



(21) 申請案號：108106896 (22) 申請日：中華民國 108 (2019) 年 02 月 27 日

(51) Int. Cl. : *C12Q1/68 (2018.01)* *G16B20/00 (2019.01)*
G01N33/53 (2006.01) *A61K48/00 (2006.01)*

(30) 優先權：2018/03/01 美國 62/637,121
 2018/03/02 美國 15/910,822
 2018/06/01 美國 15/995,835

(71) 申請人：美商瑞科循製藥公司 (美國) RECURSION PHARMACEUTICALS, INC. (US)
 美國

(72) 發明人：維多 梅森 VICTORS, MASON L. (US)；伯吉森 布雷克 BORGESON, BLAKE C.
 (US)；珍布蘭克 瑟德瑞克 ST-JEAN-LEBLANC, CEDRIC (CA)

(74) 代理人：林志剛

申請實體審查：無 申請專利範圍項數：30 項 圖式數：6 共 113 頁

(54) 名稱

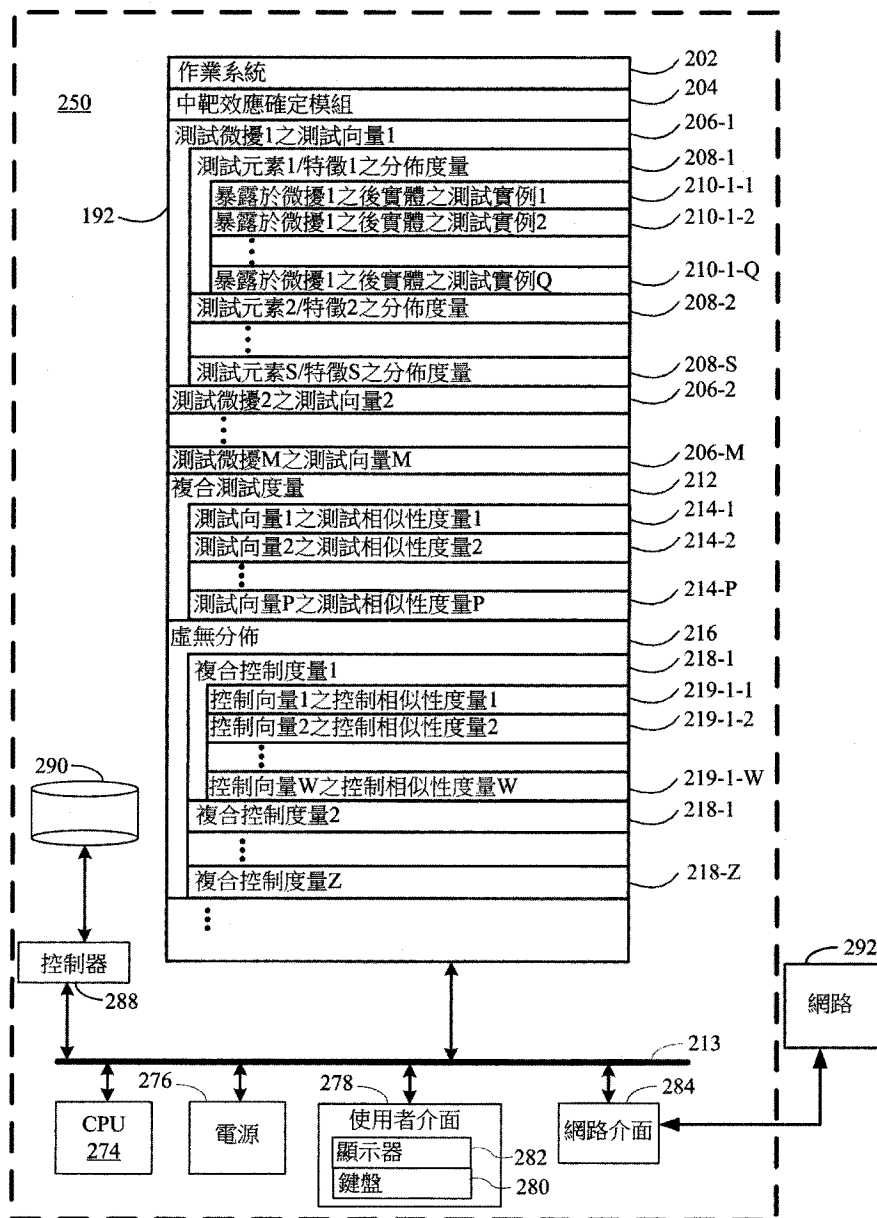
用於區別對靶標之效應的系統及方法

(57) 摘要

揭示用於確定一組測試微擾是否相對於虛無分佈區別對實體之第一成分的中靶效應的系統及方法。該等微擾為該第一成分之微擾，且該實體包含複數個成分。對於該組中之各微擾，獲得包含複數個元素的對應向量。各元素包含在暴露於該各別微擾之後該實體實例中特徵之量測值之分佈度量，或(ii)使用在微擾暴露之後該實體實例之該複數個特徵之該量測值所計算的各別維度縮減成分之分佈度量。使用該等向量計算複合度量，並與虛無分佈相比較。當該複合度量不同於該虛無分佈時，認為該組微擾相對於該虛無分佈區別對該第一成分的該中靶效應。

Systems and methods for determining whether a set of test perturbations discriminates over a null distribution for an on target effect against a first component of an entity are disclosed. The perturbations are perturbations of the first component and the entity comprises a plurality of components. For each perturbation in the set, a corresponding vector comprising a plurality of elements, is obtained. Each element comprises a distribution metric of measurements of a feature across instances of the entity upon exposure to the respective perturbation or (ii) a distribution