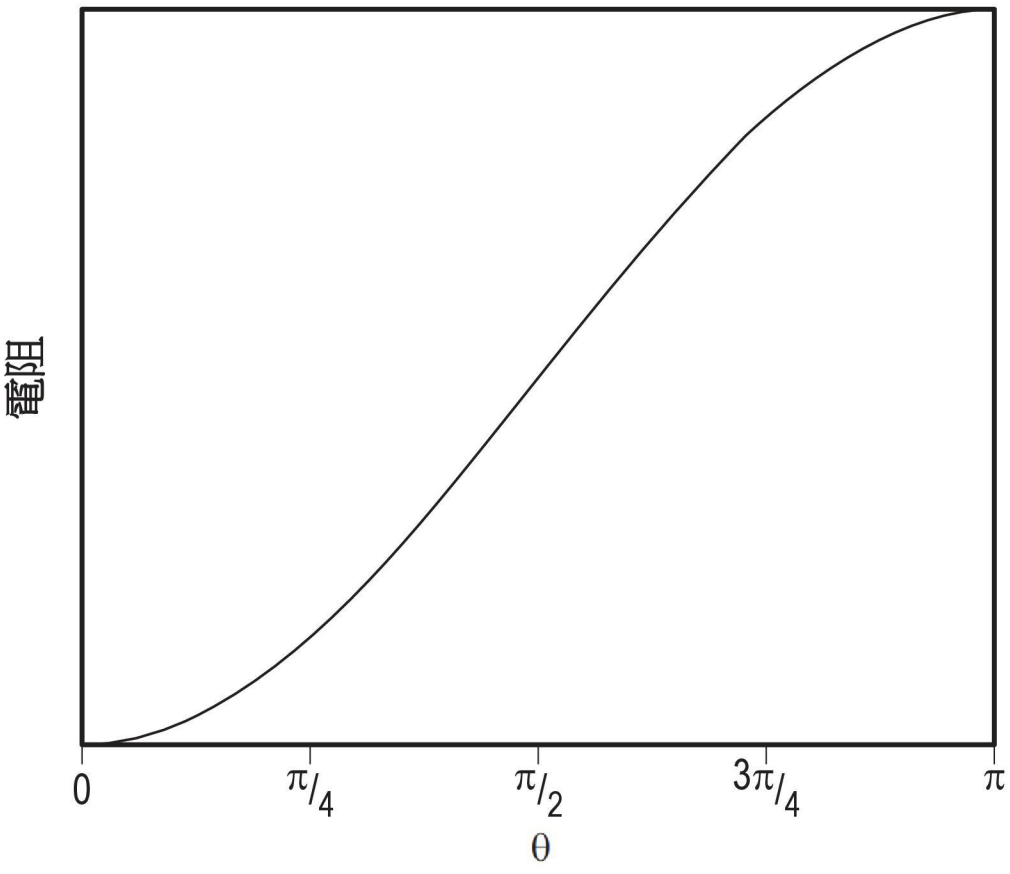
【發明圖式】



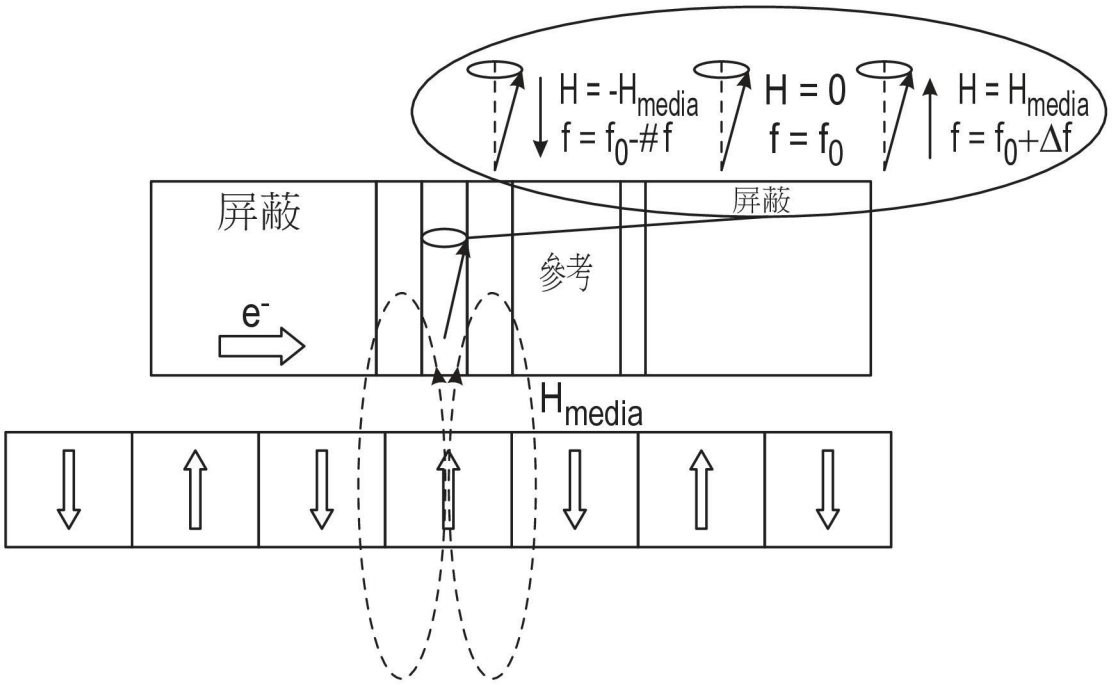
【圖1】



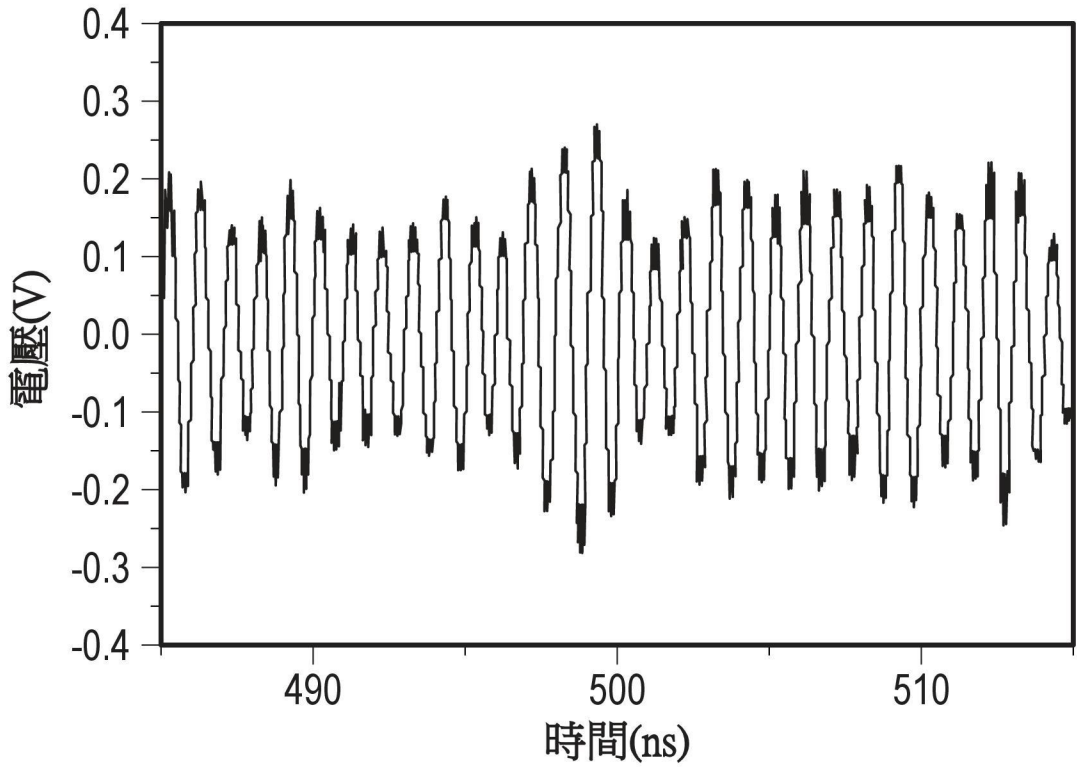
【圖2A】



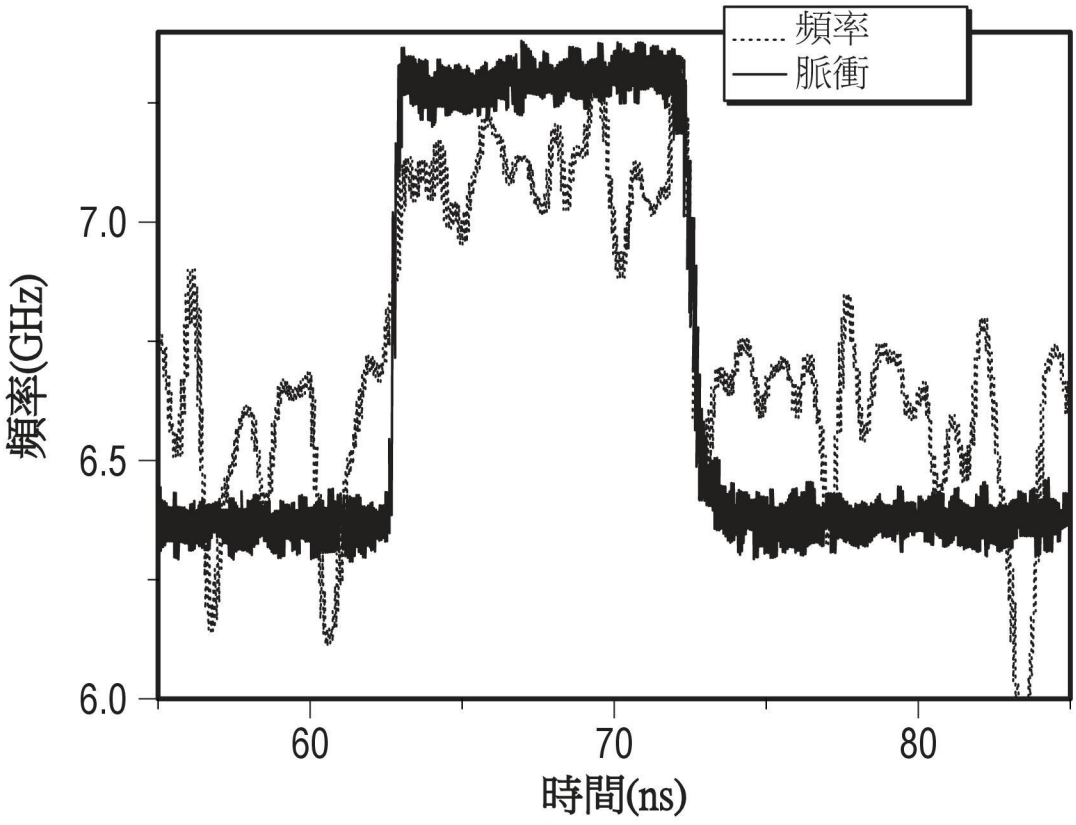
【圖2B】


H = -Hmedia
f = f0-#f
H = 0
f = f0
H = Hmedia
f = f0+Δf
屏蔽
屏蔽
參考
e-
Hmedia

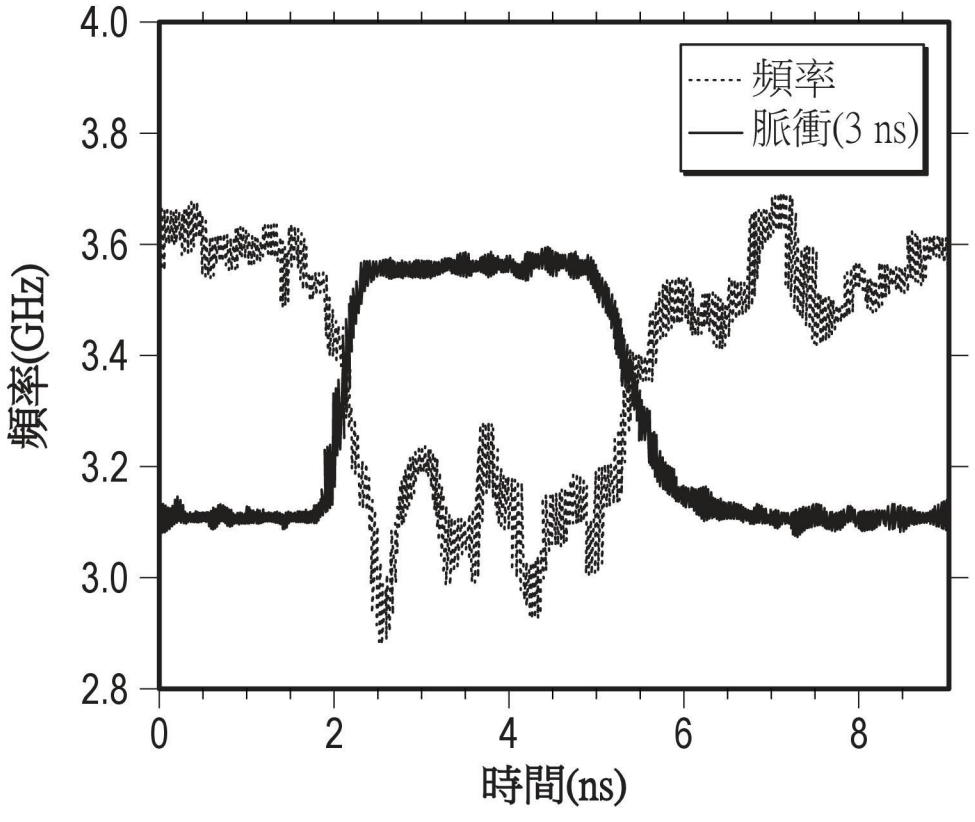
【圖3A】


0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
電壓(V)
490
500
510
時間(ns)

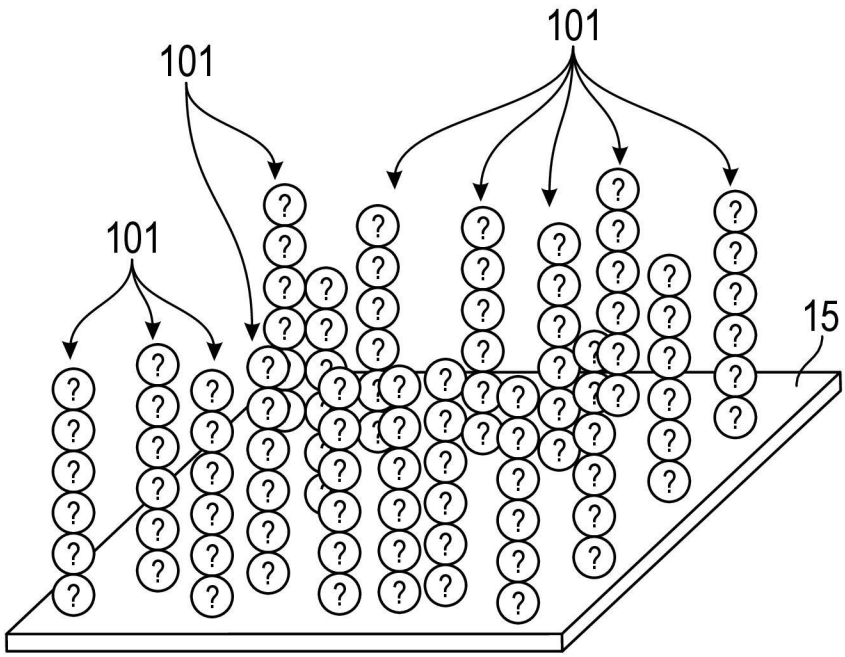
【圖3B】


頻率
脈衝
7.0
6.5
6.0
頻率(GHz)
60
70
80
時間(ns)

【圖3C】

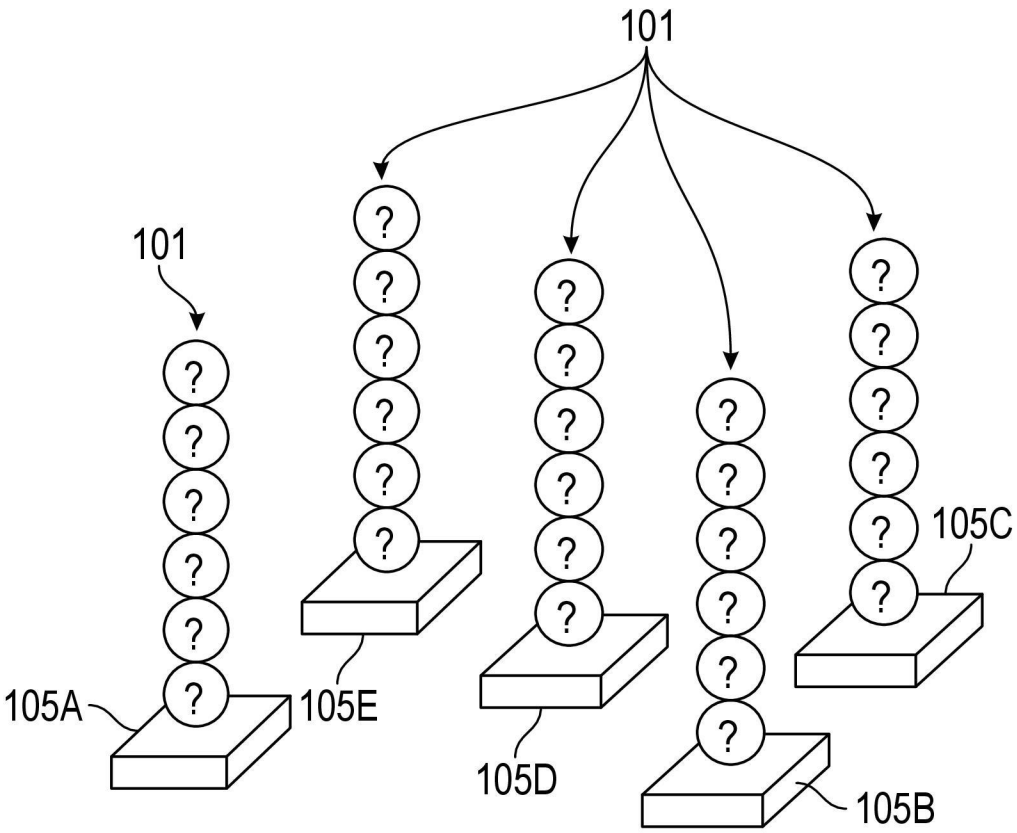

頻率
脈衝(3 ns)
4.0
3.8
3.6
3.4
3.2
3.0
2.8
頻率(GHz)
0
2
4
6
8
時間(ns)

【圖3D】



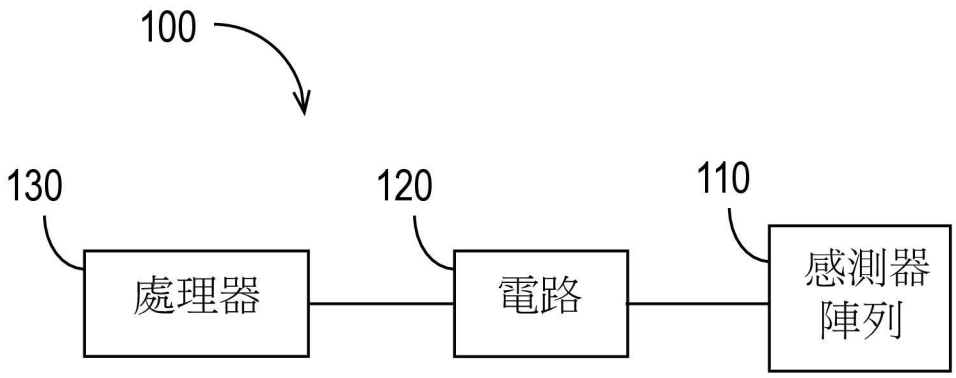
N股，1個感測器

【圖4A】

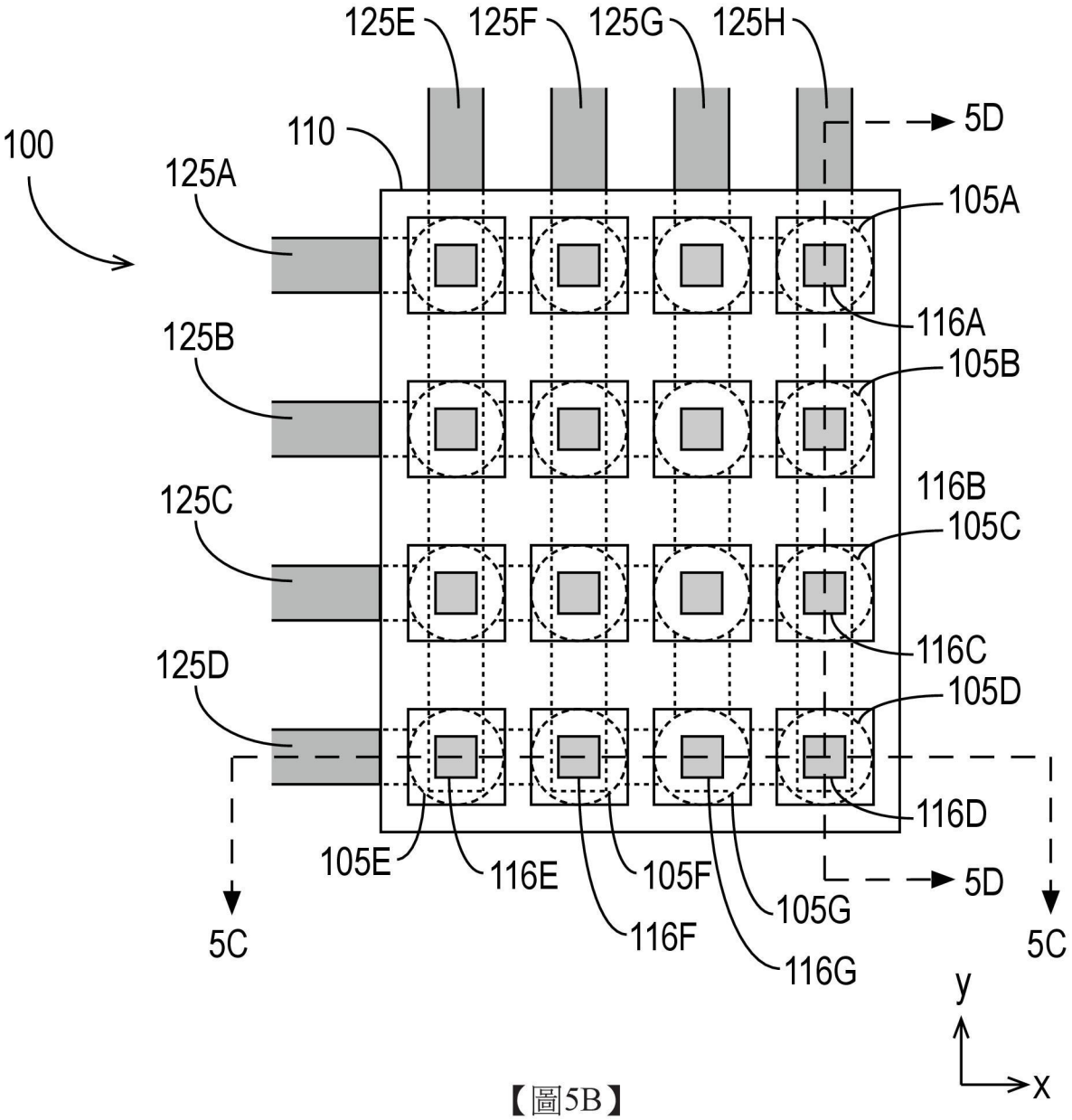


S股，S個感測器

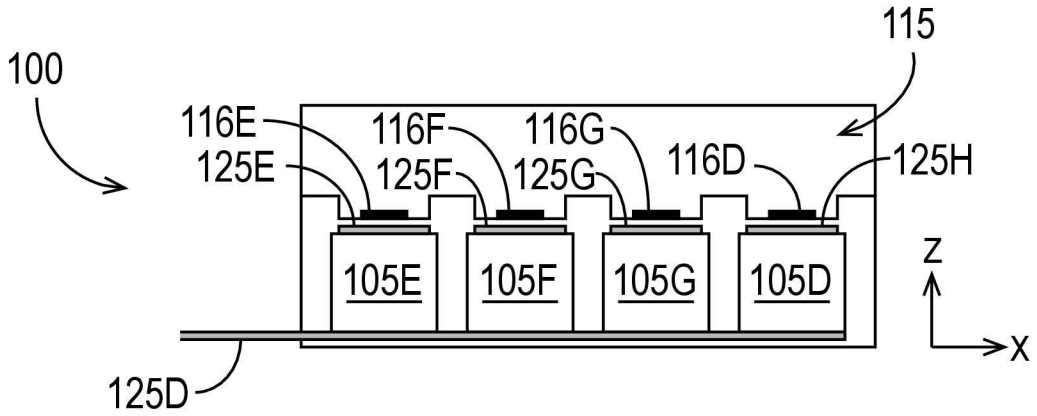
【圖4B】


100
130
處理器
120
電路
110
感測器
陣列

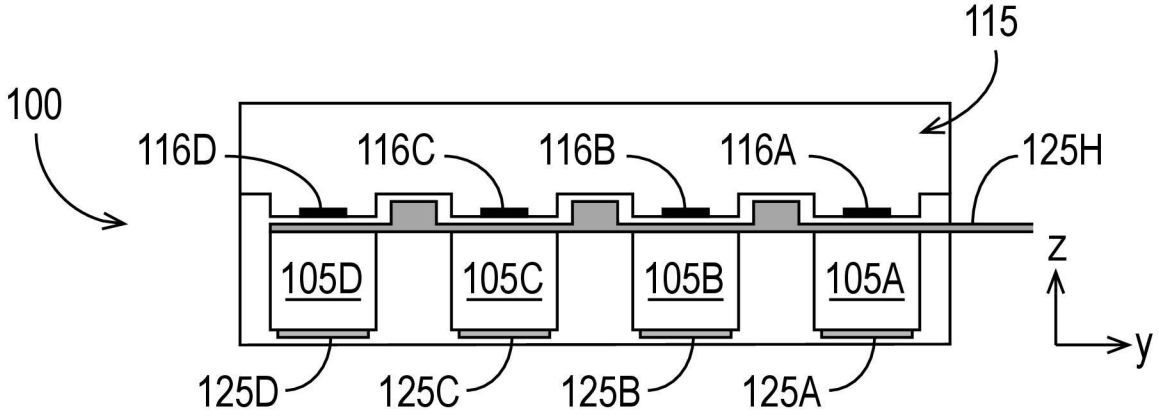
【圖5A】


100
110
125A
125B
125C
125D
125E
125F
125G
125H
5D
105A
116A
105B
116B
105C
116C
105D
116D
105E
116E
105F
116F
105G
116G
5C
5C
5D
y
x