metric of a respective dimension reduction component computed using the measurement of the plurality of features across instances of the entity upon the perturbation exposure. A composite metric is computed, using the vectors, and compared to a null distribution. When the composite metric is differentiated from the null distribution, the set of perturbations is deemed to discriminate the on target effect against the first component over the null distribution.

指定代表圖：



【圖 2】

符號簡單說明：

- 192 . . . 記憶體
- 202 . . . 作業系統
- 204 . . . 中靶效應確定模組
- 206 . . . 測試向量/第一測試向量/向量
- 208 . . . 測試元素/第一測試元素/特徵
- 210 . . . 測試實例
- 212 . . . 複合測試度量
- 213 . . . 通信匯流排
- 214 . . . 測試相似性度量
- 216 . . . 虛無分佈
- 218 . . . 複合控制度量
- 219 . . . 控制相似性度量
- 220 . . . 一組控制微擾
- 222 . . . 控制微擾
- 224 . . . 成分
- 226 . . . 控制向量
- 228 . . . 控制元素
- 230 . . . 控制實例
- 250 . . . 系統/數位影像/裝置/電腦系統
- 274 . . . 處理單元/中央處理單元/處理器
- 276 . . . 電源
- 278 . . . 使用者介面
- 280 . . . 輸入
- 282 . . . 顯示器
- 284 . . . 網路或其他通信介面
- 288 . . . 控制器

290 . . . 磁碟存儲
及/或持續裝置/記憶體
292 . . . 元件

【發明說明書】

【中文發明名稱】

用於區別對靶標之效應的系統及方法

【英文發明名稱】

SYSTEMS AND METHODS FOR DISCRIMINATING EFFECTS ON
TARGETS

【技術領域】

【0001】本揭露大致上係關於用於識別出具有對選定靶標的中靶(on target)效應的一組微擾的系統及方法。

【先前技術】

【0002】許多疾病呈現出特定的結構表型，與狄蘭氏症候群(Cornelia de Lange Syndrome)患者之明顯面部特徵一樣。類似地，許多疾病之特定結構標識可經由細胞影像獲得。因此，自動化顯微成像為研究此等特定細胞結構標識的高通量篩選平台提供合適的基礎。顯微成像提供關於功能資料點的資訊以及在x、y、及z維度上的相關空間資訊。