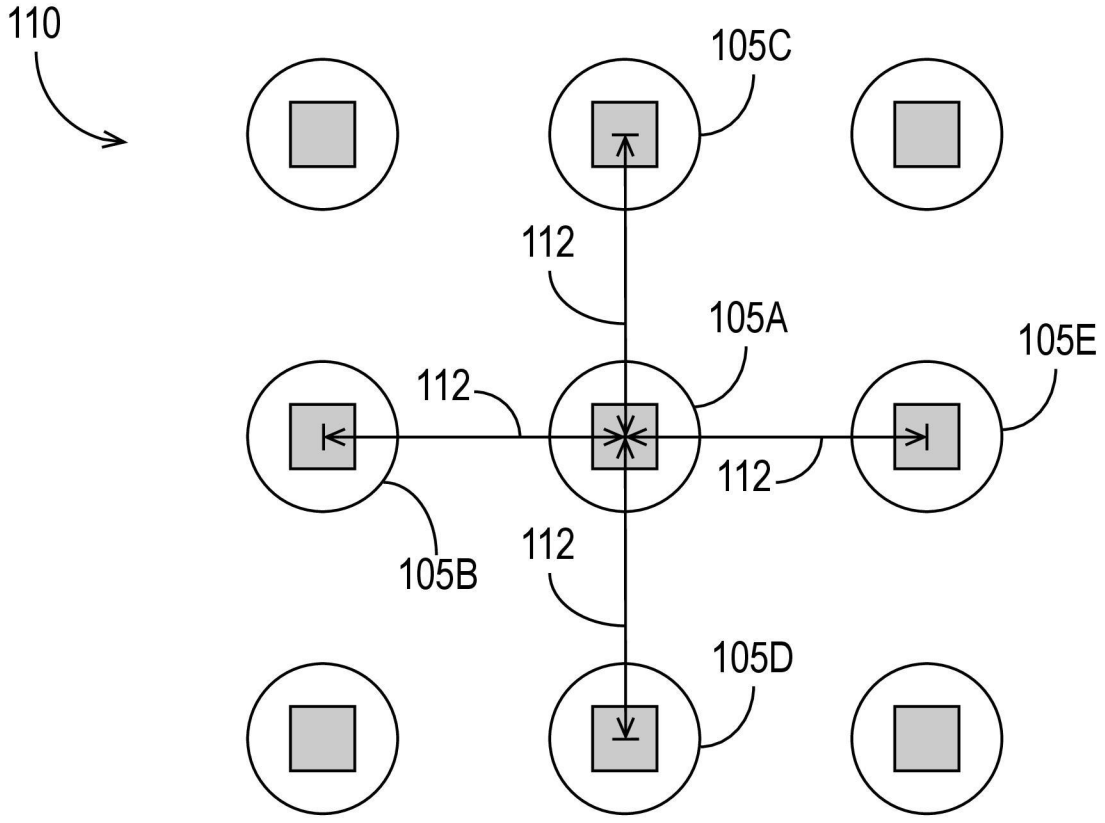
【圖5B】


100
115
116E
116F
116G
116D
125H
125E
125F
125G
105E
105F
105G
105D
z
x
125D

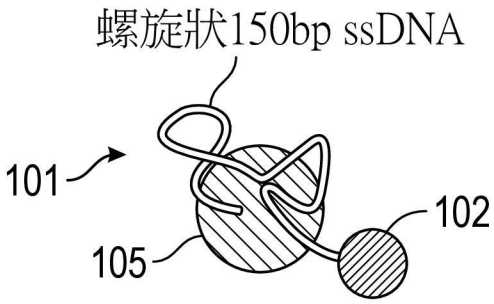
【圖5C】


100
115
116D
116C
116B
116A
125H
105D
105C
105B
105A
z
y
125D
125C
125B
125A

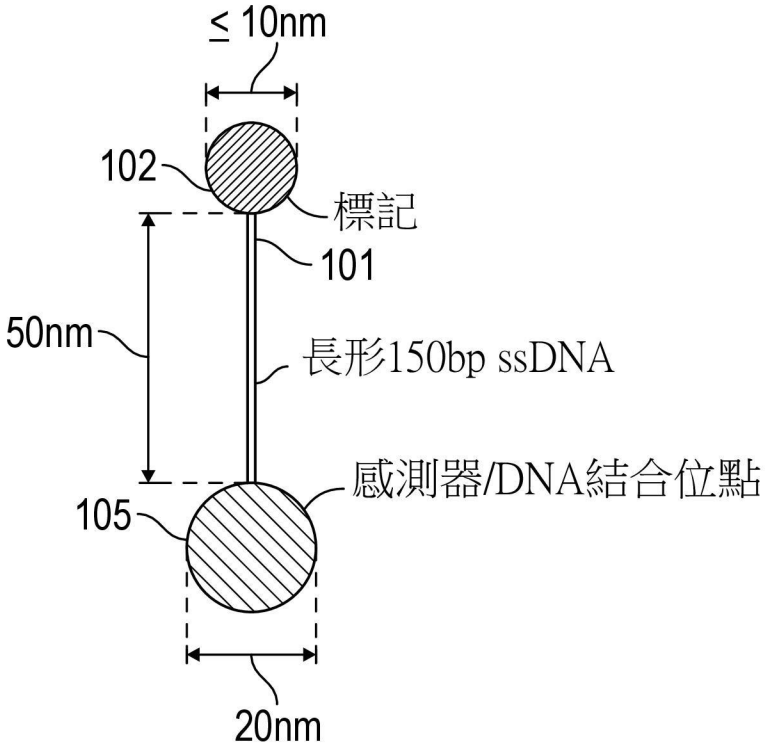
【圖5D】


110
105C
112
105A
105E
112
112
112
105B
112
105D

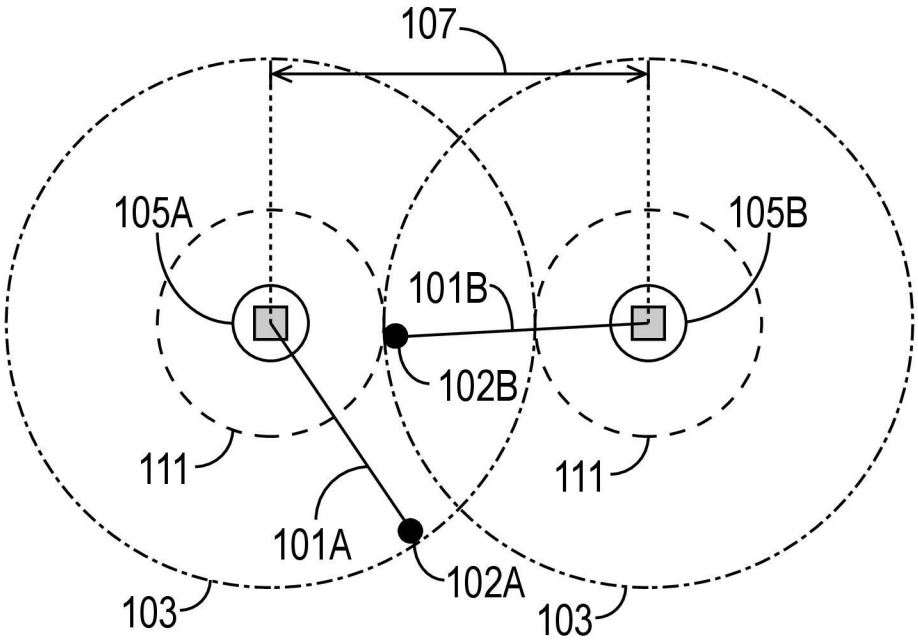
【圖5E】


螺旋狀150bp ssDNA
101
102
105

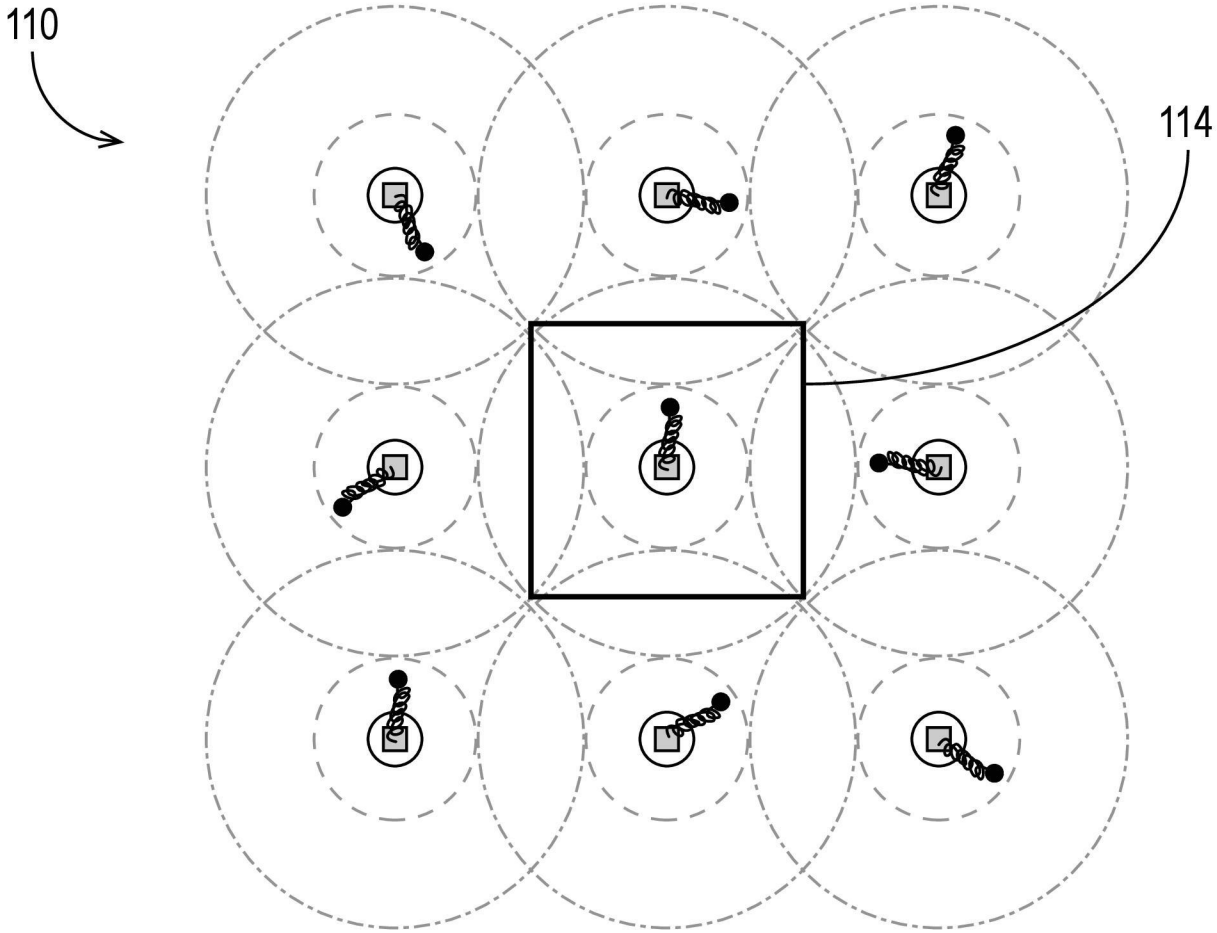
【圖6A】


≤ 10nm
102
標記
101
50nm
長形150bp ssDNA
感測器/DNA結合位點
105
20nm

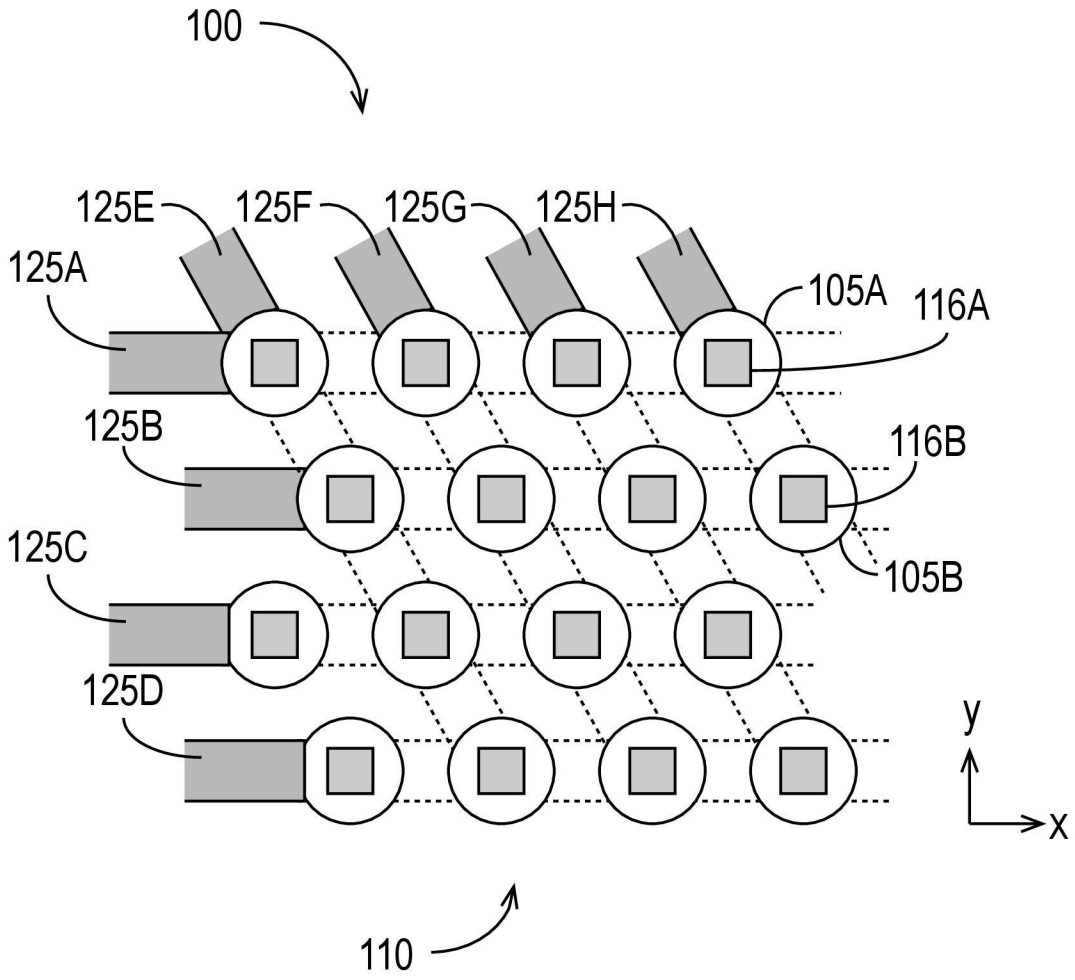
【圖6B】



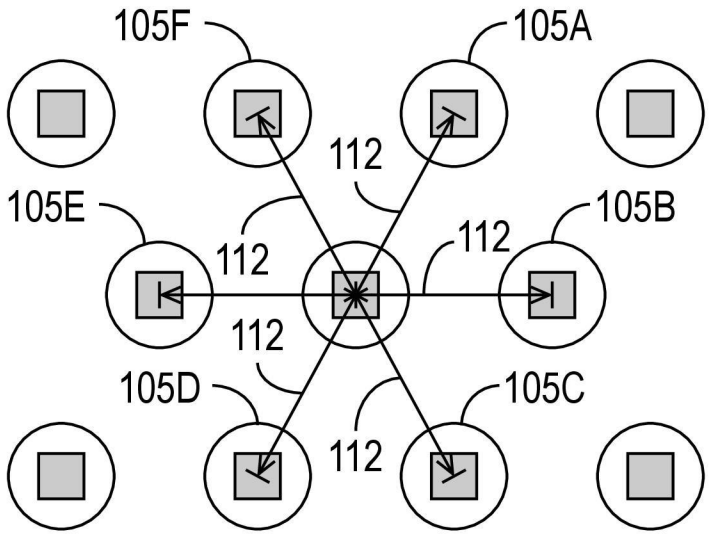
【圖7A】



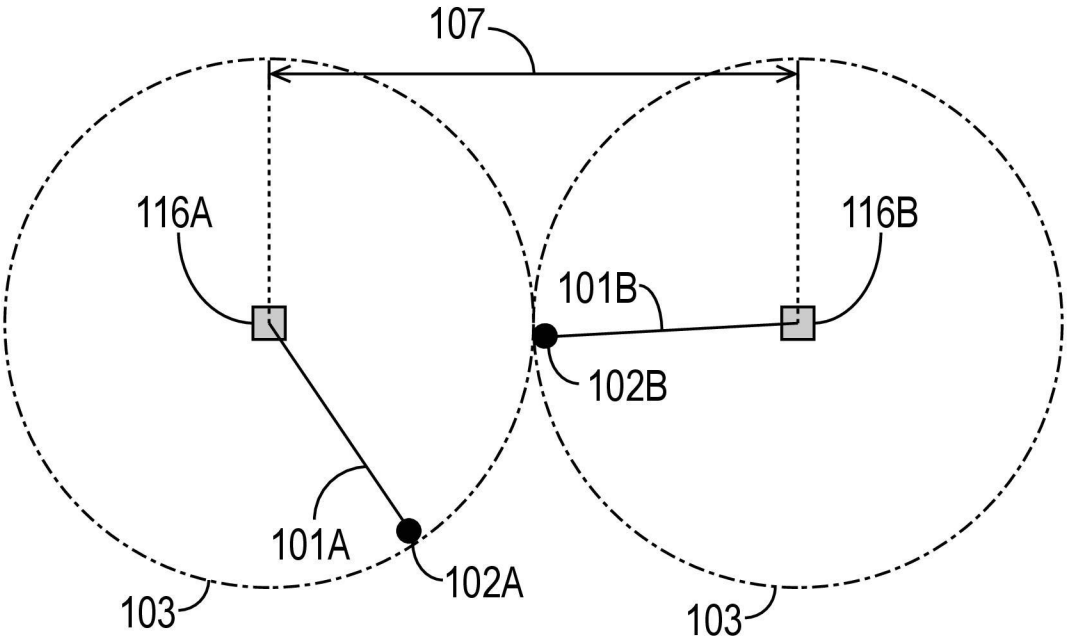
【圖7B】


100
125E
125F
125G
125H
125A
105A
116A
125B
116B
125C
105B
125D
y
x
110

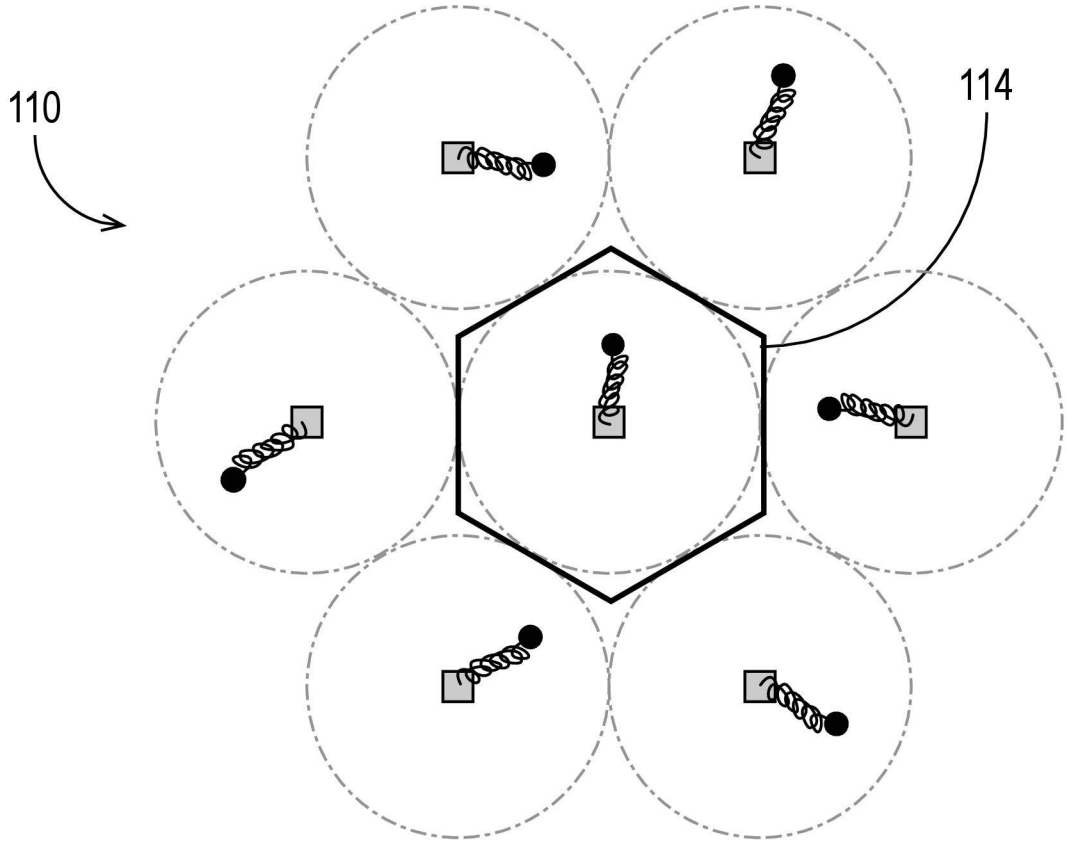
【圖8A】


105F
105A
112
105E
105B
112
112
112
105D
105C
112

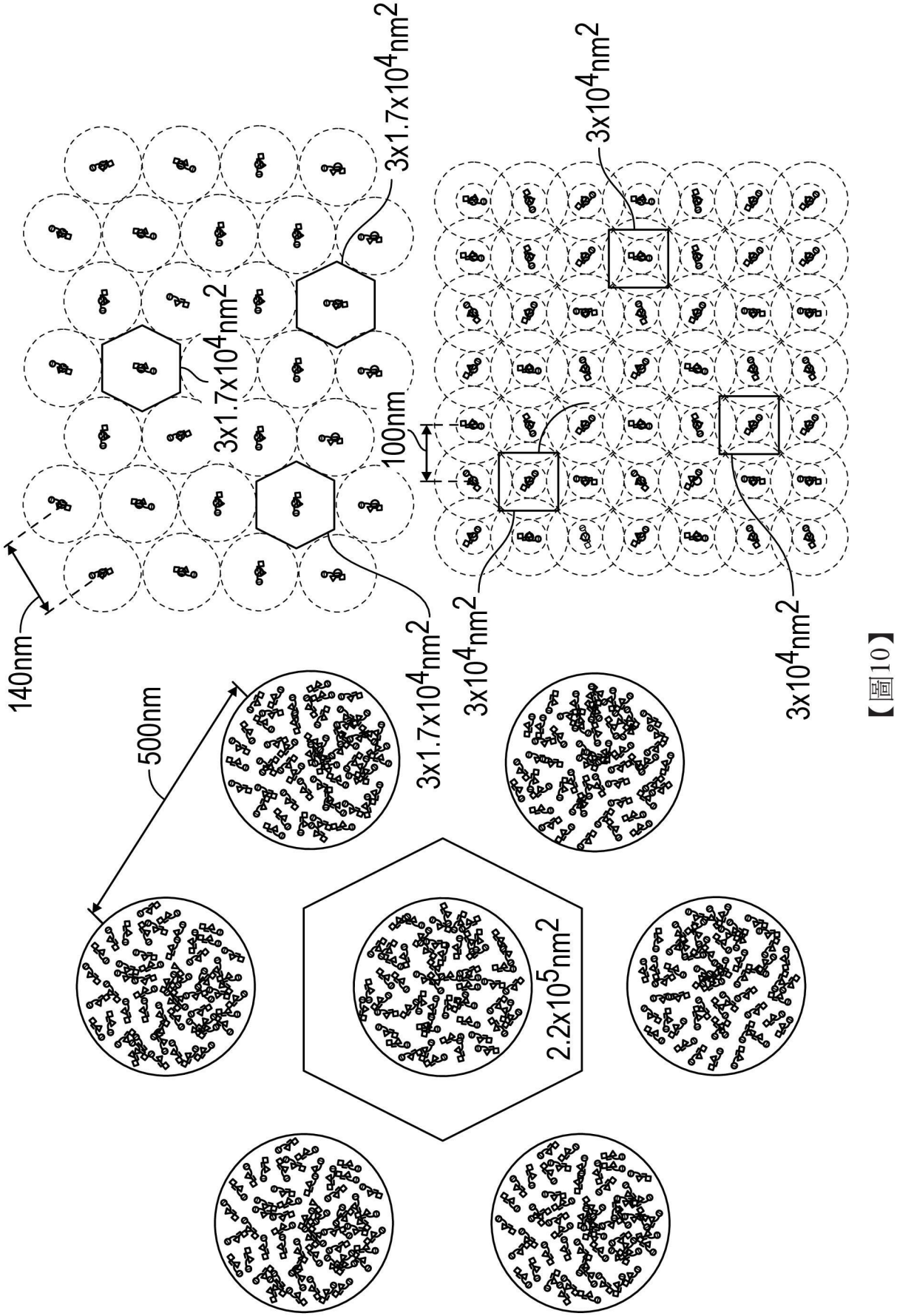
【圖8B】



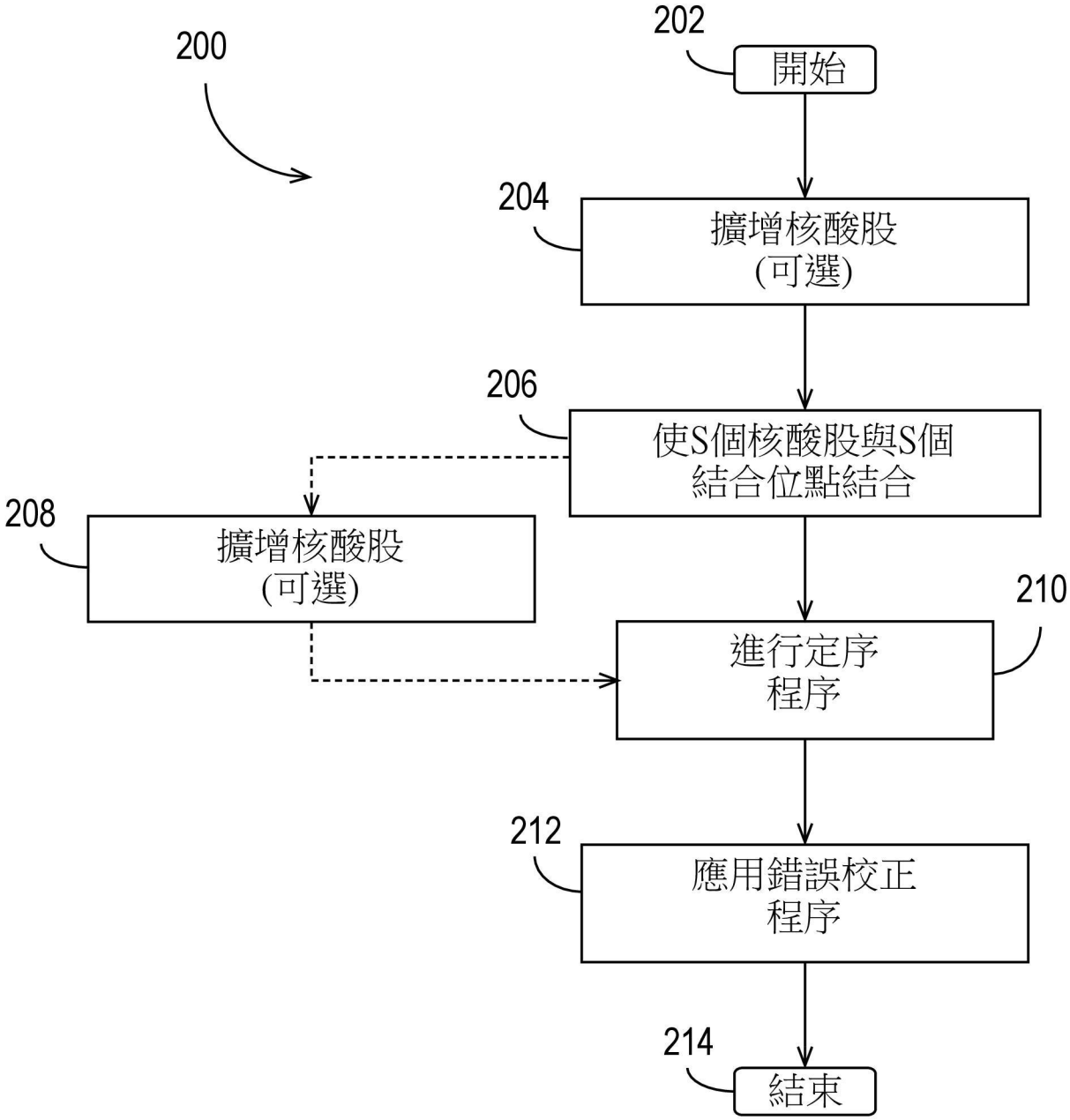
【圖9A】



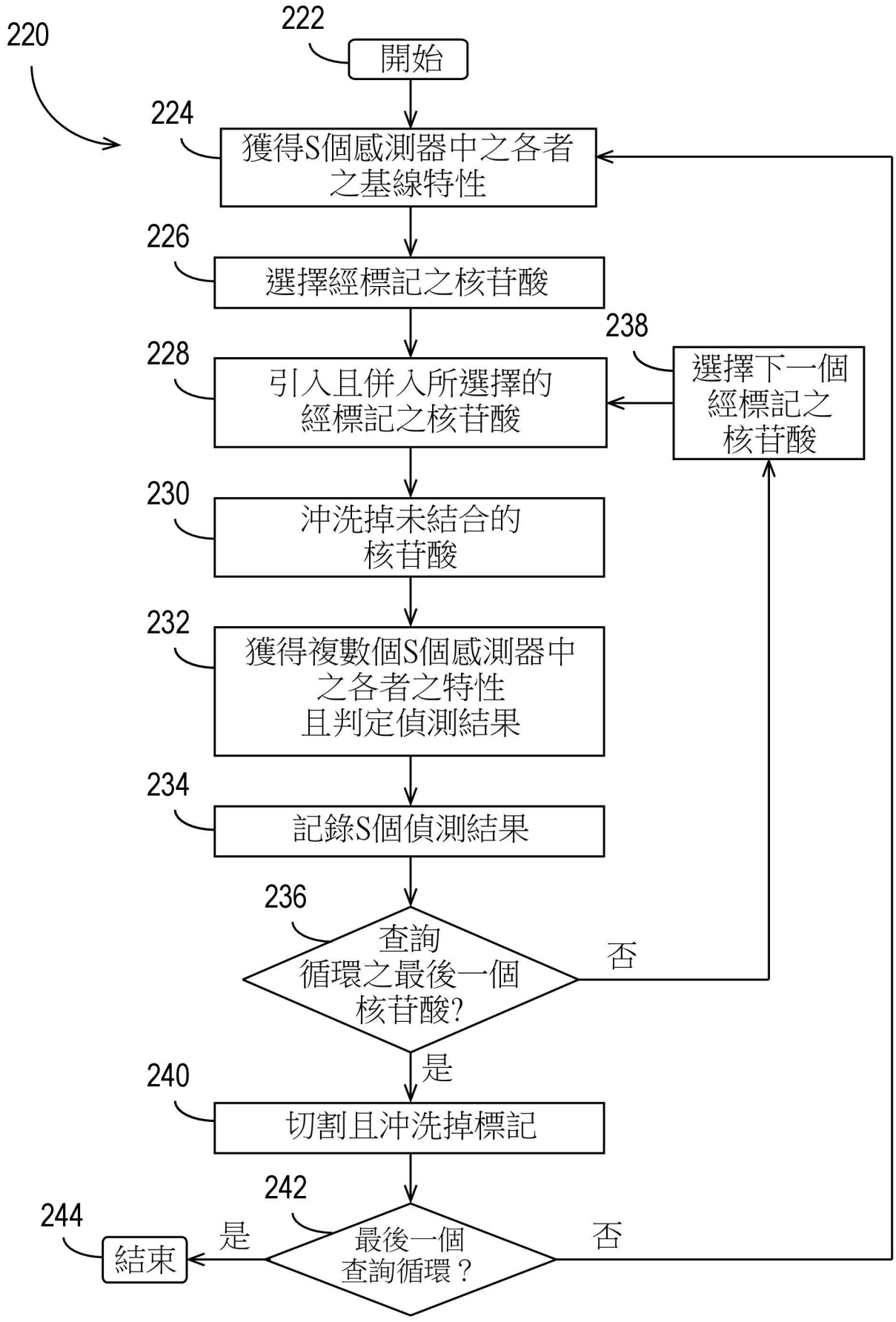
【圖9B】


3x1.7x10^4nm^2
3x10^4nm^2
3x1.7x10^4nm^2
100nm
140nm
3x10^4nm^2
3x10^4nm^2
500nm
3x1.7x10^4nm^2
2.2x10^5nm^2

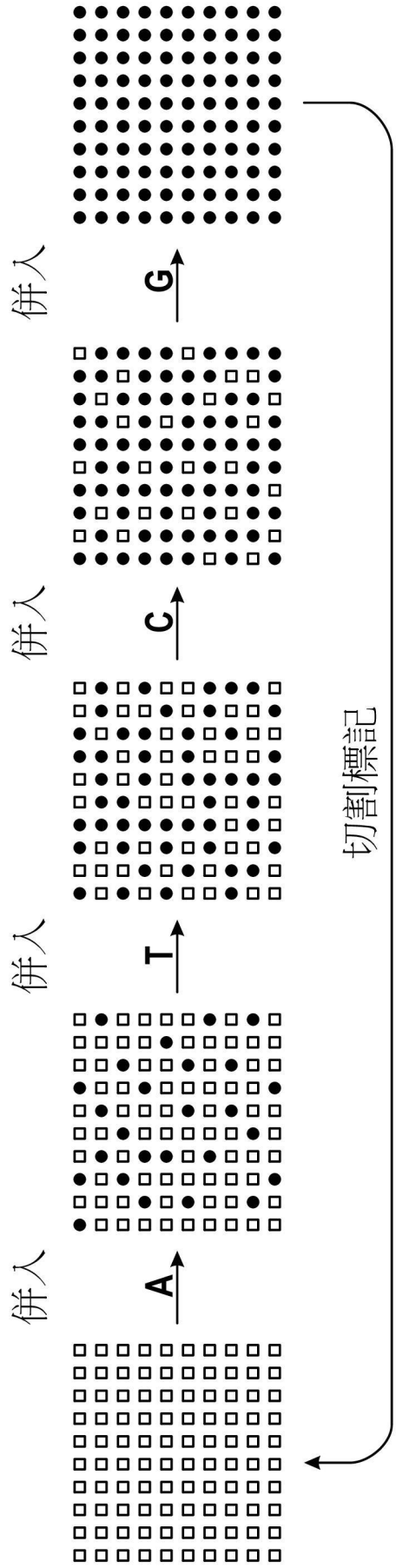
【圖10】


200
202
開始
204
擴增核酸股
(可選)
206
使S個核酸股與S個
結合位點結合
208
擴增核酸股
(可選)
210
進行定序
程序
212
應用錯誤校正
程序
214
結束

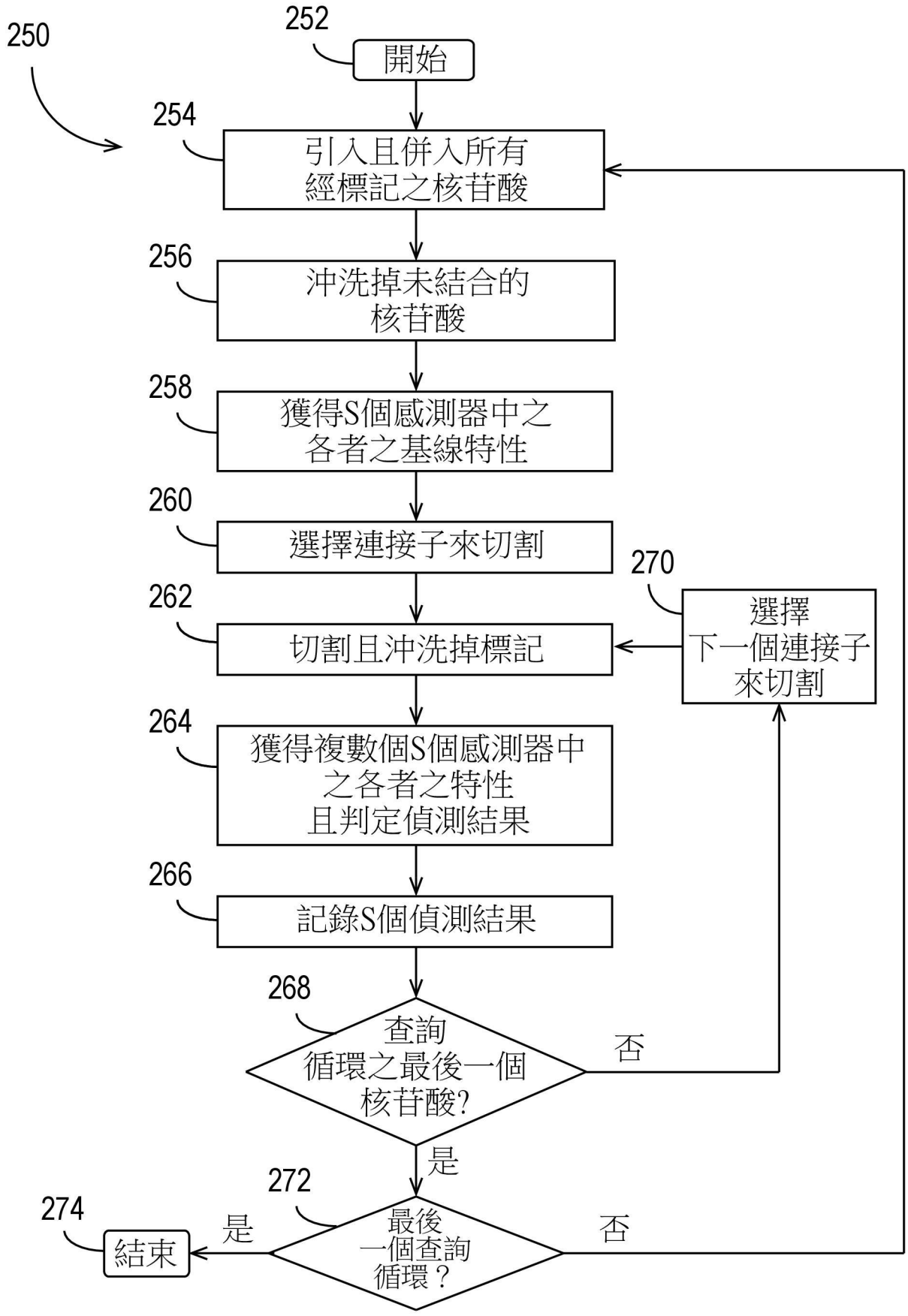
【圖11】


220
222
開始
224
獲得S個感測器中之各者之基線特性
226
選擇經標記之核苷酸
228
引入且併入所選擇的經標記之核苷酸
238
選擇下一個經標記之核苷酸
230
沖洗掉未結合的核苷酸
232
獲得複數個S個感測器中之各者之特性且判定偵測結果
234
記錄S個偵測結果
236
查詢循環之最後一個核苷酸?
否
是
240
切割且沖洗掉標記
242
最後一個查詢循環？
是
否
244
結束

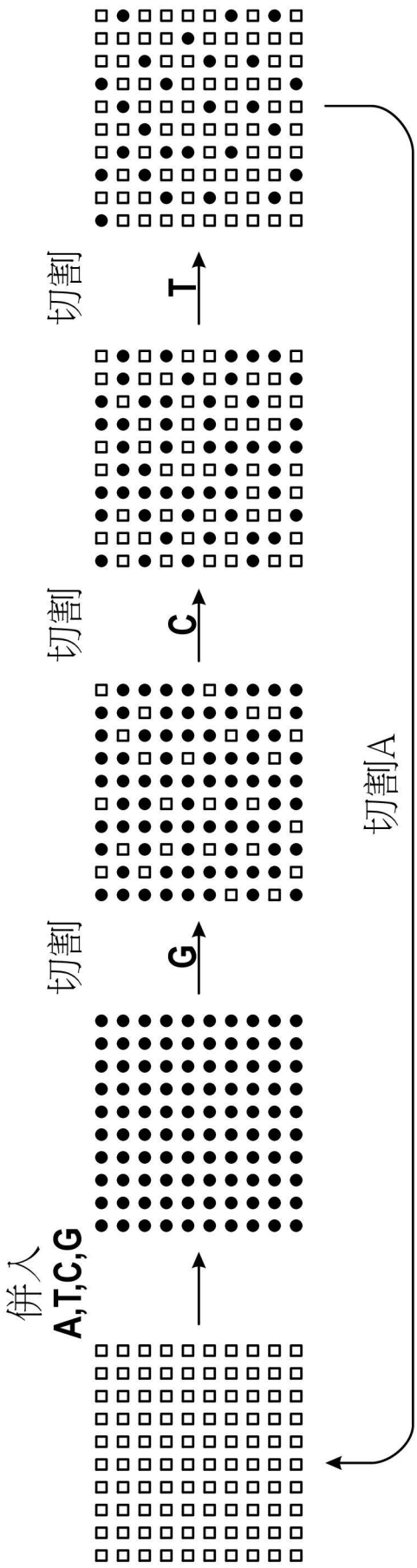
【圖12】



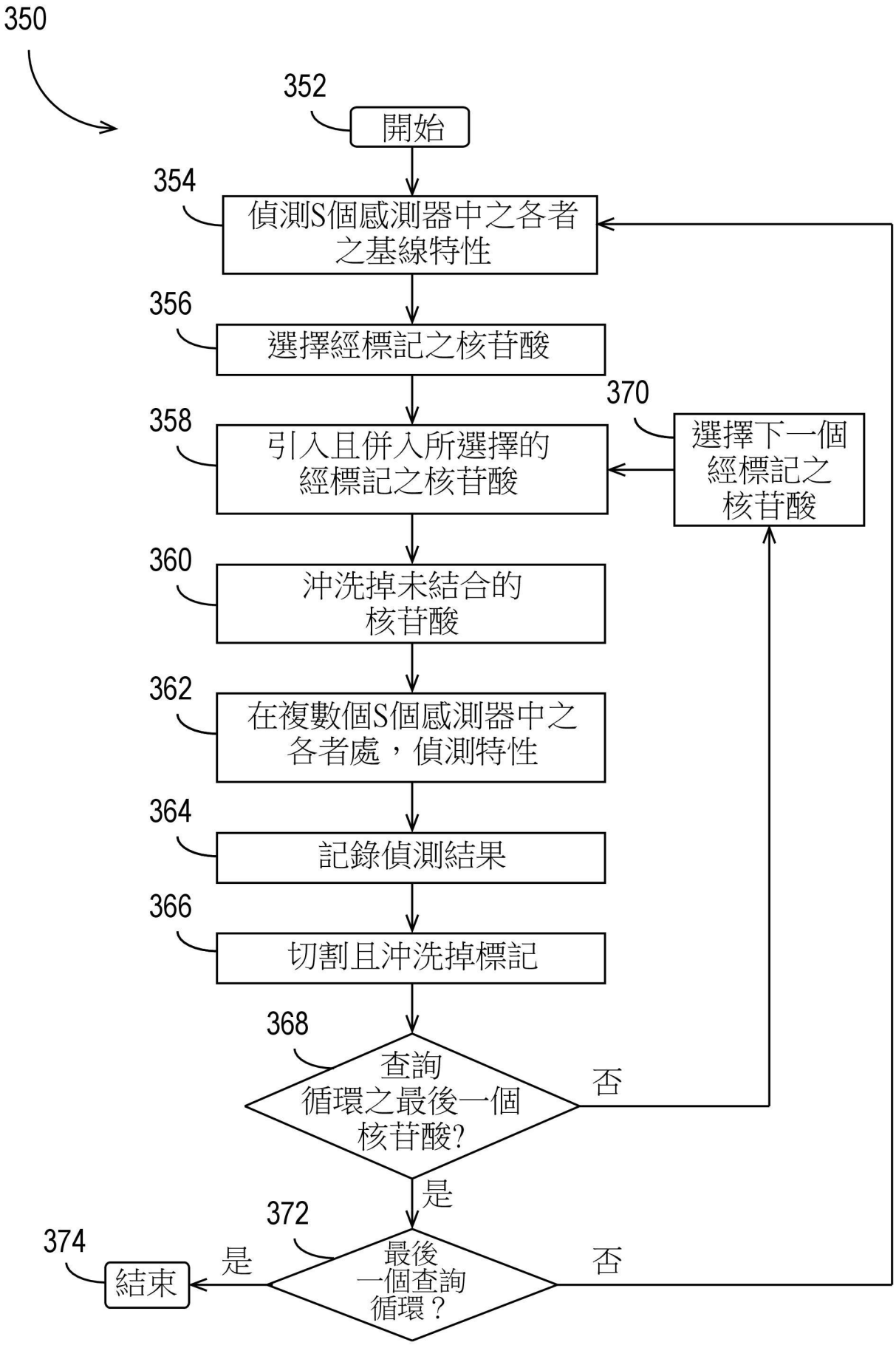
【圖13】


250
252
開始
254
引入且併入所有
經標記之核苷酸
256
沖洗掉未結合的
核苷酸
258
獲得S個感測器中之
各者之基線特性
260
選擇連接子來切割
270
選擇
下一個連接子
來切割
262
切割且沖洗掉標記
264
獲得複數個S個感測器中
之各者之特性
且判定偵測結果
266
記錄S個偵測結果
268
查詢
循環之最後一個
核苷酸?
否
是
272
最後
一個查詢
循環？
是
否
274
結束

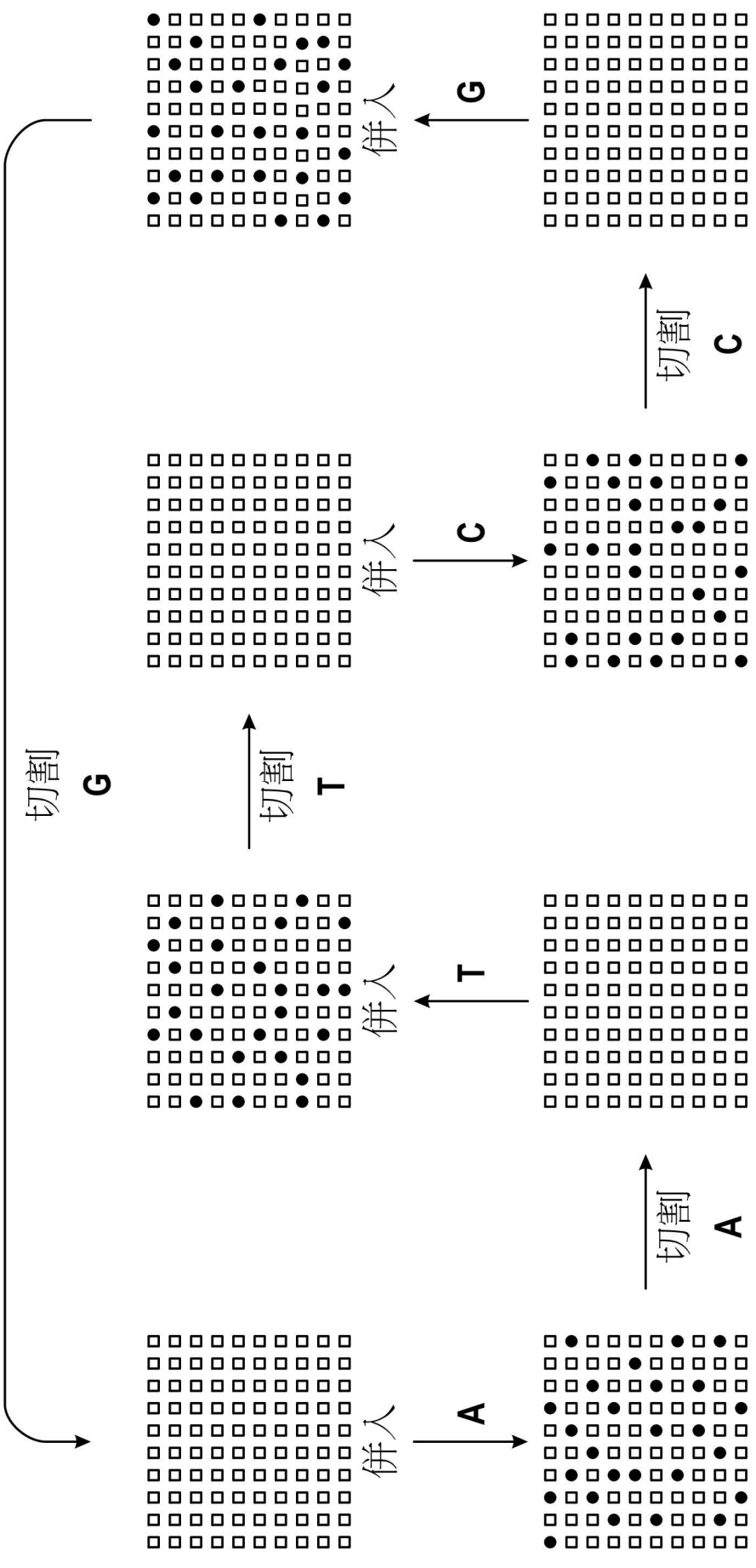
【圖14】



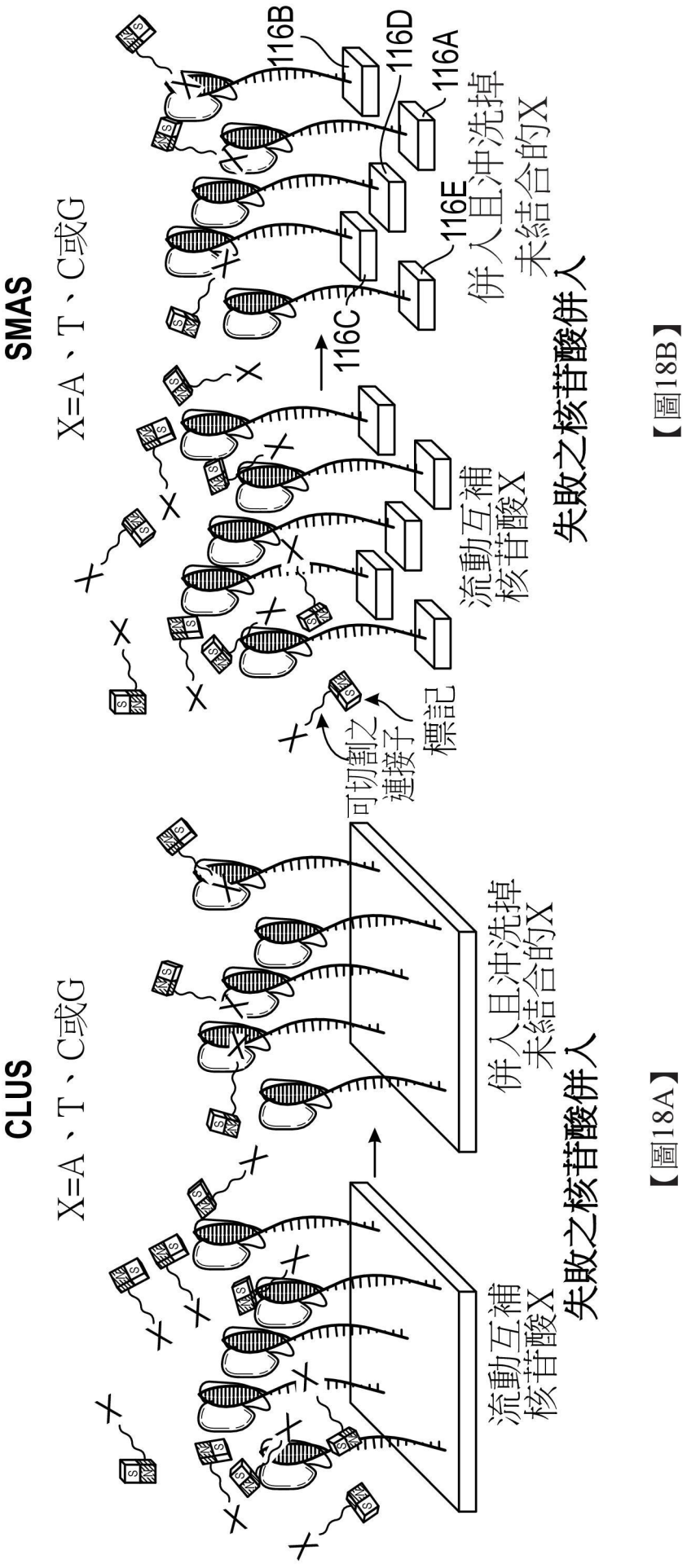
【圖15】


350
352
開始
354
偵測S個感測器中之各者之基線特性
356
選擇經標記之核苷酸
370
選擇下一個經標記之核苷酸
358
引入且併入所選擇的經標記之核苷酸
360
沖洗掉未結合的核苷酸
362
在複數個S個感測器中之各者處，偵測特性
364
記錄偵測結果
366
切割且沖洗掉標記
368
查詢循環之最後一個核苷酸?
否
是
372
最後一個查詢循環？
是
否
374
結束

【圖16】

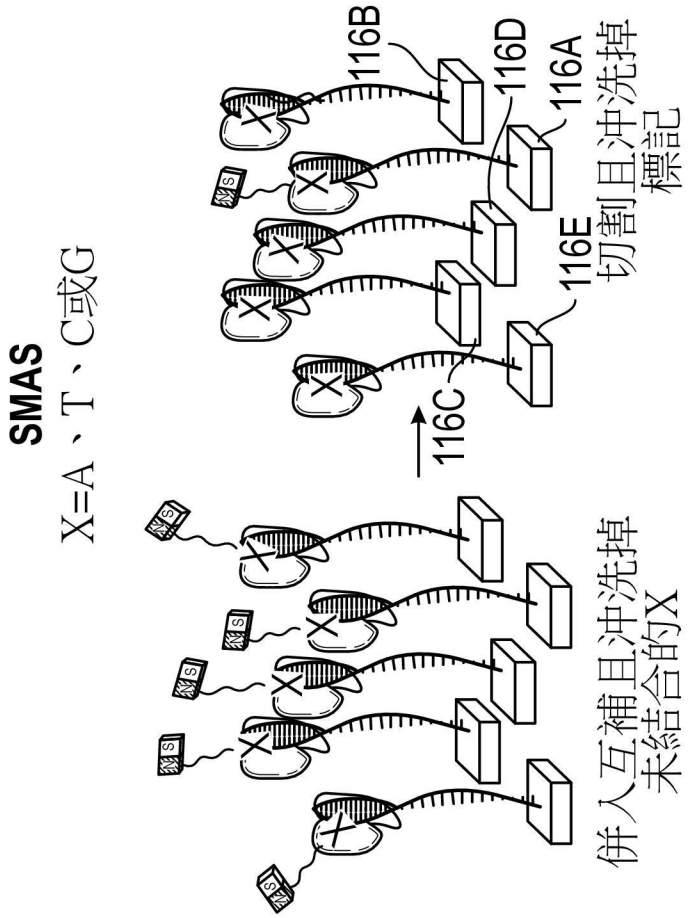


【圖17】

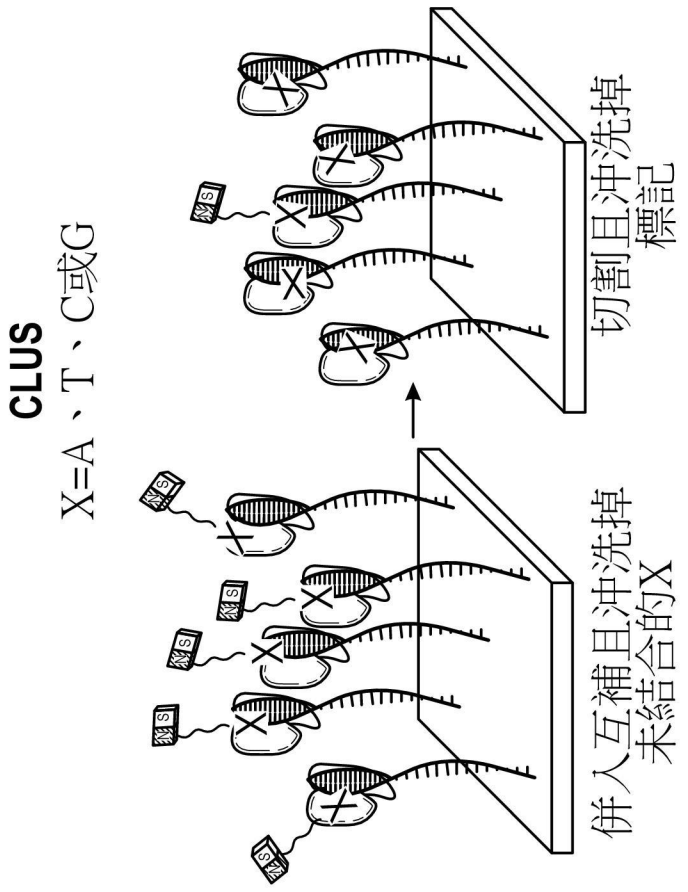

CLUS
X=A、T、C或G
流動互補
核苷酸X
併入且沖洗掉
未結合的X
失敗之核苷酸併入
SMAS
X=A、T、C或G
可切割之
連接子
標記
116C
116B
116D
116A
116E
流動互補
核苷酸X
併入且沖洗掉
未結合的X
失敗之核苷酸併入

【圖18A】

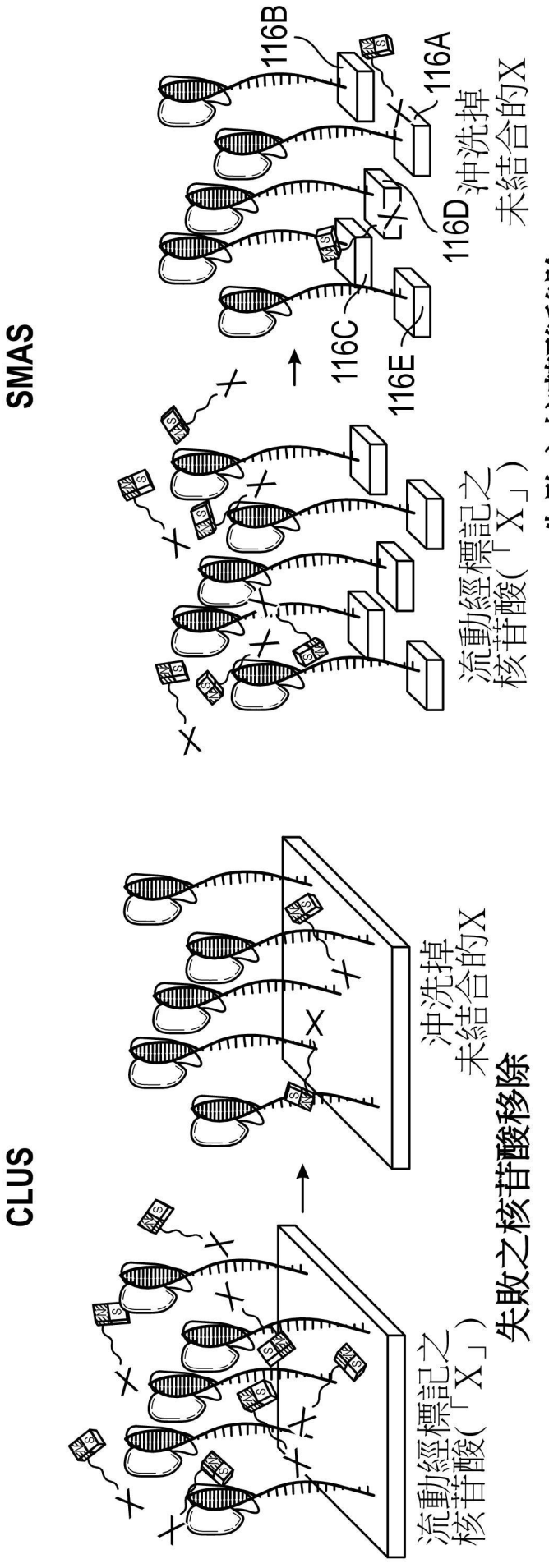
【圖18B】


SMAS
X=A、T、C或G
116B
116D
116A
116E
116C
切割且沖洗掉
標記
併入互補且沖洗掉
未結合的X
失敗之標記移除

【圖18D】

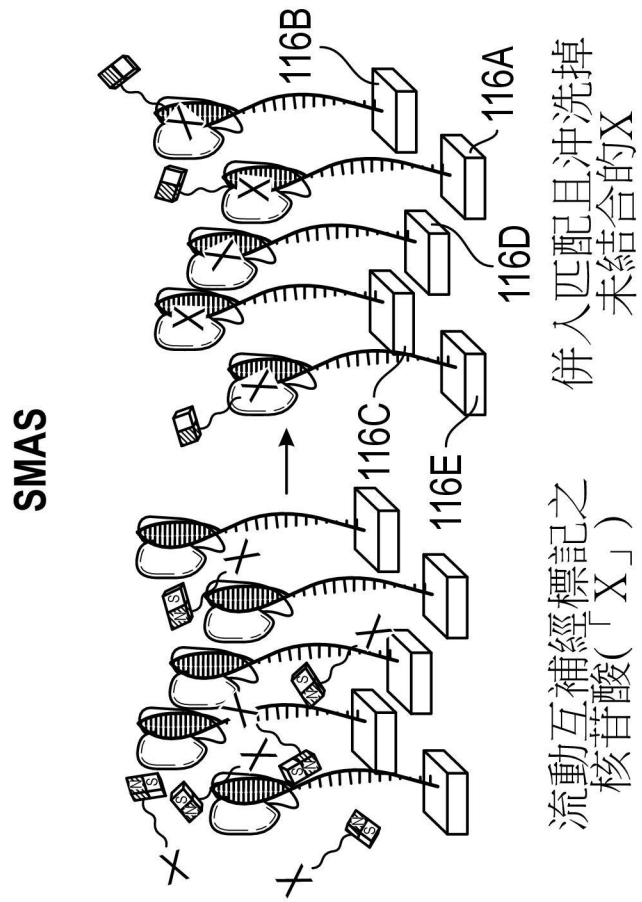

CLUS
X=A、T、C或G
切割且沖洗掉
標記
併入互補且沖洗掉
未結合的X
失敗之標記移除

【圖18C】


CLUS
流動經標記之
核苷酸(「X」)
沖洗掉
未結合的X
失敗之核苷酸移除
SMAS
流動經標記之
核苷酸(「X」)
116C
116E
116B
116A
116D
沖洗掉
未結合的X
失敗之核苷酸移除

【圖18E】

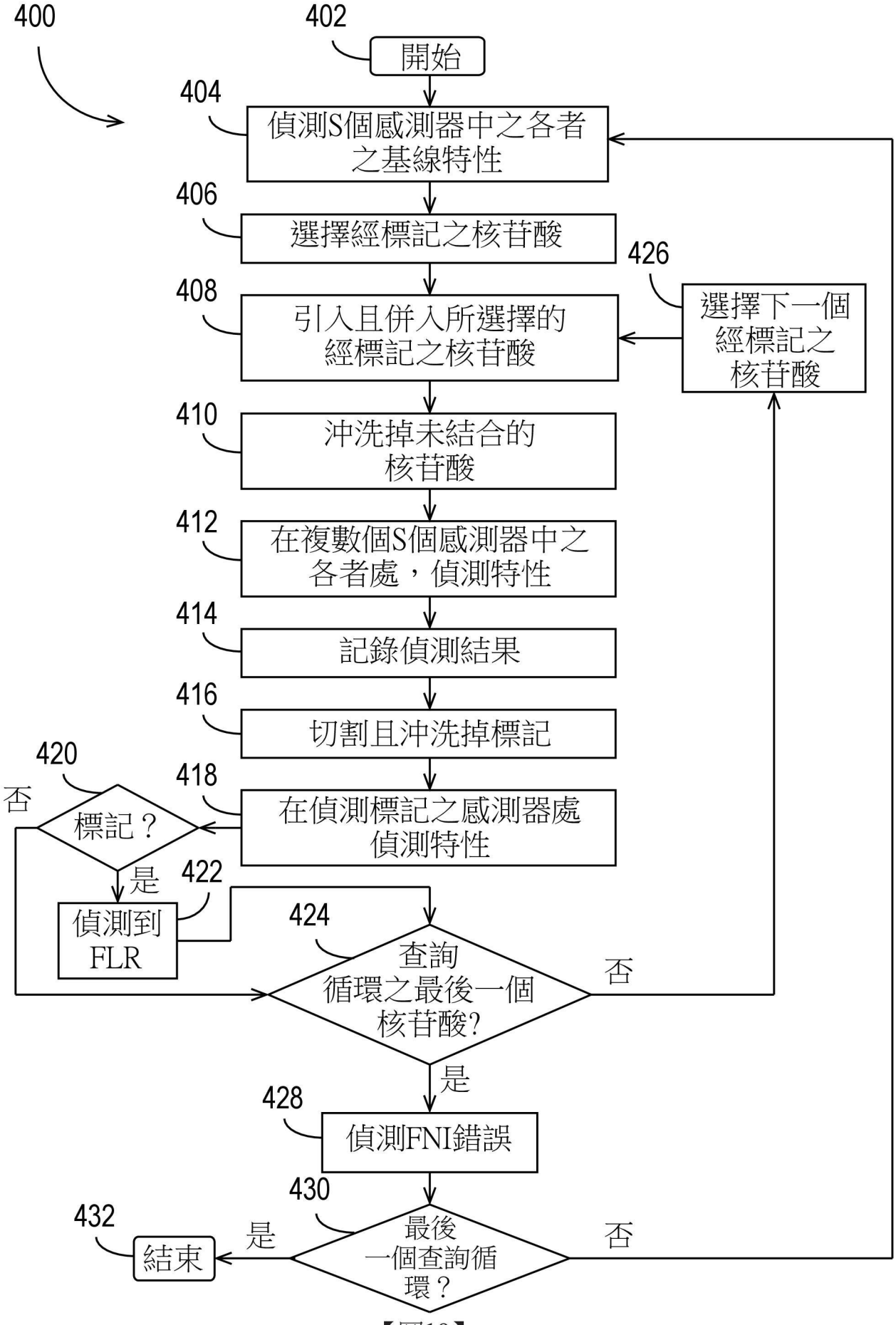
【圖18F】


CLUS
流動互補經標記之核苷酸(「X」)
併入匹配且沖洗掉未結合的X
失敗之核苷酸偵測

【圖18G】

SMAS
116B
116A
116D
116C
116E
流動互補經標記之核苷酸(「X」)
併入匹配且沖洗掉未結合的X
失敗之核苷酸偵測

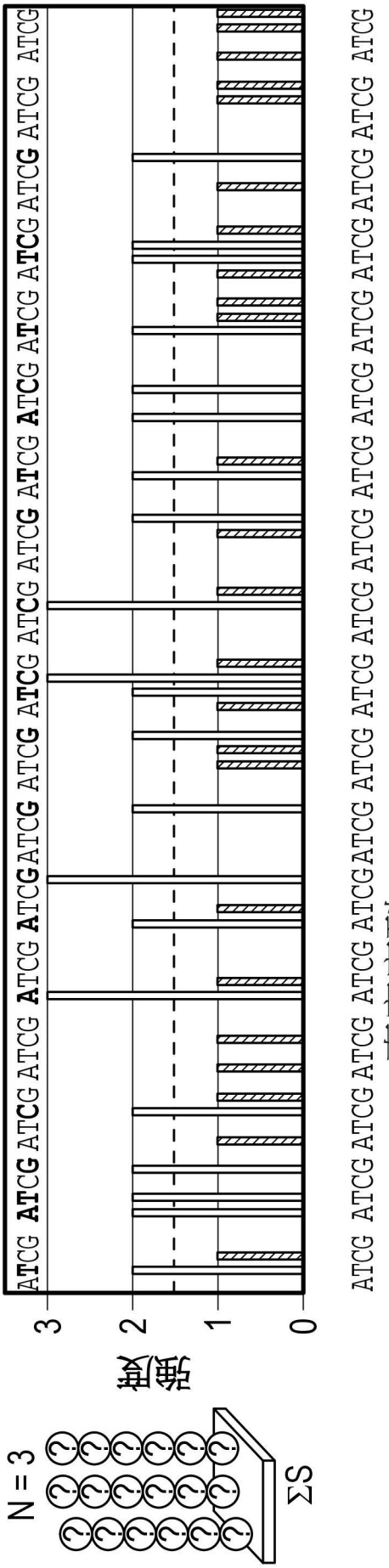
【圖18H】


400
402
開始
404
偵測S個感測器中之各者之基線特性
406
選擇經標記之核苷酸
426
選擇下一個經標記之核苷酸
408
引入且併入所選擇的經標記之核苷酸
410
沖洗掉未結合的核苷酸
412
在複數個S個感測器中之各者處，偵測特性
414
記錄偵測結果
416
切割且沖洗掉標記
420
否
418
在偵測標記之感測器處偵測特性
標記？
是
422
偵測到FLR
424
查詢循環之最後一個核苷酸?
否
是
428
偵測FNI錯誤
430
最後一個查詢循環？
否
是
432
結束

【圖19】

● 併入正確核苷酸
○ 未併入正確核苷酸
ʻ 失敗之核苷酸併入
⊘ 失敗之標記移除
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3

【圖20】


N = 3
ΣS
強度
3
2
1
0
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
真實序列: TAG CAA GGT CCG CTA CTG GCA GAC TGG
識別之序列: TAT GCA AGG GTC CGT ACT TCG
正確識別21個鹼基中的6個(~29%)

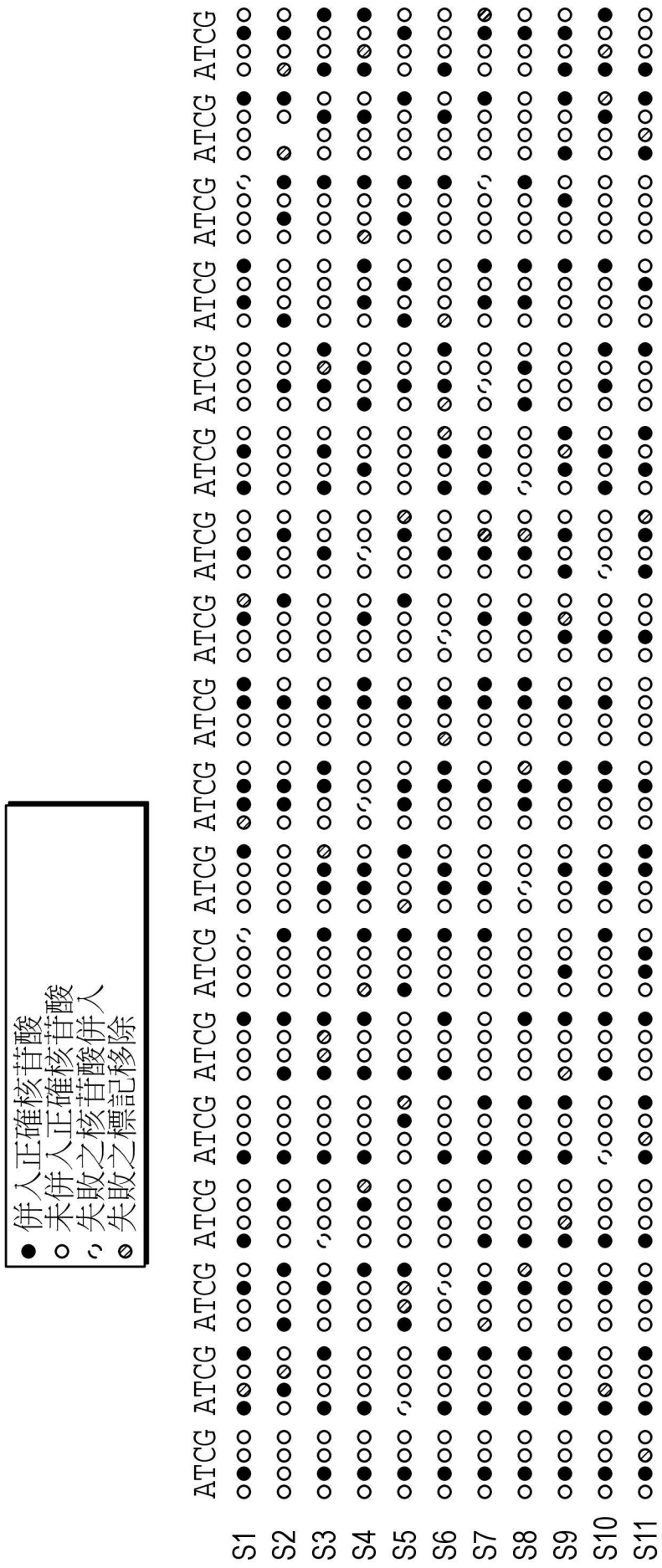
【圖21】

● 偵測到標記
○ 未偵測到標記
⊘ 失敗之標記移除
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
FLR錯誤校正
● 偵測到標記
○ 未偵測到標記
K = 3
S1
S2
S3
真實序列: TAG CAA GGT CCG CTA CTG GCA GAC TGG
識別之序列: TAG CAA GGT CCT ACT G
正確識別16個鹼基中的11個(~69%)

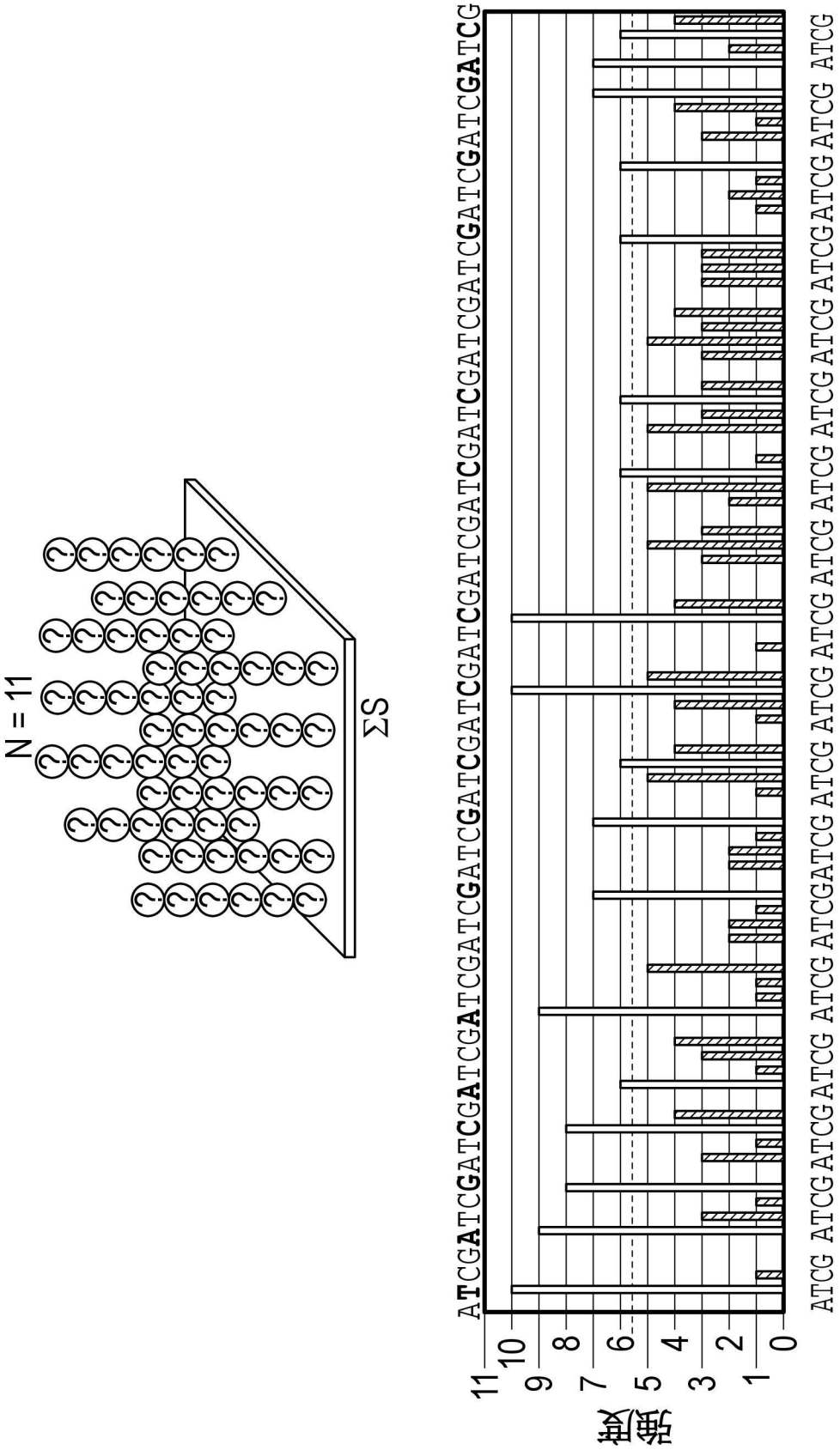
【圖22】

● 偵測到標記
○ 未偵測到標記
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
FNI錯誤校正
K = 3
S1
S2
S3
● 偵測到標記
○ 未偵測到標記
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
真實序列: TAG CAA GGT CCG CTA CTG GC
識別之序列: TAG CAA GGT CCG CTA CTG GC
正確識別20個鹼基中的20個(~100%)

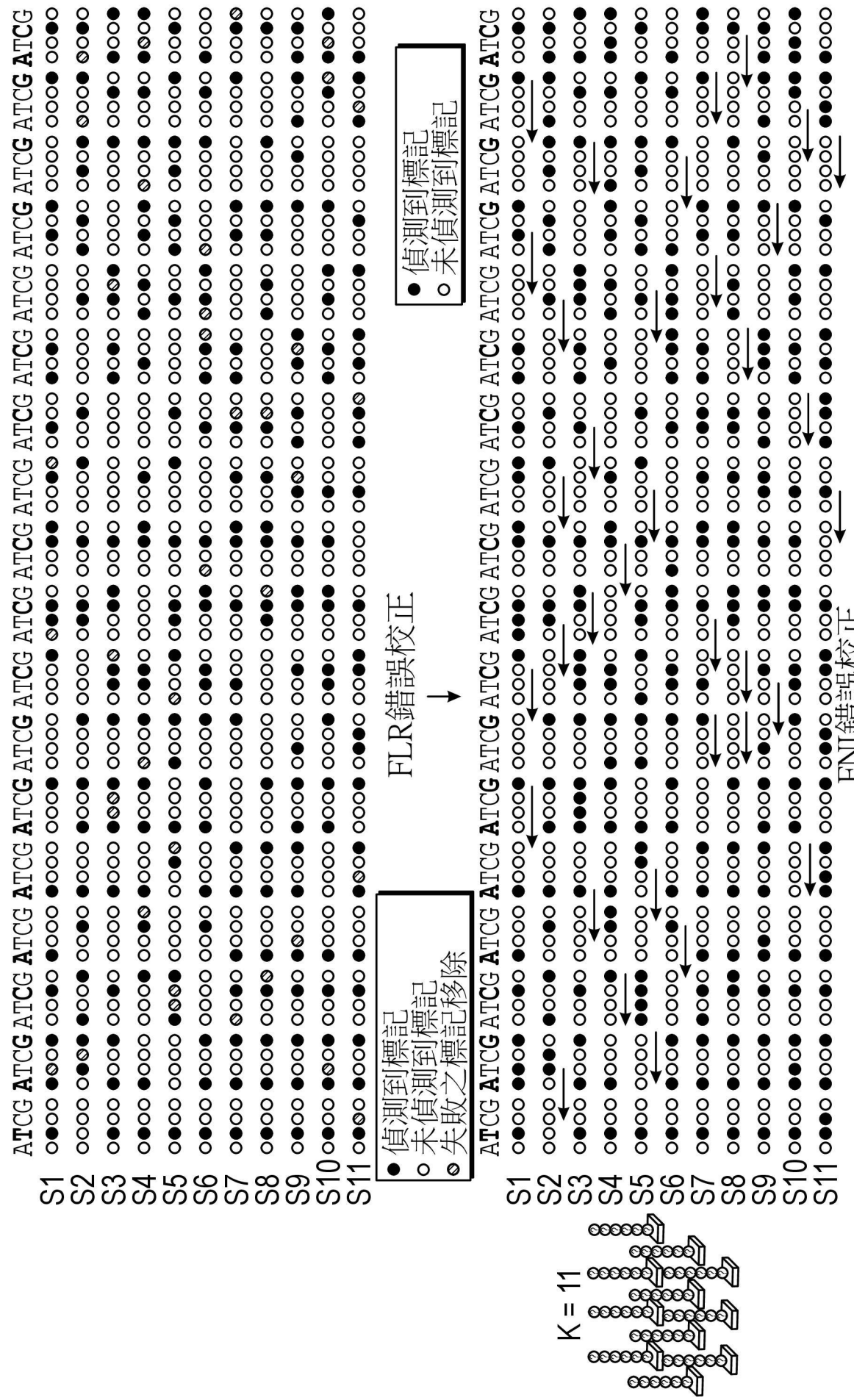
【圖23】


● 併入正確核苷酸
○ 未併入正確核苷酸
ʃ 失敗之核苷酸併入
⊘ 失敗之標記移除
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11

【圖24】


N = 11
ΣS
強度
0
1
2
3
4
5
6
7
8
9
10
11
ATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG
ATCG ATCG ATCG ATCG ATCG ATCGATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
真實序列: TAG CAA GGT CCG CTA CTG
識別之序列: TAG CAA GGC CCC CGG GAC
正確識別18個鹼基中的11個(~61%)

【圖25】


ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
● 偵測到標記
○ 未偵測到標記
⊘ 失敗之標記移除
FLR錯誤校正
● 偵測到標記
○ 未偵測到標記
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
K = 11
FNI錯誤校正
下一頁

【圖26】

● 偵測到標記
○ 未偵測到標記
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
真實序列: TAG CAA GGT CCG CTA CTG GCA G
識別之序列: TAG CAA GGT CCG CTA CTG GCA G
正確識別22個鹼基中的22個(~100%)

【圖26】
(續圖)

● 併入正確核苷酸
○ 未併入正確核苷酸
ↄ 失敗之核苷酸併入
⊘ 失敗之標記移除
⊘ 失敗之核苷酸移除
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3

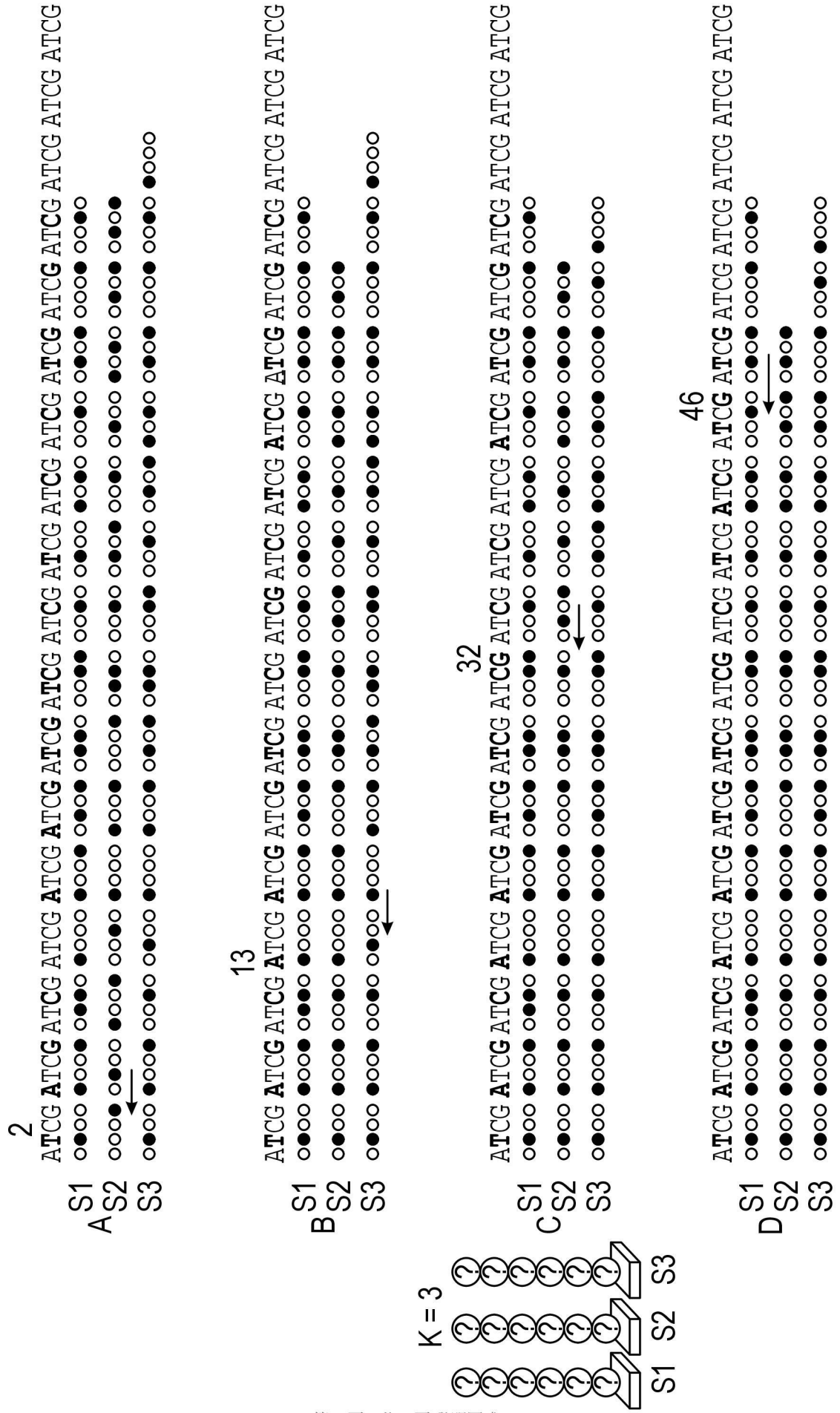
【圖27】

● 偵測到標記
○ 未偵測到標記
⊘ 失敗之標記移除
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
FLR錯誤校正
● 偵測到標記
○ 未偵測到標記
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
K = 3
S1
S2
S3
真實序列: TAG CAA GGT CCG CTA CTG GCA GAC TGG
識別之序列: TAG CAA GTG TCC TAC TG
正確識別17個鹼基中的8個(~47%)

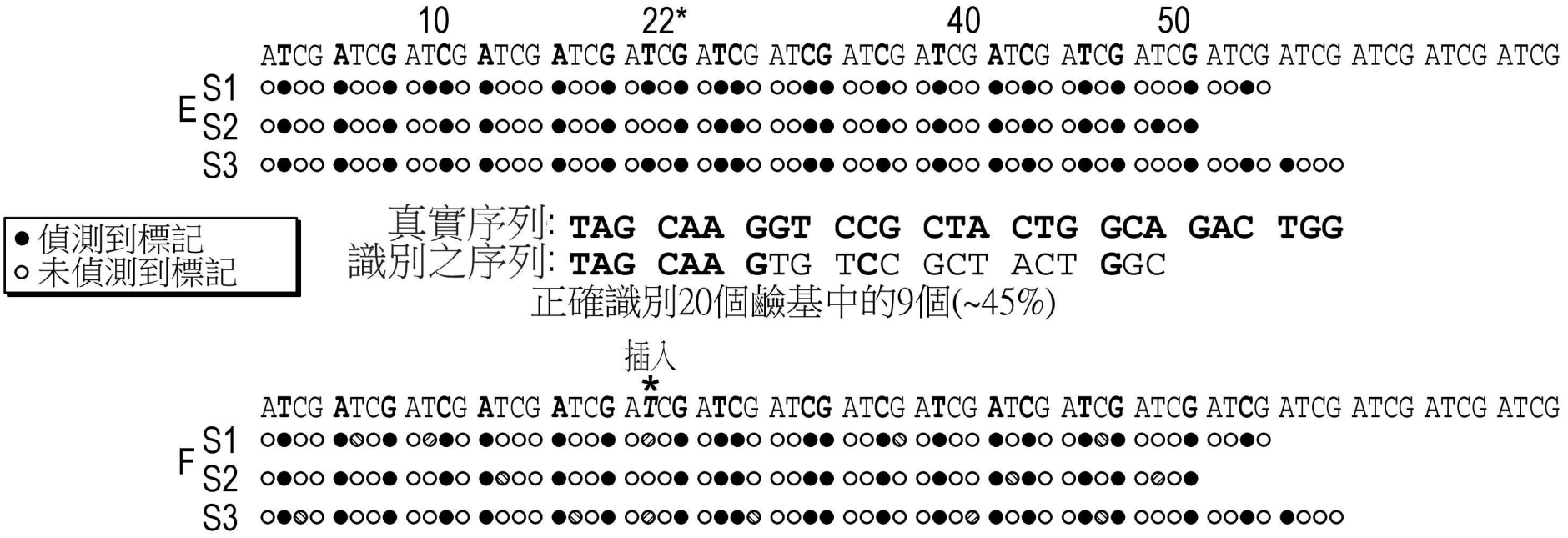
【圖28】

● 偵測到標記
○ 未偵測到標記
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
FNI錯誤校正
● 偵測到標記
○ 未偵測到標記
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
K = 3
S1 S2 S3
真實序列: TAG CAA GGT CCG CTA CTG GCA GAC TGG
識別之序列: TAG CAA GTG TCC TCC TGG CG
正確識別20個鹼基中的9個(~45%)

【圖29】


2
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
A
13
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
B
32
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
C
46
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3
D
K = 3
S1 S2 S3

【圖30】


10
22*
40
50
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
E S1
S2
S3
● 偵測到標記
○ 未偵測到標記
真實序列: TAG CAA GGT CCG CTA CTG GCA GAC TGG
識別之序列: TAG CAA GTG TCC GCT ACT GGC
正確識別20個鹼基中的9個(~45%)
插入
*
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
F S1
S2
S3
● 併入正確核苷酸
○ 未併入正確核苷酸
⊘ 失敗之標記移除
⊘ 失敗之核苷酸移除

【圖30】
(續圖)

● 併入正確核苷酸
○ 未併入正確核苷酸
⊘ 失敗之標記移除
ↄ 失敗之核苷酸併入
⊘ 失敗之核苷酸移除
⊖ 失敗之標記偵測
循環
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
S1
S2
S3

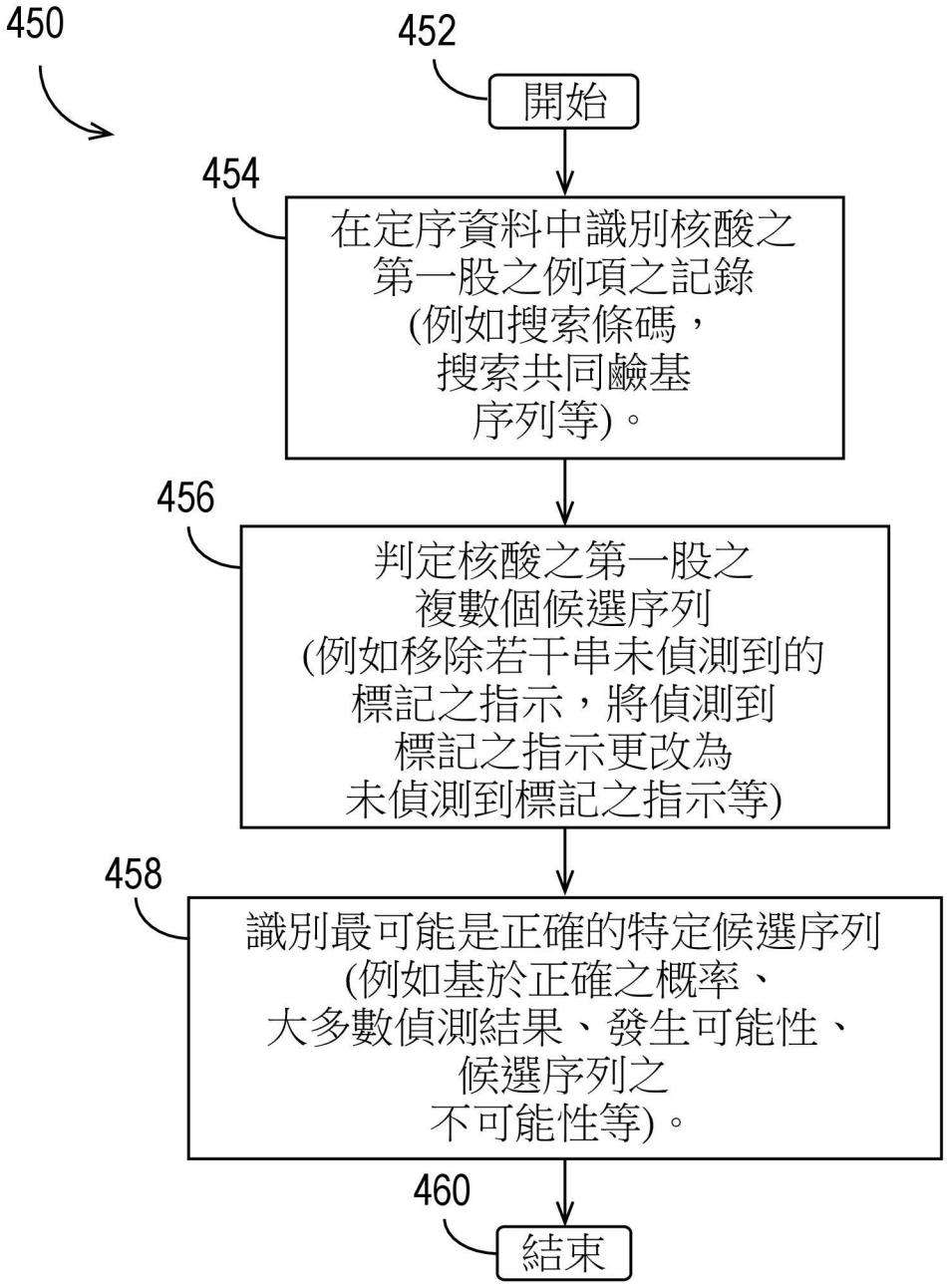
【圖31】

A
B
S1
S2
S3
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
● 偵測到標記
○ 未偵測到標記
⊘ 失敗之標記移除
真實序列: TAG CAA GGT CCG CTA CTG GCA GAC TGG
識別之序列: TAG CAA GTG TCT ACT TG
正確識別17個鹼基中的8個(~47%)
算法錯誤校正

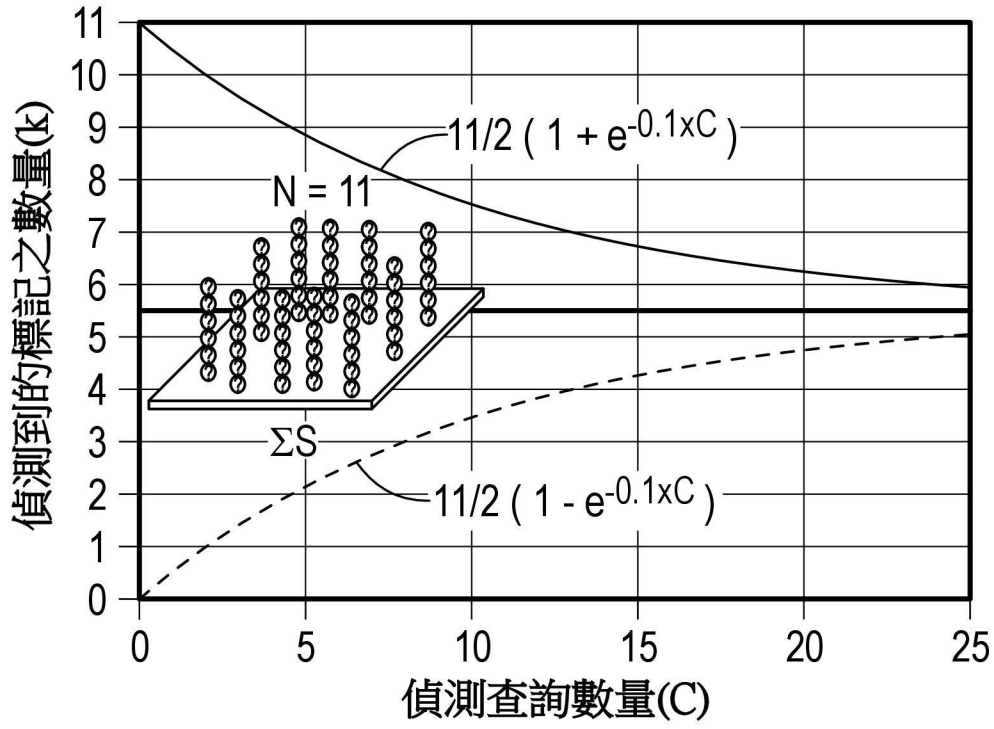
【圖32】

5 10 13 20 22* 27 32* 40 41 48 50
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
C S1
S2
S3
真實序列: TAG CAA GGT CCG CTA CTG GCA GAC TGG
識別之序列: TAG CAA GTG TCC CTA CTG GC
正確識別20個鹼基中的16個(~80%)
偵測到標記
未偵測到標記
K = 3
S1 S2 S3
插入
刪除
ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG ATCG
D S1
S2
S3
併入正確核苷酸
未併入正確核苷酸
失敗之標記移除
失敗之核苷酸移除
失敗之標記偵測

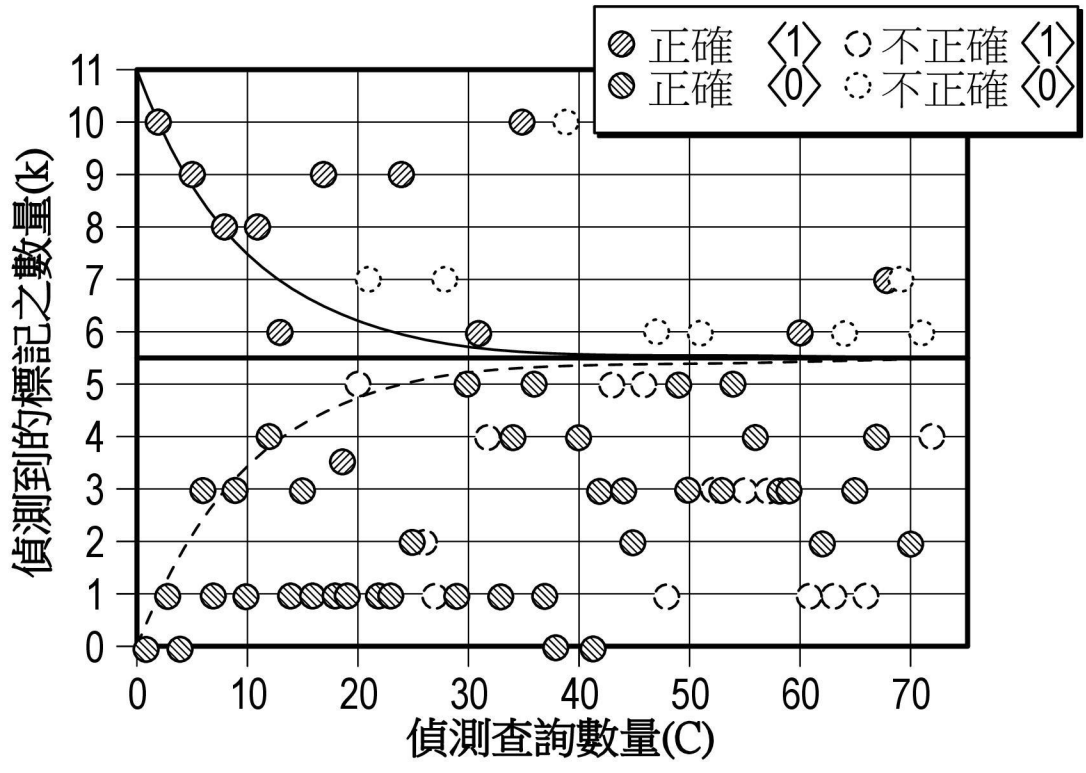
【圖32】
(續圖)


450
452
開始
454
在定序資料中識別核酸之
第一股之例項之記錄
(例如搜索條碼，
搜索共同鹼基
序列等)。
456
判定核酸之第一股之
複數個候選序列
(例如移除若干串未偵測到的
標記之指示，將偵測到
標記之指示更改為
未偵測到標記之指示等)
458
識別最可能是正確的特定候選序列
(例如基於正確之概率、
大多數偵測結果、發生可能性、
候選序列之
不可能性等)。
460
結束

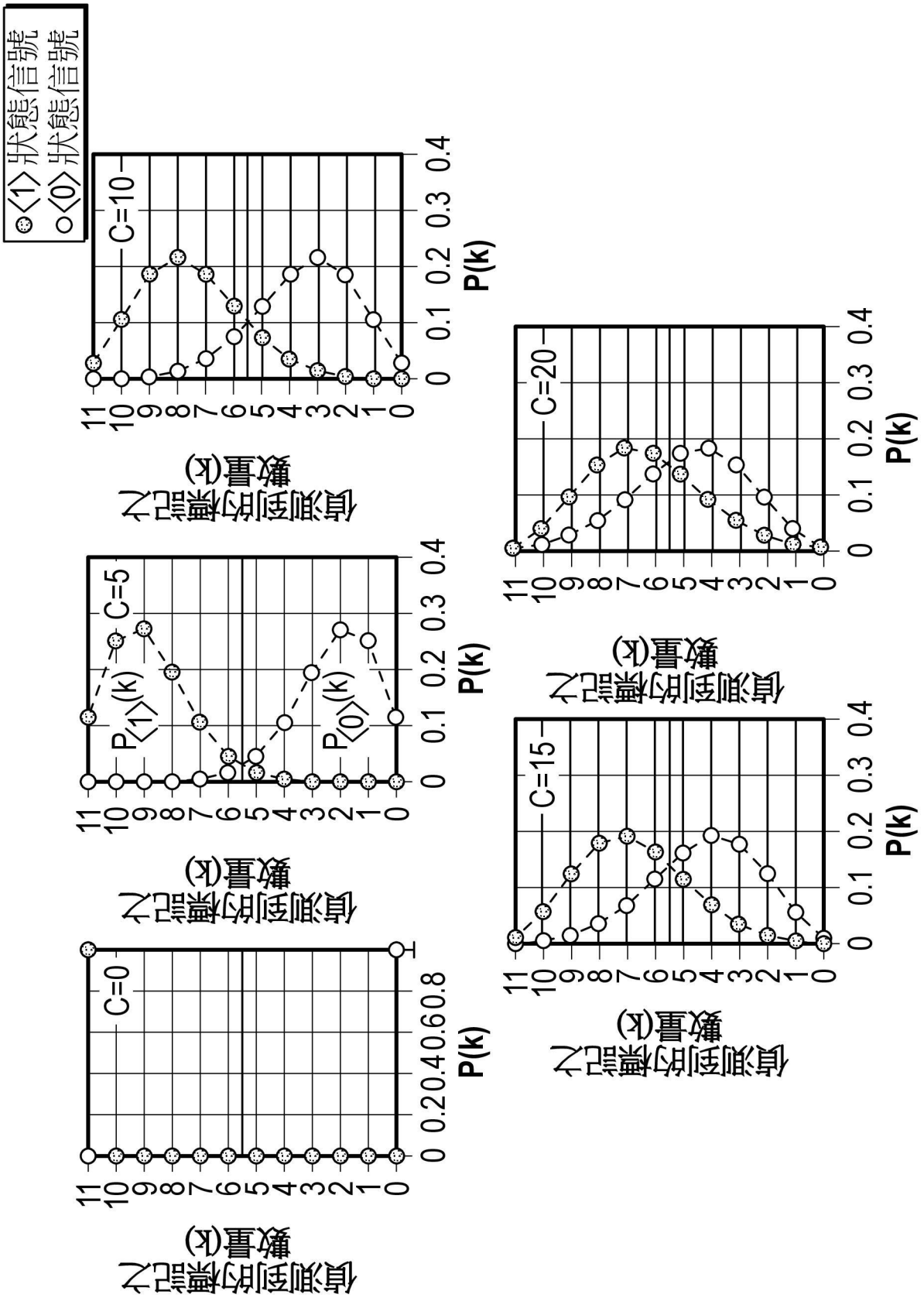
【圖33】


11/2 (1 + e-0.1xC)
N = 11
ΣS
11/2 (1 - e-0.1xC)
偵測到的標記之數量(k)
偵測查詢數量(C)

【圖34A】


正確 〈1〉
不正確 〈1〉
正確 〈0〉
不正確 〈0〉
偵測到的標記之數量(k)
偵測查詢數量(C)

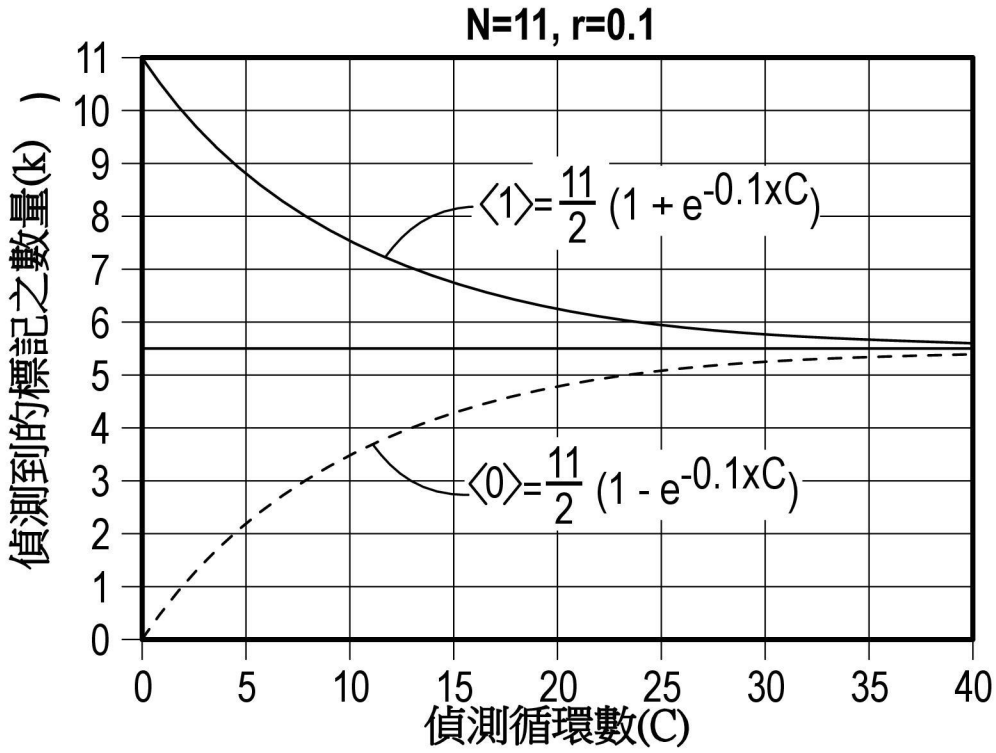
【圖34B】

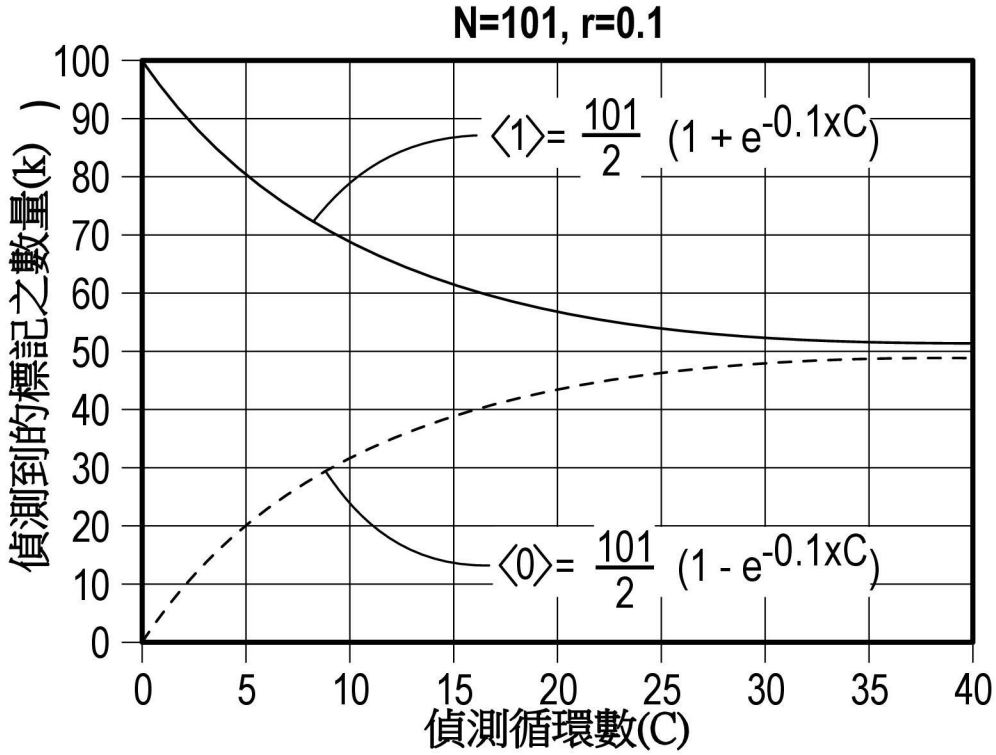

◎〈1〉狀態信號
○〈0〉狀態信號
C=0
C=5
C=10
C=15
C=20
$P_{\langle 1\rangle}(k)$
$P_{\langle 0\rangle}(k)$
P(k)
偵測到的標記之數量(k)

【圖35】

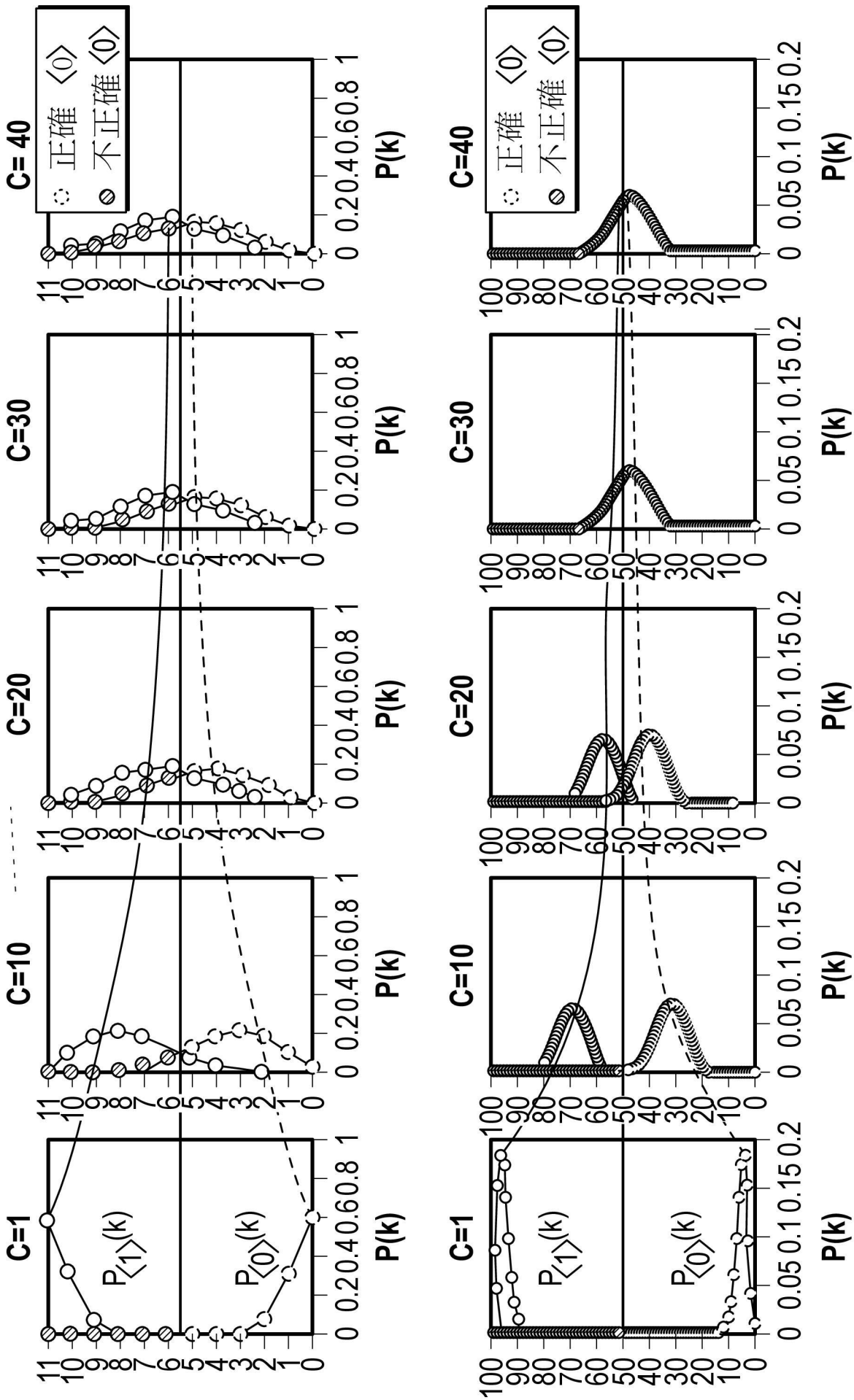
不正確〈1〉
不正確〈0〉
C=0
C=5
C=10
C=15
C=20
$P_{\langle 1\rangle}(k)$
$P_{\langle 0\rangle}(k)$
P(k)
偵測到的標記之數量(k)

【圖36】

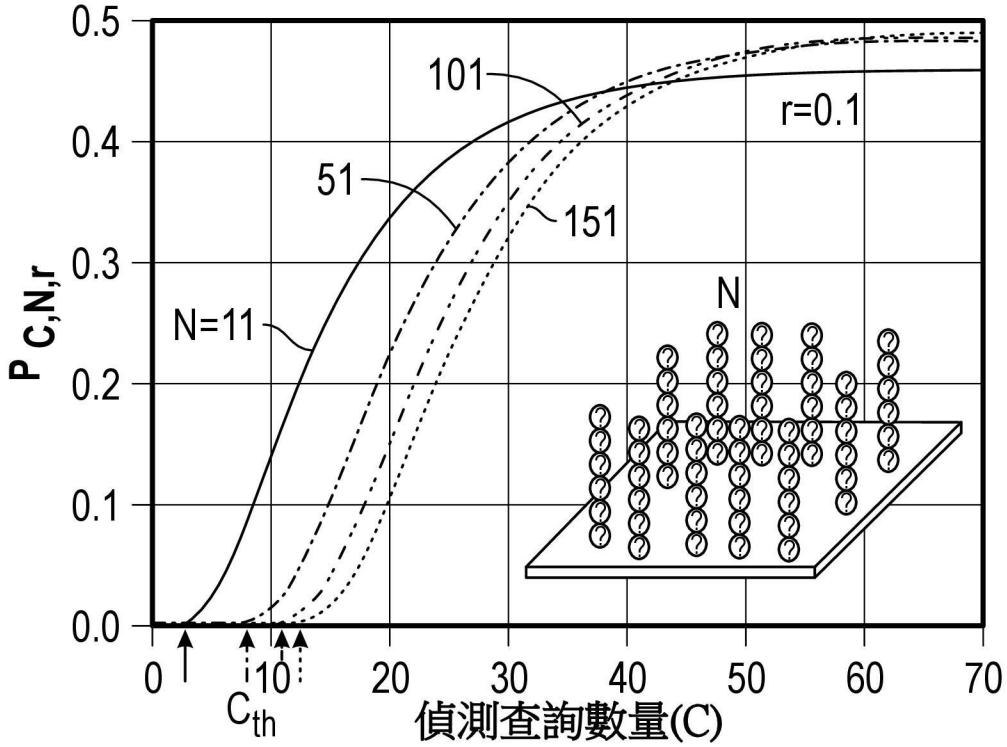

N=11, r=0.1
偵測到的標記之數量(k)
偵測循環數(C)
〈1〉= 11/2 (1 + e-0.1xC)
〈0〉= 11/2 (1 - e-0.1xC)
0
1
2
3
4
5
6
7
8
9
10
11
0
5
10
15
20
25
30
35
40


N=101, r=0.1
偵測到的標記之數量(k)
偵測循環數(C)
〈1〉= 101/2 (1 + e-0.1xC)
〈0〉= 101/2 (1 - e-0.1xC)
0
10
20
30
40
50
60
70
80
90
100
0
5
10
15
20
25
30
35
40

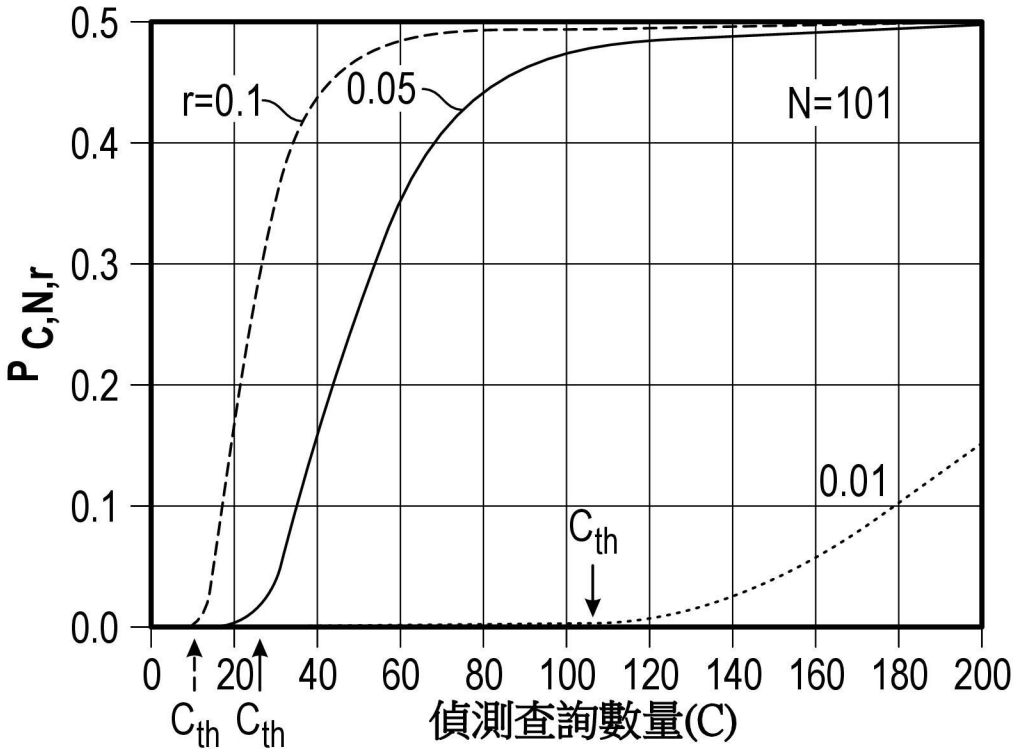
【圖37A】



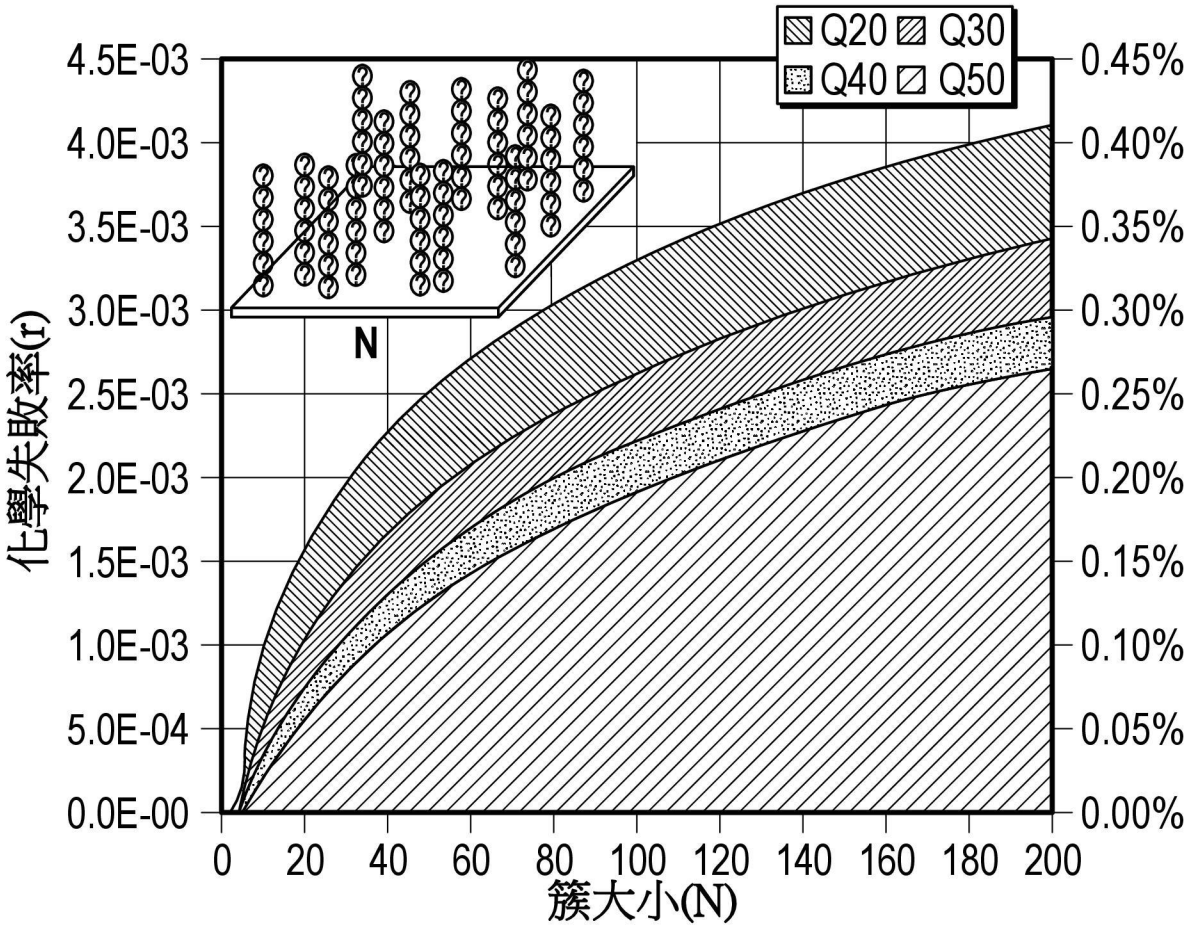
【圖37B】


0.5
0.4
0.3
0.2
0.1
0.0
$P_{C,N,r}$
101
r=0.1
51
151
N
N=11
0
10
20
30
40
50
60
70
C_{th}
偵測查詢數量(C)

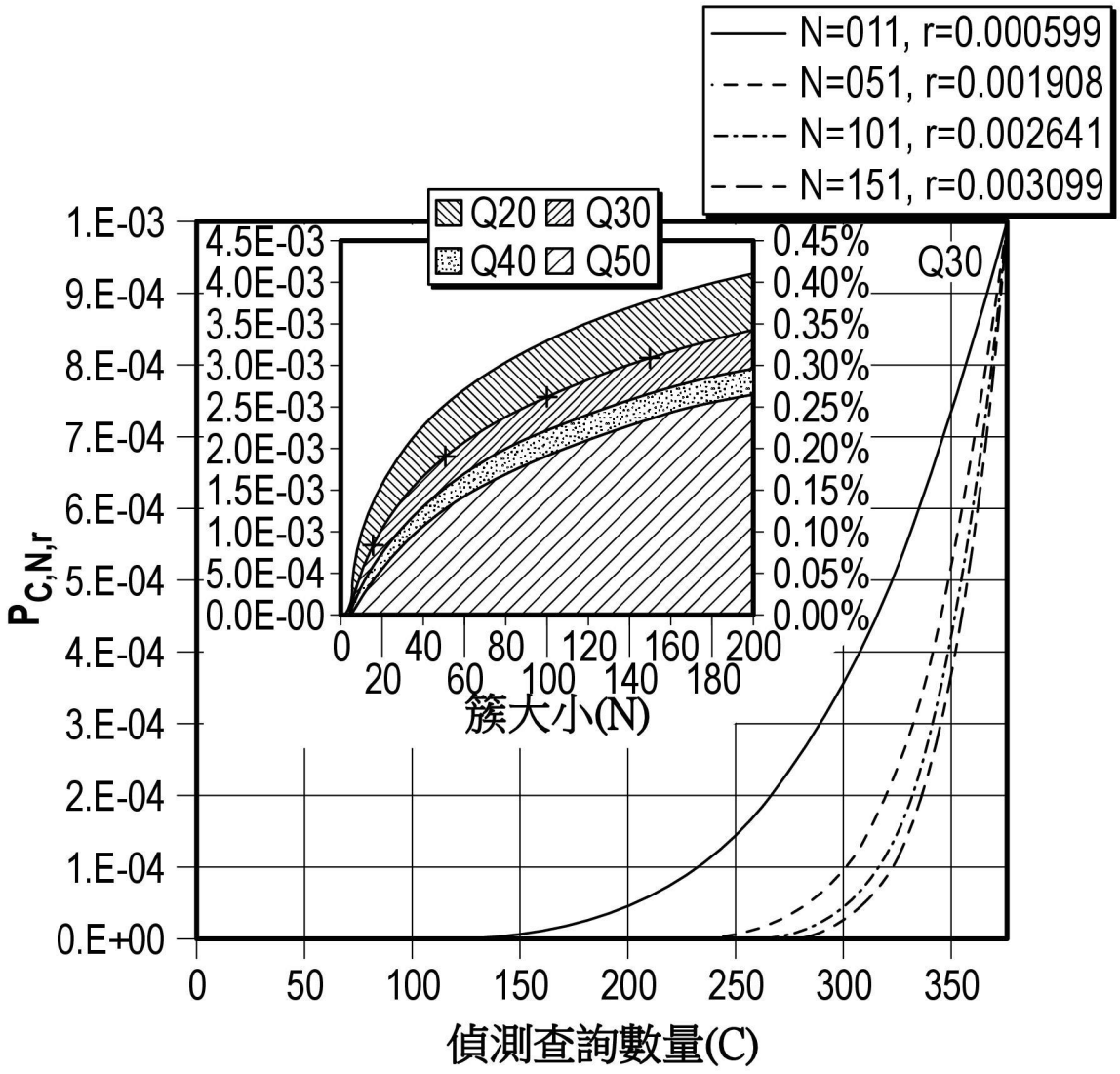
【圖38A】


0.5
0.4
0.3
0.2
0.1
0.0
$P_{C,N,r}$
r=0.1
0.05
N=101
0.01
C_{th}
0
20
40
60
80
100
120
140
160
180
200
C_{th} C_{th}
偵測查詢數量(C)

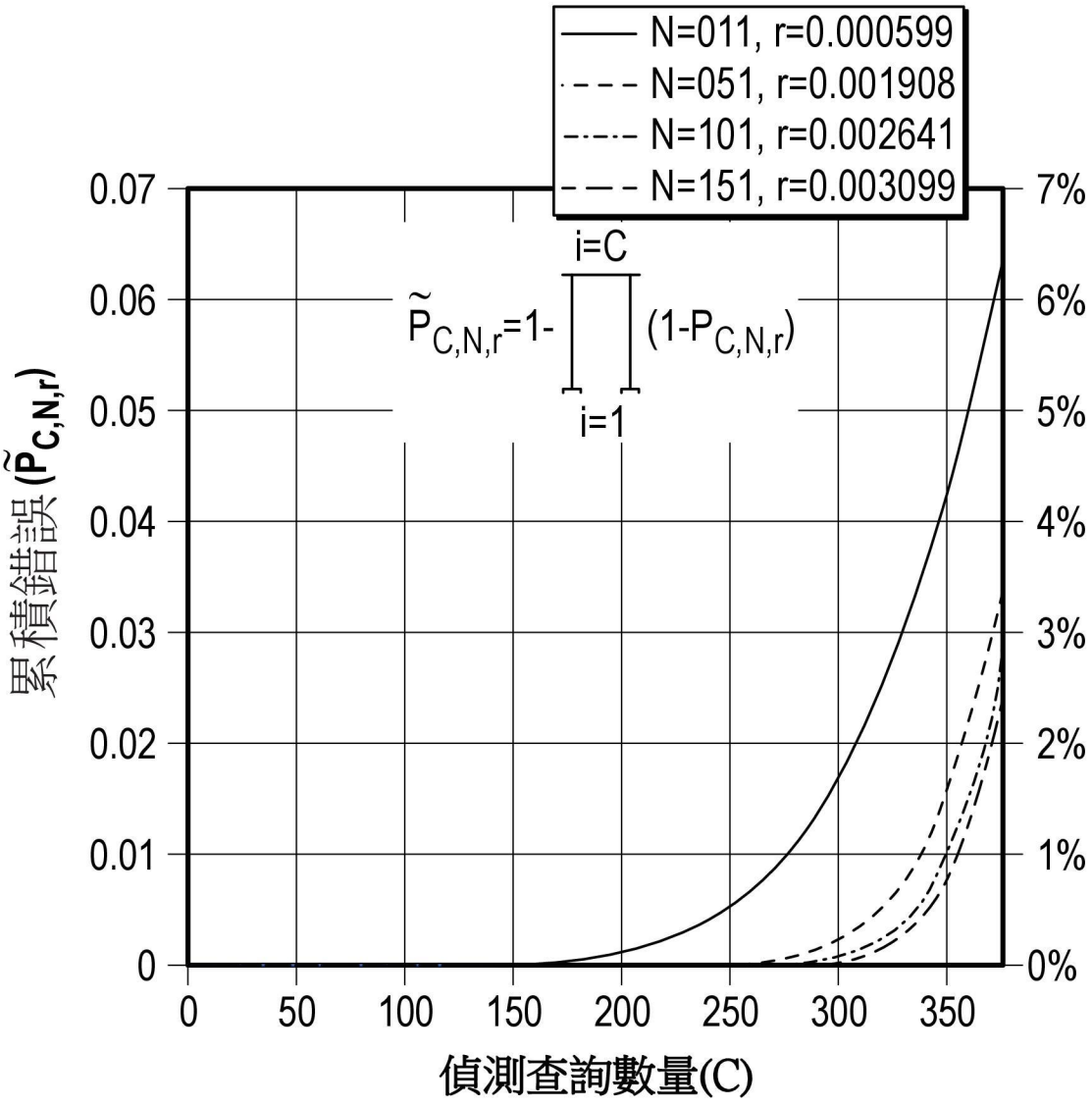
【圖38B】


Q20
Q30
Q40
Q50
N
4.5E-03
4.0E-03
3.5E-03
3.0E-03
2.5E-03
2.0E-03
1.5E-03
1.0E-03
5.0E-04
0.0E-00
0.45%
0.40%
0.35%
0.30%
0.25%
0.20%
0.15%
0.10%
0.05%
0.00%
0
20
40
60
80
100
120
140
160
180
200
化學失敗率(r)
簇大小(N)

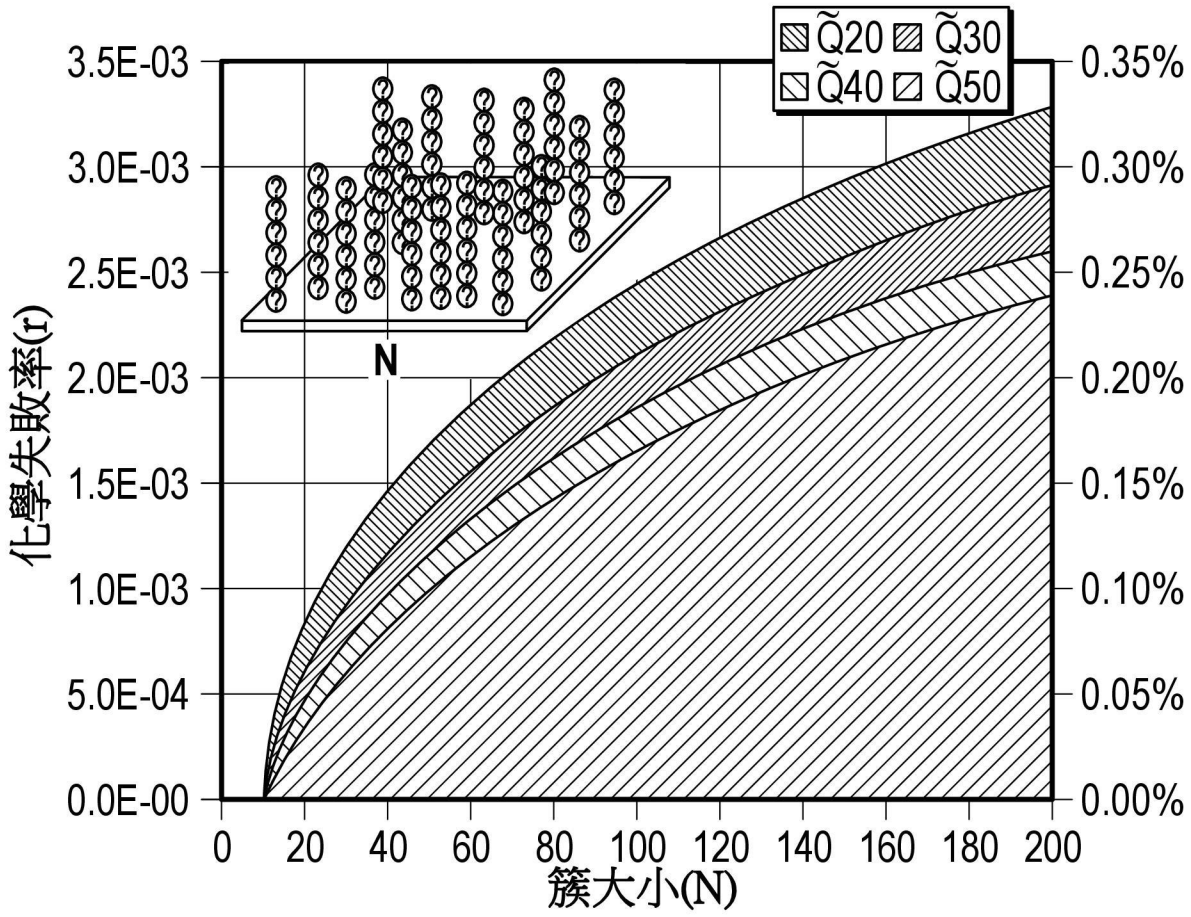
【圖39】


N=011, r=0.000599
N=051, r=0.001908
N=101, r=0.002641
N=151, r=0.003099
Q20
Q30
Q40
Q50
Q30
1.E-03
9.E-04
8.E-04
7.E-04
6.E-04
5.E-04
4.E-04
3.E-04
2.E-04
1.E-04
0.E+00
$P_{C,N,r}$
4.5E-03
4.0E-03
3.5E-03
3.0E-03
2.5E-03
2.0E-03
1.5E-03
1.0E-03
5.0E-04
0.0E-00
0.45%
0.40%
0.35%
0.30%
0.25%
0.20%
0.15%
0.10%
0.05%
0.00%
0 20 40 60 80 100 120 140 160 180 200
簇大小(N)
0 50 100 150 200 250 300 350
偵測查詢數量(C)

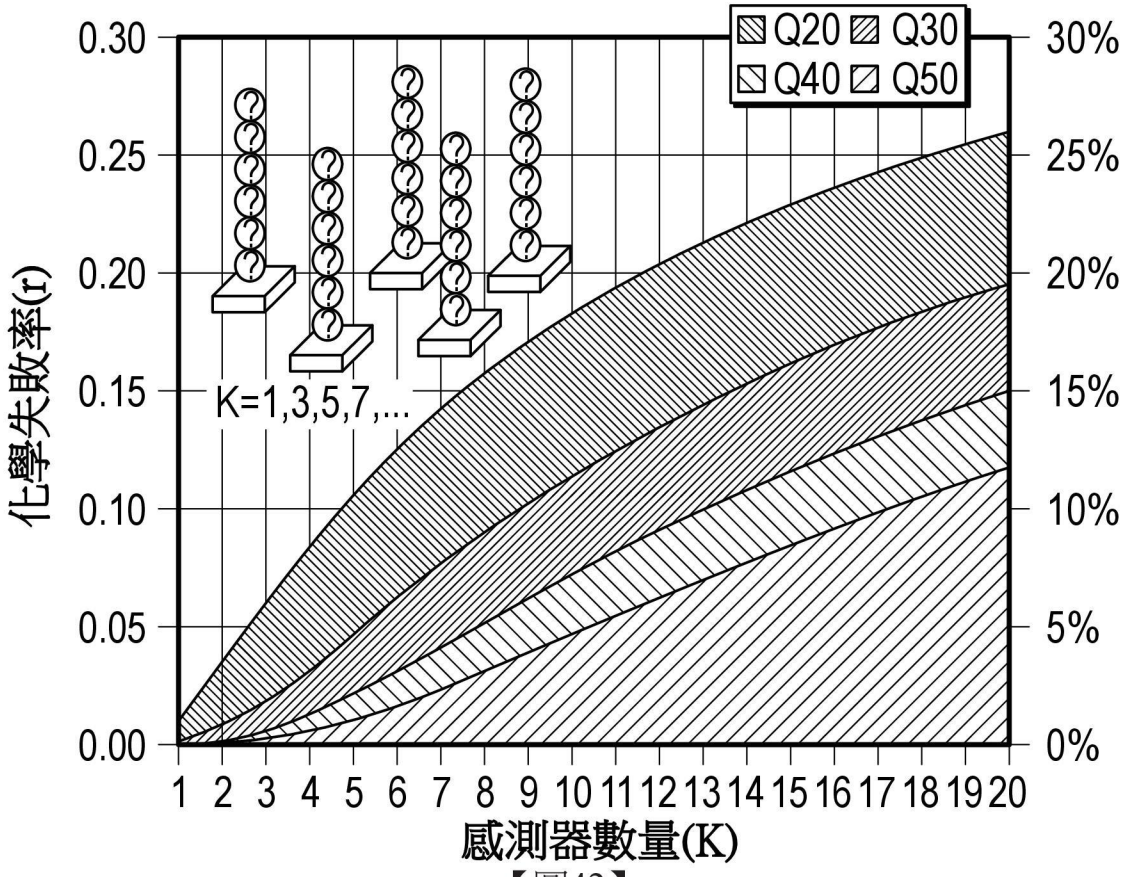
【圖40A】


N=011, r=0.000599
N=051, r=0.001908
N=101, r=0.002641
N=151, r=0.003099
i=C
$\tilde{P}_{C,N,r}=1-\prod_{i=1}(1-P_{C,N,r})$
i=1
累積錯誤($\tilde{P}_{C,N,r}$)
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
7%
6%
5%
4%
3%
2%
1%
0%
0
50
100
150
200
250
300
350
偵測查詢數量(C)

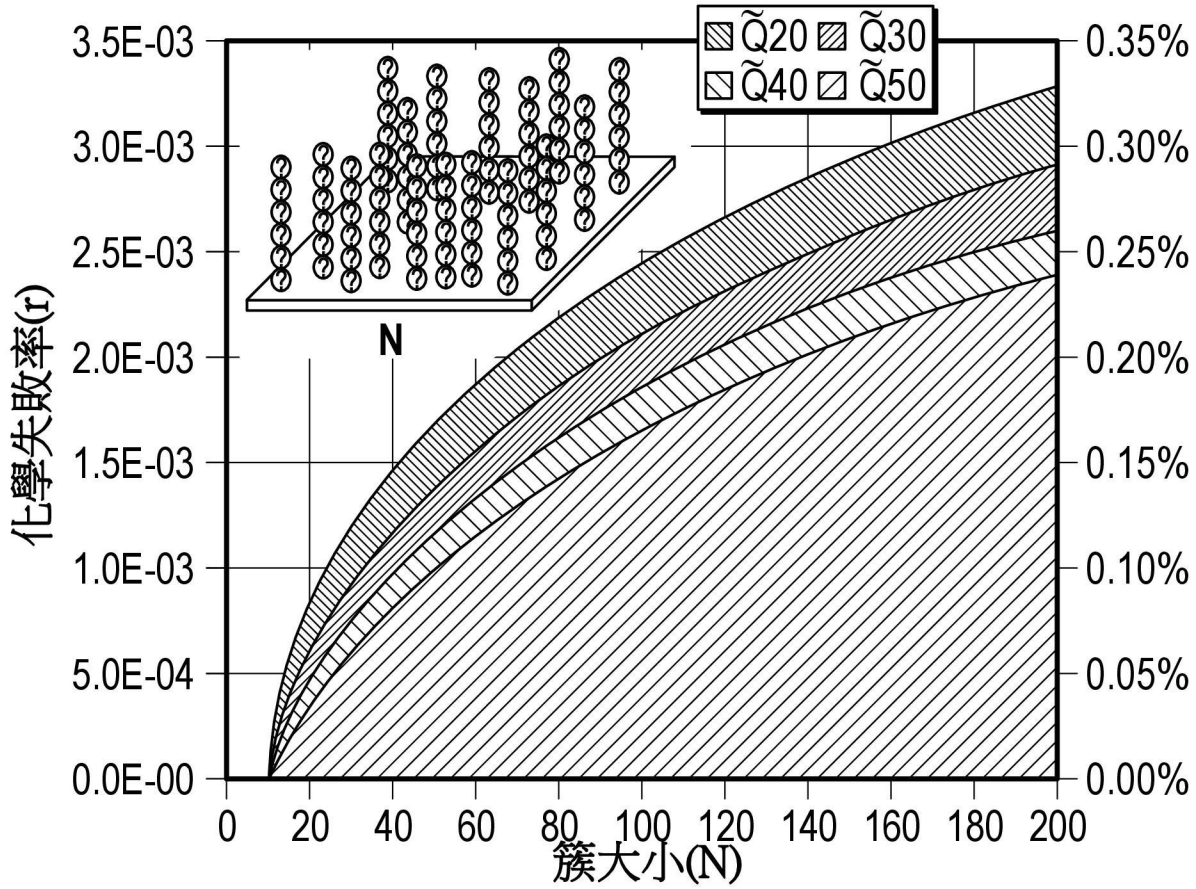
【圖40B】


Q̃20 Q̃30
Q̃40 Q̃50
3.5E-03
3.0E-03
2.5E-03
2.0E-03
1.5E-03
1.0E-03
5.0E-04
0.0E-00
0.35%
0.30%
0.25%
0.20%
0.15%
0.10%
0.05%
0.00%
N
0 20 40 60 80 100 120 140 160 180 200
化學失敗率(r)
簇大小(N)

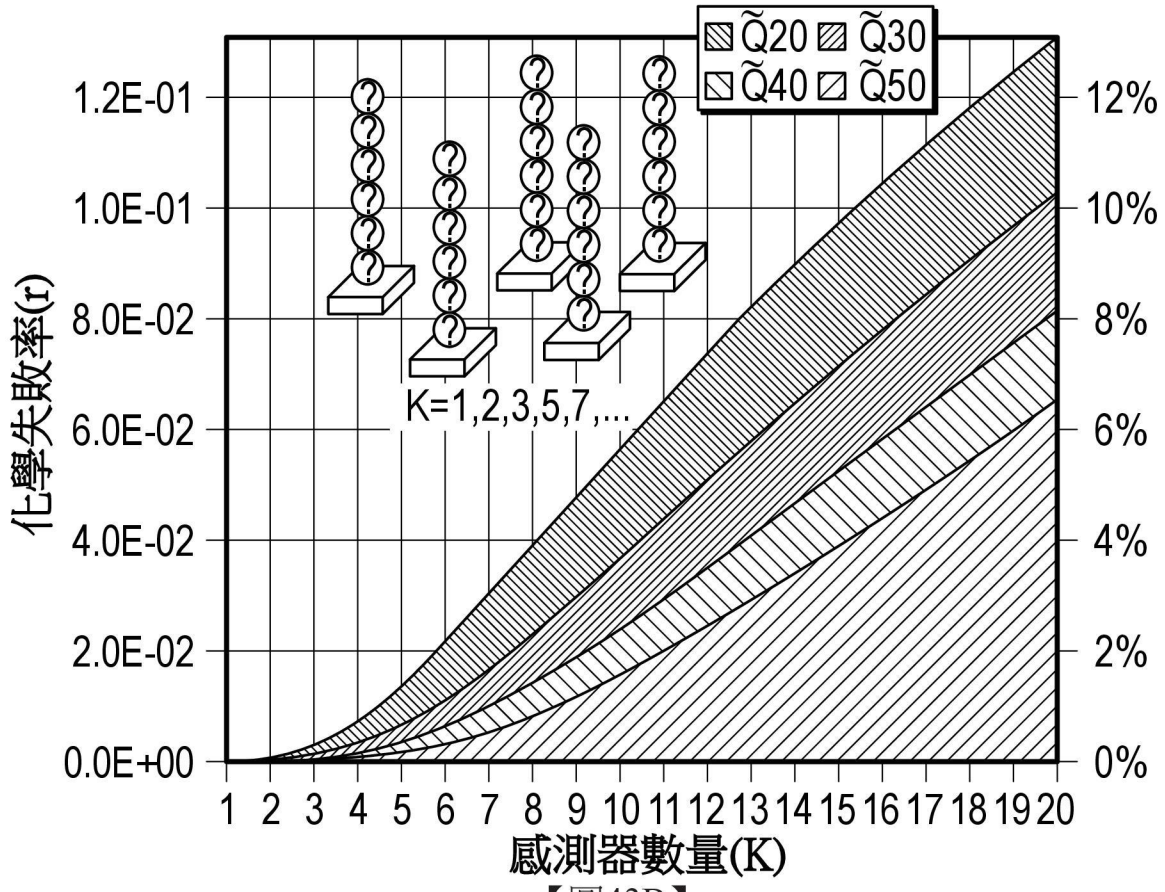
【圖41】


Q20 Q30
Q40 Q50
0.30
0.25
0.20
0.15
0.10
0.05
0.00
30%
25%
20%
15%
10%
5%
0%
K=1,3,5,7,...
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
化學失敗率(r)
感測器數量(K)

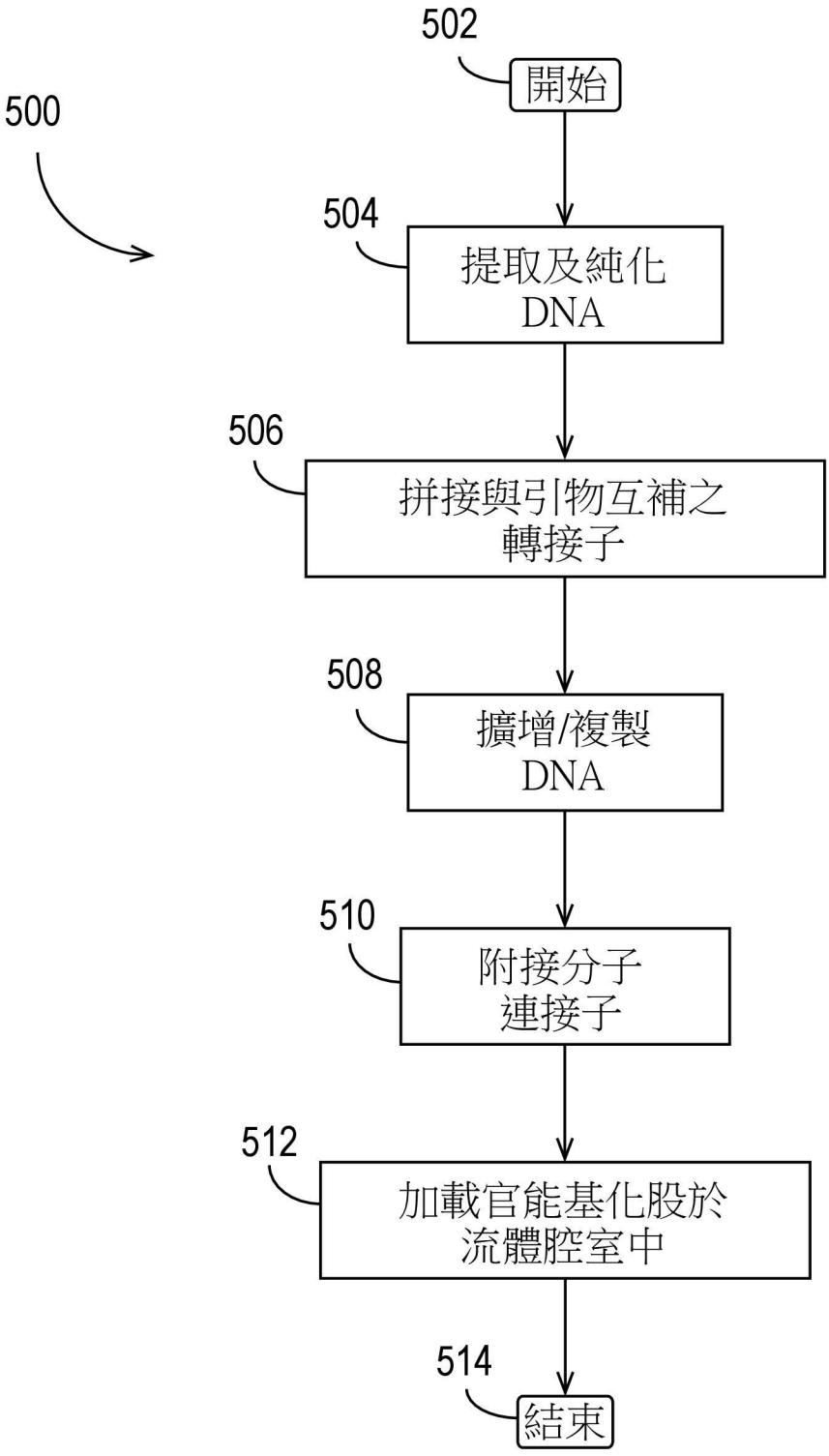
【圖42】


Q̃20 Q̃30
Q̃40 Q̃50
N
化學失敗率(r)
3.5E-03
3.0E-03
2.5E-03
2.0E-03
1.5E-03
1.0E-03
5.0E-04
0.0E-00
0.35%
0.30%
0.25%
0.20%
0.15%
0.10%
0.05%
0.00%
0 20 40 60 80 100 120 140 160 180 200
簇大小(N)

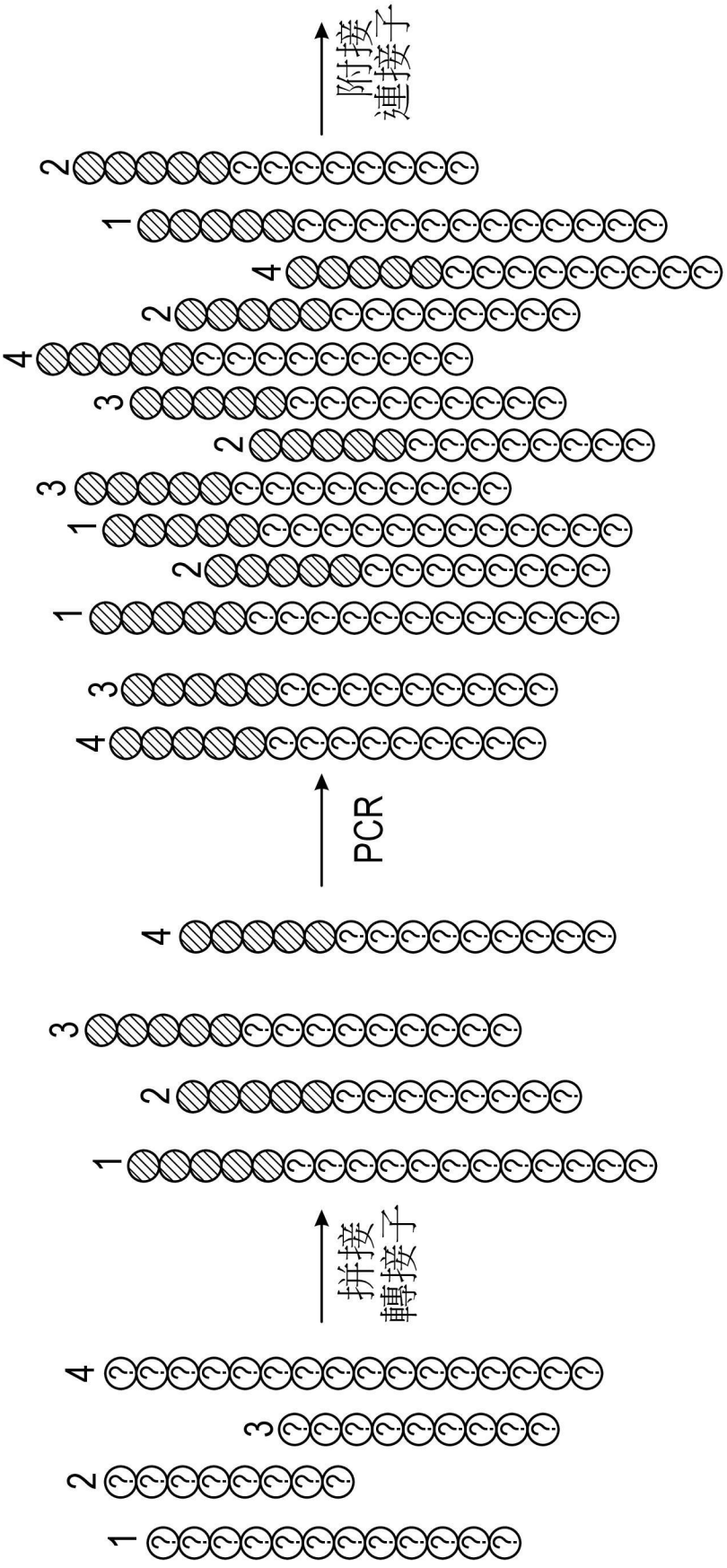
【圖43A】


Q̃20 Q̃30
Q̃40 Q̃50
K=1,2,3,5,7,...
化學失敗率(r)
1.2E-01
1.0E-01
8.0E-02
6.0E-02
4.0E-02
2.0E-02
0.0E+00
12%
10%
8%
6%
4%
2%
0%
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
感測器數量(K)

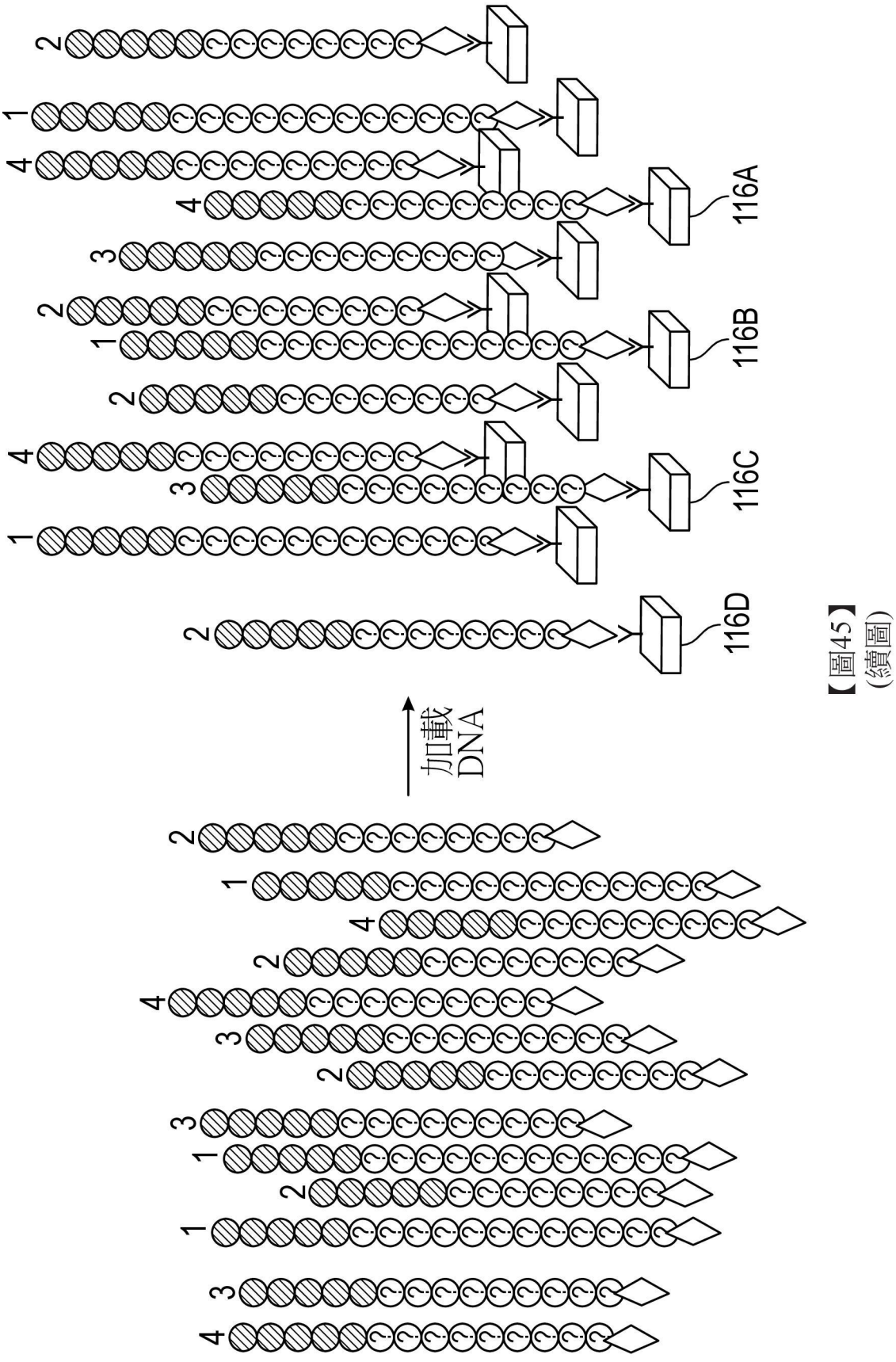
【圖43B】


500
502
開始
504
提取及純化
DNA
506
拼接與引物互補之
轉接子
508
擴增/複製
DNA
510
附接分子
連接子
512
加載官能基化股於
流體腔室中
514
結束

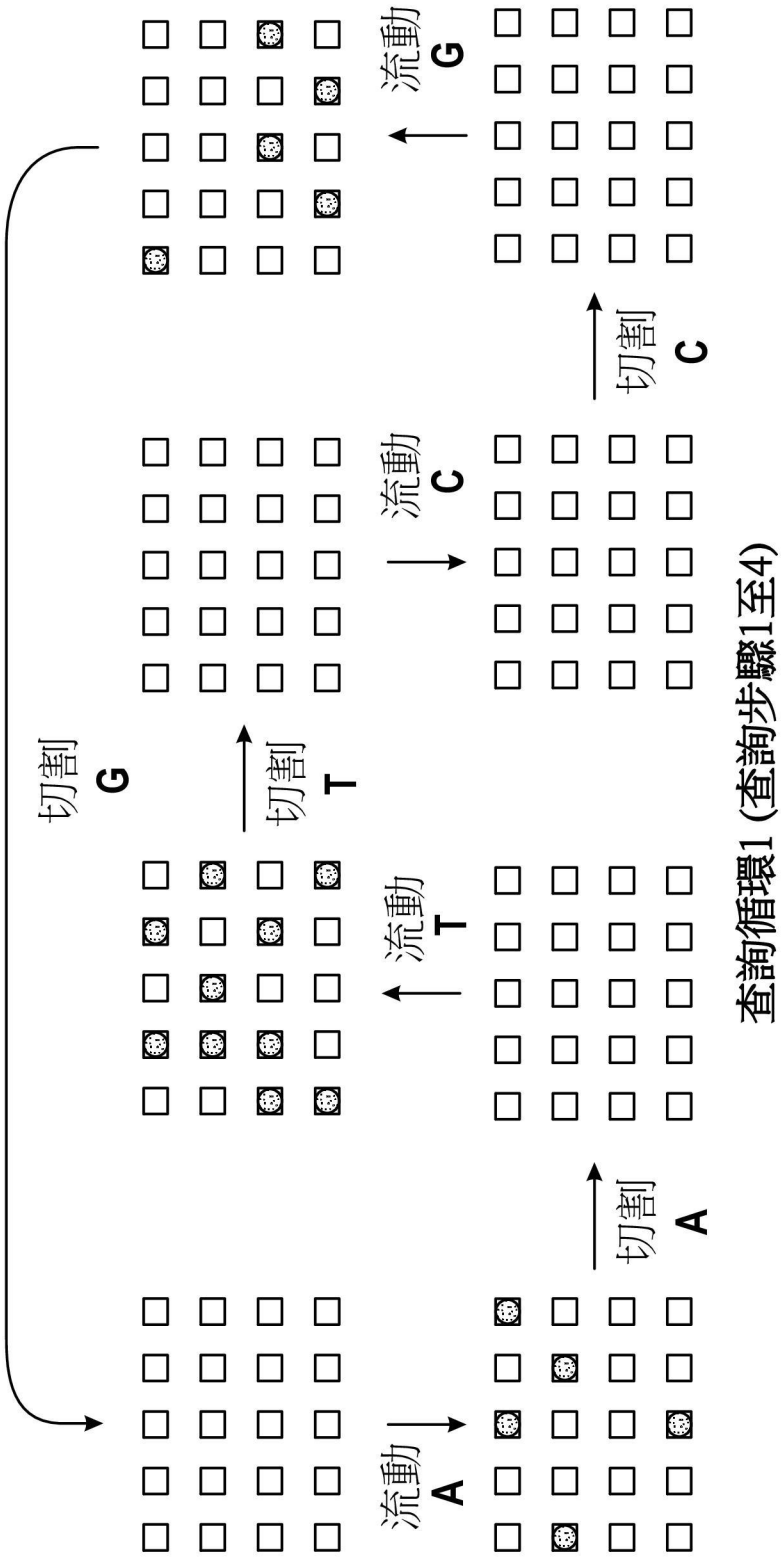
【圖44】



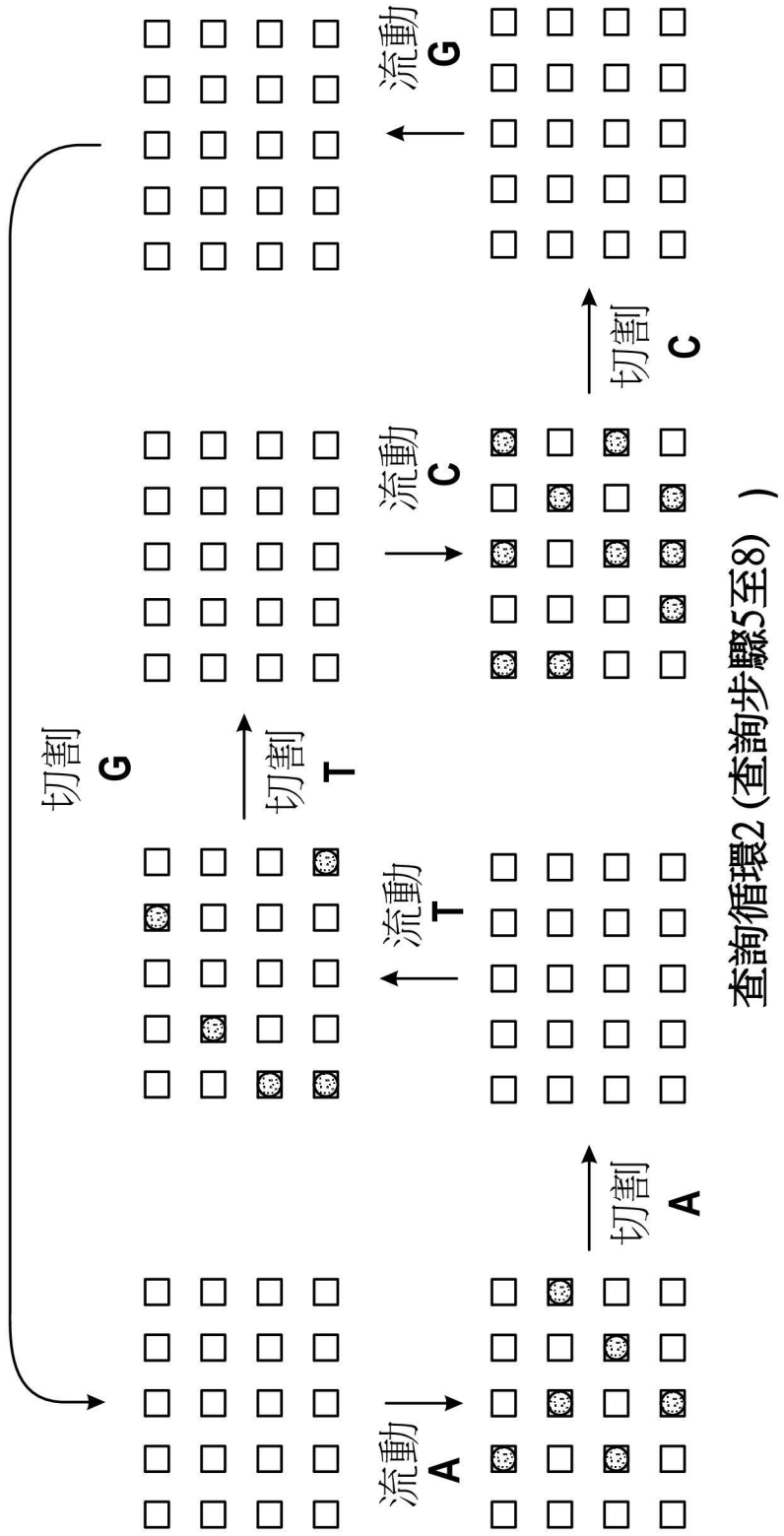
【圖45】



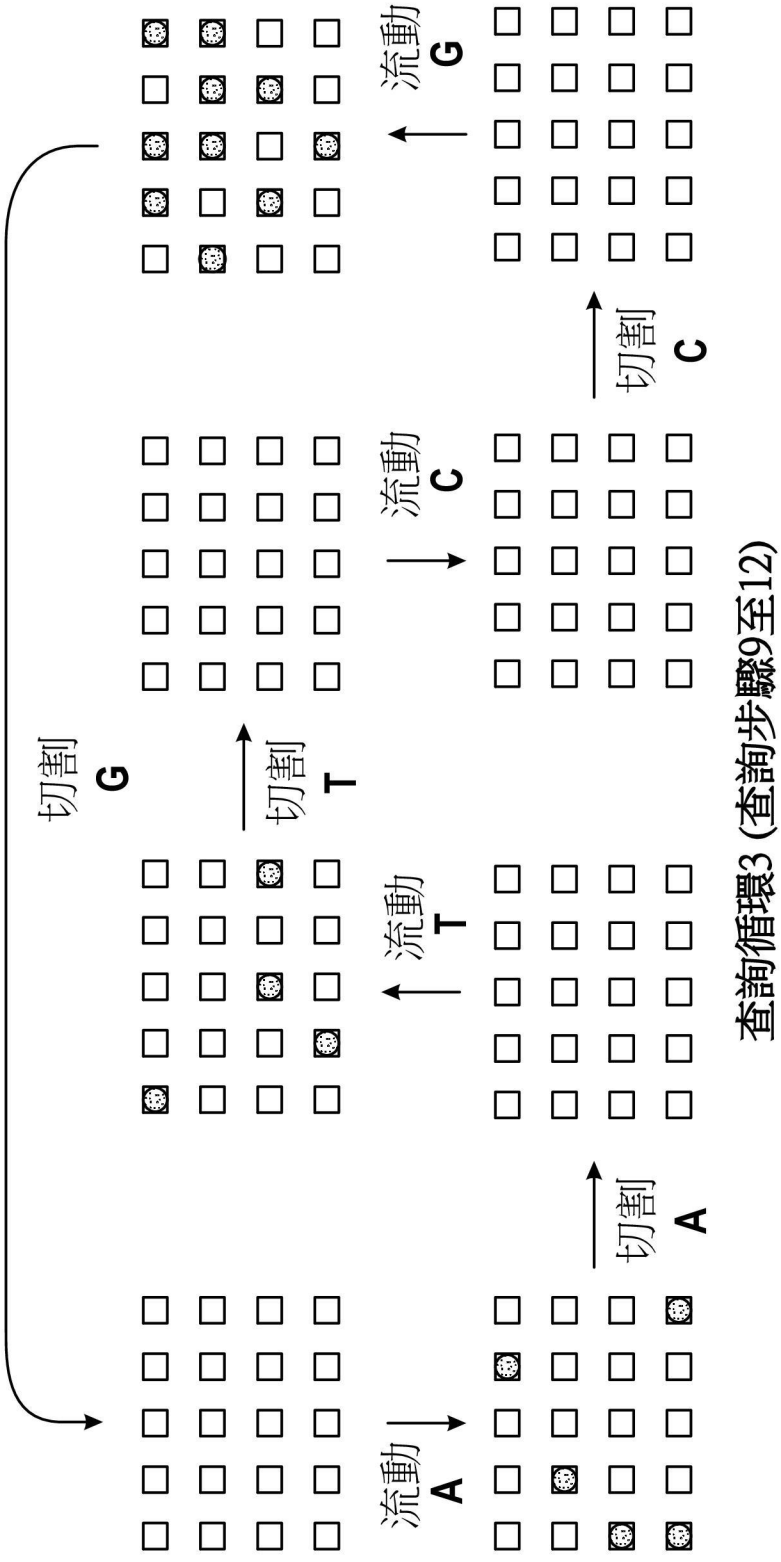
【圖45】
(續圖)



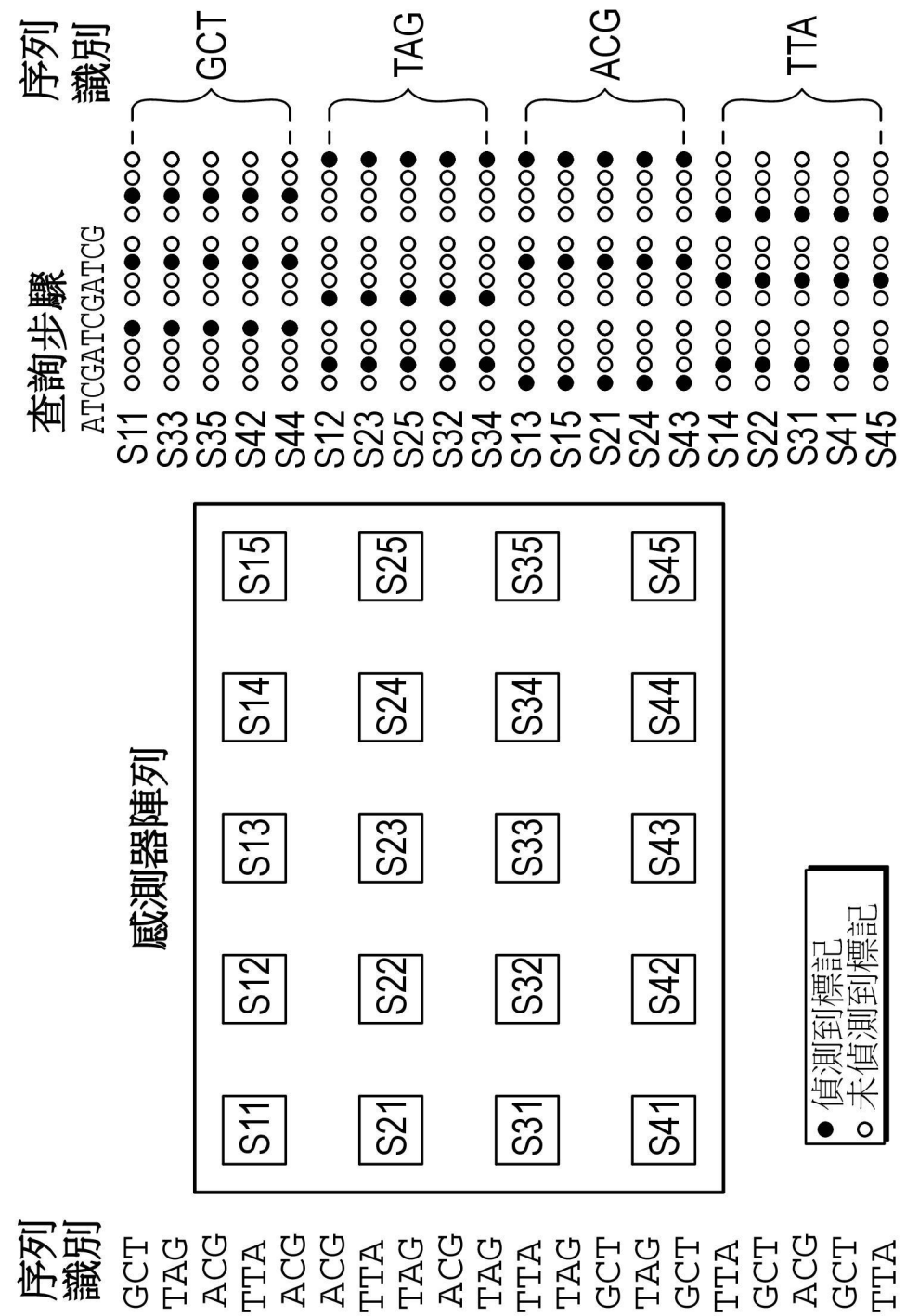
【圖46A】



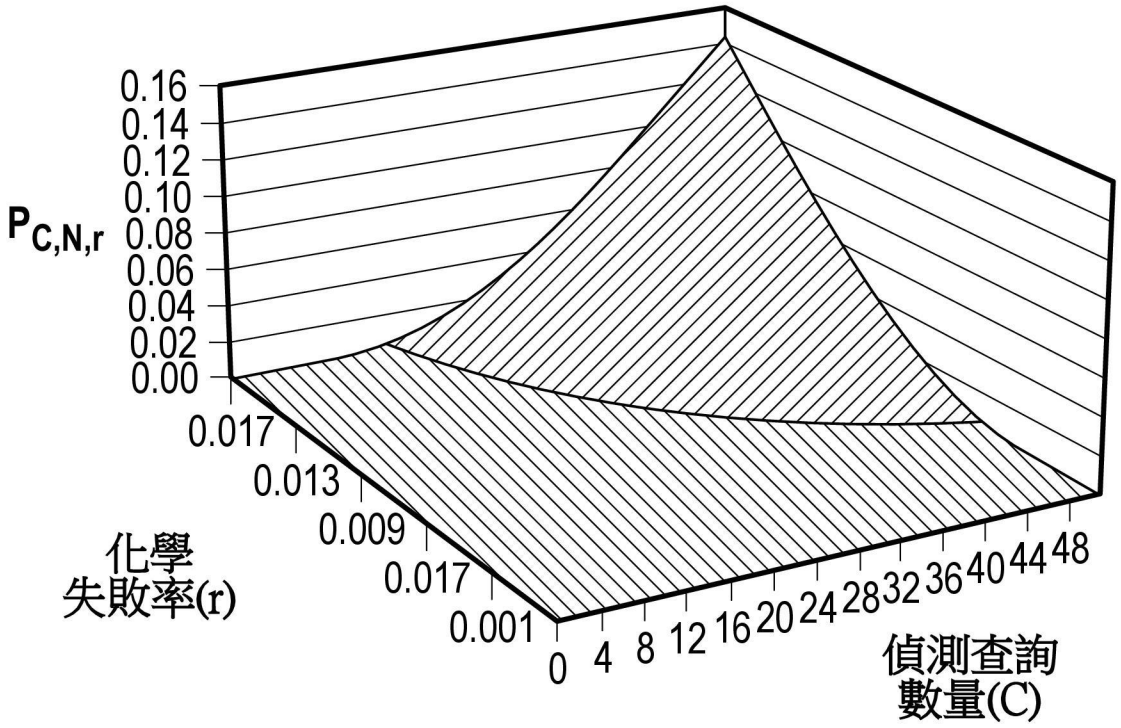
【圖46B】



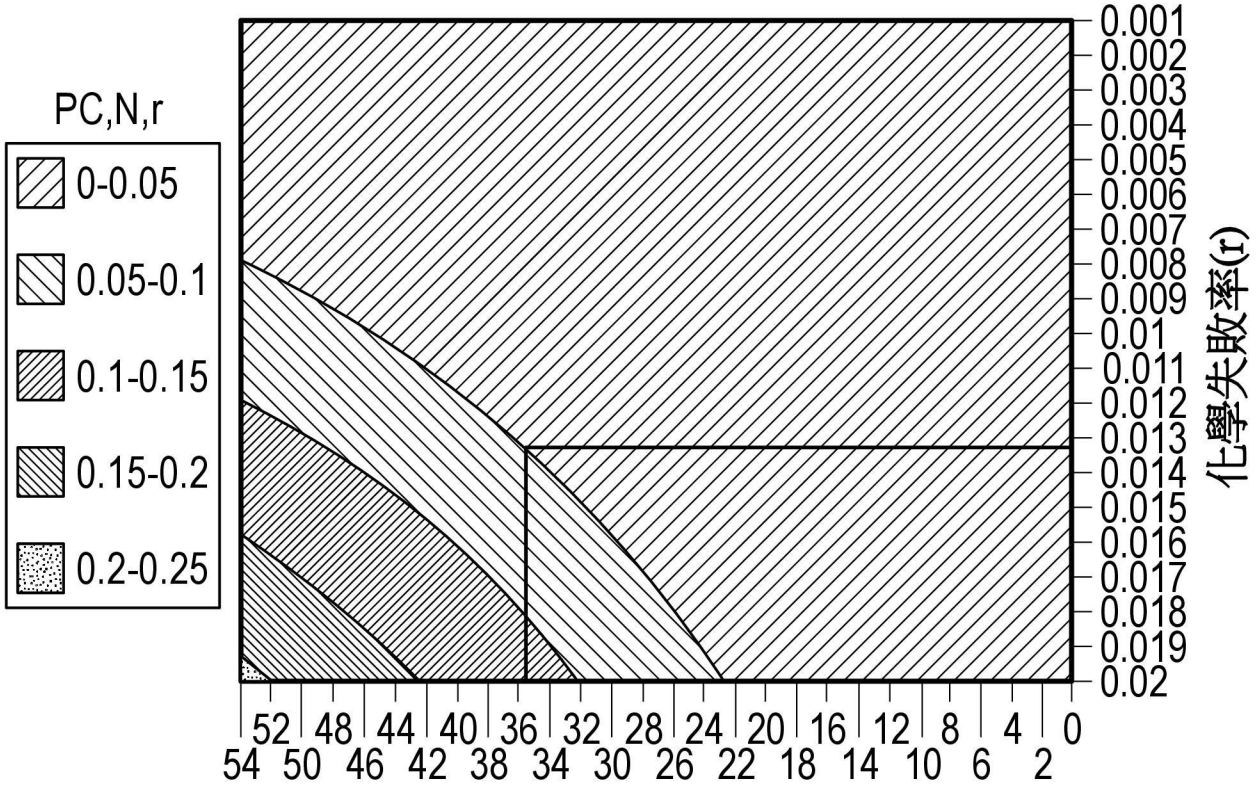
【圖46C】


查詢步驟
ATCG ATCG ATCG
序列識別
S11 ooo● oo●o o●oo GCT
S12 o●oo ●ooo oooo TAG
S13 ●ooo oo●o oooo ACG
S14 o●oo o●oo ●ooo TTA
S15 ●ooo oo●o oooo ACG
S21 ●ooo oo●o oooo ACG
S22 o●oo o●oo ●ooo TTA
S23 o●oo ●ooo oooo TAG
S24 ●ooo oo●o oooo ACG
S25 o●oo ●ooo oooo TAG
S31 o●oo o●oo ●ooo TTA
S32 o●oo ●ooo oooo TAG
S33 ooo● oo●o o●oo GCT
S34 o●oo ●ooo oooo TAG
S35 ooo● oo●o o●oo GCT
S41 o●oo o●oo ●ooo TTA
S42 ooo● oo●o o●oo GCT
S43 ●ooo oo●o oooo ACG
S44 ooo● oo●o o●oo GCT
S45 o●oo o●oo ●ooo TTA
感測器陣列
S11 S12 S13 S14 S15
S21 S22 S23 S24 S25
S31 S32 S33 S34 S35
S41 S42 S43 S44 S45
● 偵測到標記
o 未偵測到標記
查詢步驟
ATCGATCGATCG
序列識別
S11 ooo● oo●o o●oo –
S33 ooo● oo●o o●oo
S35 ooo● oo●o o●oo GCT
S42 ooo● oo●o o●oo
S44 ooo● oo●o o●oo –
S12 o●oo ●ooo oooo –
S23 o●oo ●ooo oooo
S25 o●oo ●ooo oooo TAG
S32 o●oo ●ooo oooo
S34 o●oo ●ooo oooo –
S13 ●ooo oo●o oooo –
S15 ●ooo oo●o oooo
S21 ●ooo oo●o oooo ACG
S24 ●ooo oo●o oooo
S43 ●ooo oo●o oooo –
S14 o●oo o●oo ●ooo –
S22 o●oo o●oo ●ooo
S31 o●oo o●oo ●ooo TTA
S41 o●oo o●oo ●ooo
S45 o●oo o●oo ●ooo –

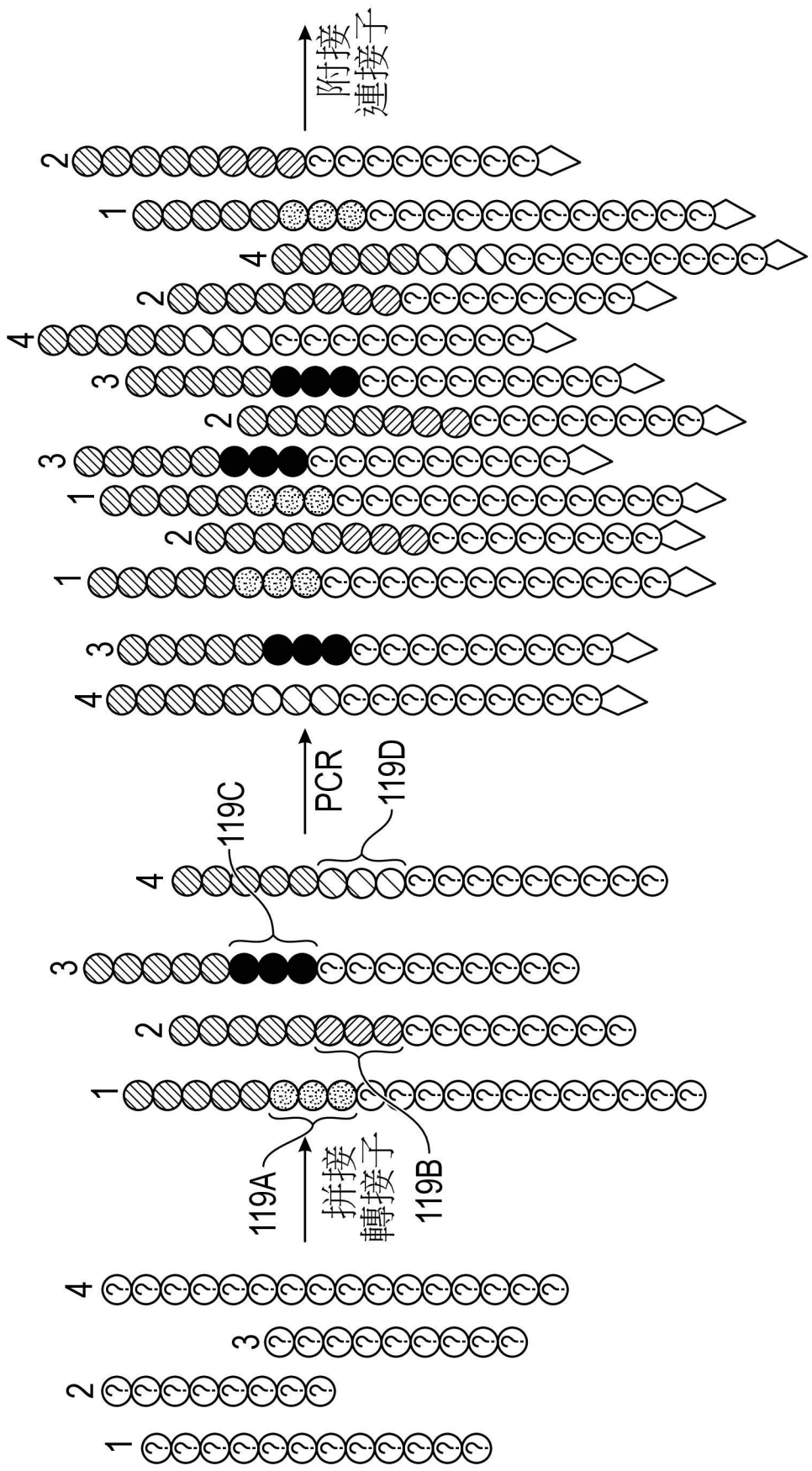
【圖47】


0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
$P_{C,N,r}$
0.017
0.013
0.009
0.017
0.001
化學
失敗率(r)
0 4 8 12 16 20 24 28 32 36 40 44 48
偵測查詢
數量(C)

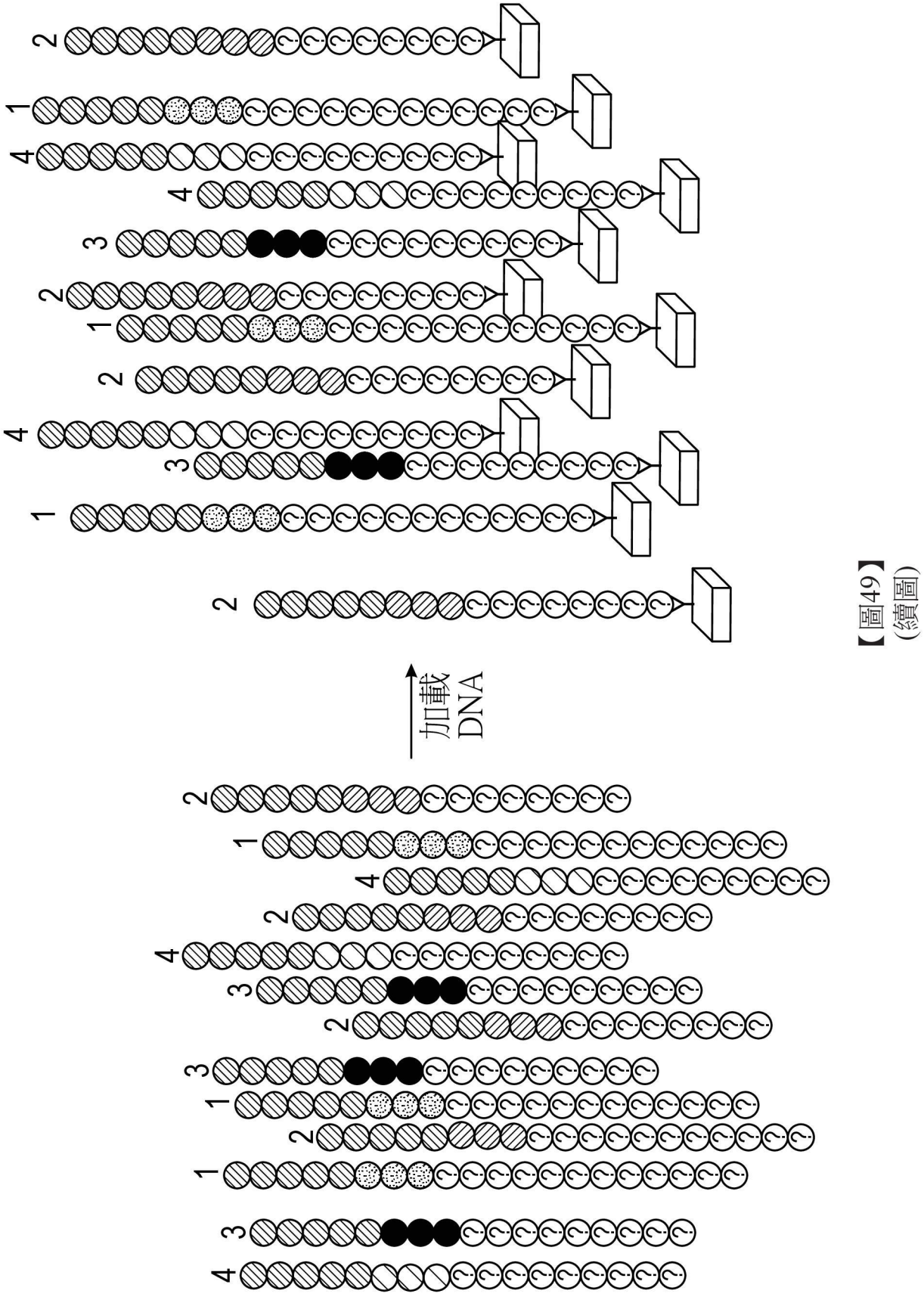
【圖48A】


PC,N,r
0-0.05
0.05-0.1
0.1-0.15
0.15-0.2
0.2-0.25
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0.008
0.009
0.01
0.011
0.012
0.013
0.014
0.015
0.016
0.017
0.018
0.019
0.02
化學失敗率(r)
52 48 44 40 36 32 28 24 20 16 12 8 4 0
54 50 46 42 38 34 30 26 22 18 14 10 6 2
偵測查詢數量(C)

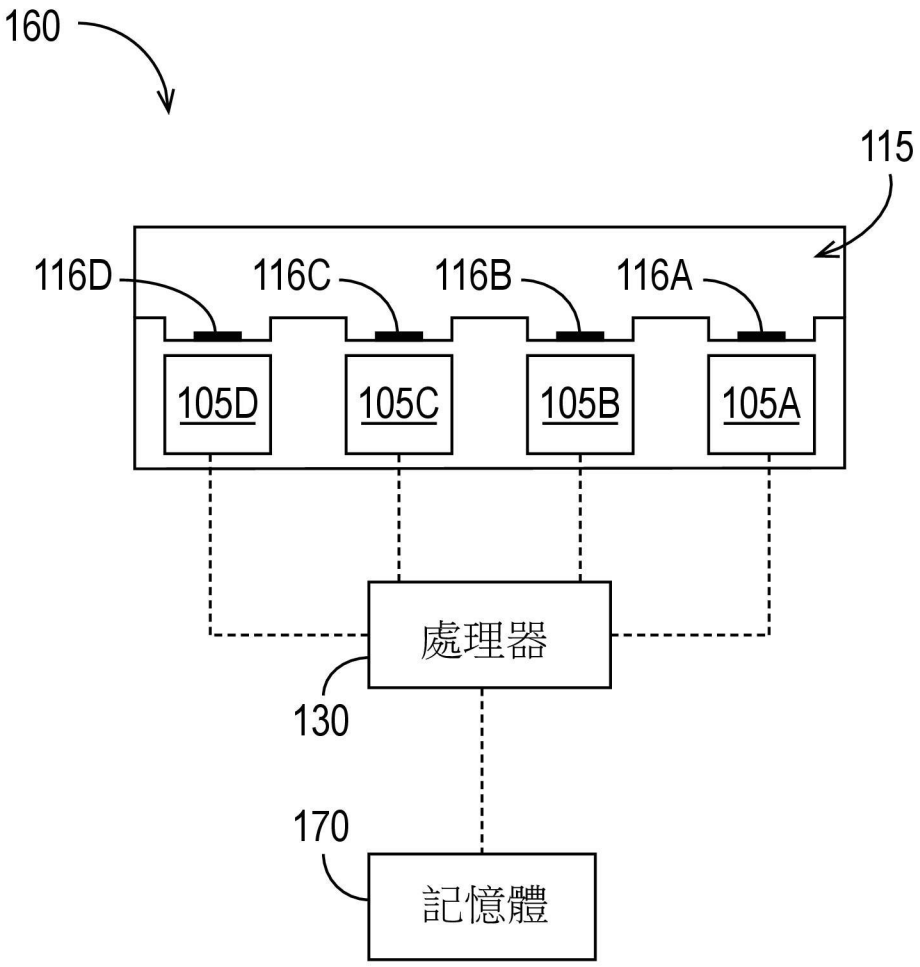
【圖48B】



【圖49】



【圖49】
(續圖)


160
115
116D
116C
116B
116A
105D
105C
105B
105A
處理器
130
170
記憶體

【圖50】