【0003】過去幾年中，顯微成像技術已有所進步。光學、機器人科學、及計算技術之進步，以及擴大的對比標記庫包括功能活細胞報導子均有助於廣泛採用在基於細胞之檢定系統中提供高度動態且量化的螢光讀出的基於影像

之篩選平台。參見 Bickle, 2010, 「The beautiful cell: high-content screening in drug discovery」, *Anal. Bioanal Chem.* 398, 219-226 ; Isherwood 等人, 2011, 「Live cell in vitro and in vivo imaging applications: accelerating drug discovery」, *Pharmaceutics* 3, 141-170 (2011) ; 及 Kummel 等人, 2010, 「Integration of multiple readouts into the z' factor for assay quality assessment」, *J. Biomol. Screen* 15, 95-101。

【0004】最近已出現了非侵入式無標籤成像技術，滿足在高含量篩選情況下針對基於細胞之檢定的最小細胞操作 (manipulation) 之需求。在此等無標籤技術中，數位全像顯微術 (Rappaz 等人, 2015 Automated multi-parameter measurement of cardiomyocytes dynamics with digital holographic microscopy」, *Opt. Express* 23, 13333-13347) 提供使用 96 孔及 384 孔盤的針對端點及曠時成像而自動化的量化資訊。參見例如 Kuhn, J. 等人, 2013 「Label-free cytotoxicity screening assay by digital holographic microscopy」, *Assay Drug Dev. Technol.* 11, 101-107; Rappaz 等人, 2014 「Digital holographic microscopy: a quantitative label-free microscopy technique for phenotypic screening」, *Comb. Chem. High Throughput Screen* 17, 80-88 ; 及 Rappaz 等人, 2015, *Label-Free Biosensor Methods in Drug Discovery* (Fang, Y. 編) 307-325, Springer Science+Business Media)。類似地，可以數位方式重建並

量化無標籤光學技術諸如相位差或微分干涉差(DIC)。參見Koos, 2015, 「DIC image reconstruction using an energy minimization framework to visualize optical path length distribution」, *Sci. Rep.* 6, 30420。層光螢光顯微術(light sheet fluorescence microscopy; LSFM)有望以3D高解析度且在快速的記錄速度及最小的光誘導細胞損傷之情況下分析大量樣本。LSFM在各種研究領域中日漸流行, 包括神經科學、植物與發育生物學(plant and developmental biology)、毒物學、及藥物發現, 但是其尚不適合於自動化HTS設定。參見Pampaloni等人, 2014, 「Tissue-culture light sheet fluorescence microscopy (TC-LSFM) allows long-term imaging of three-dimensional cell cultures under controlled conditions」, *Integr. Biol. (Camb.)* 6, 988-998; Swoger等人, 2014, 「Imaging cellular spheroids with a single (selective) plane illumination microscope」, *Cold Spring Harb. Protoc.*, 106-113; 及Pampaloni等人, 2013, 「High-resolution deep imaging of live cellular spheroids with light-sheet-based fluorescence microscopy」, *Cell Tissue Res.* 352, 161-177。

【0005】細胞繪畫(cell painting)及細胞繪畫之相關變型表示有希望的成像技術之另一形式。細胞繪畫為一種形態剖析檢定, 其複用六種螢光染料(在五個通道中成像)來顯示八種廣泛相關的細胞成分(cellular component)或胞器。將細胞鋪板於多孔盤中, 用欲測試之處理進行擾動,

染色，固定並在高通量顯微鏡上成像。接著，自動化影像分析軟體識別出單獨的細胞並量測在一個與成千上萬個之間的任何數目(但最常見為約1,000個)的形態特徵(各種全細胞及亞細胞成分之大小、形狀、紋理、強度等的各種量度)，以產生合適於偵測甚至細微表型的輪廓。可比較用不同實驗微擾處理之細胞群之輪廓以合乎許多目標，諸如識別出化學或遺傳性微擾之表型影響、將化合物及/或基因分組至功能路徑、及識別出疾病標識。參見 Bray 等人, 2016, *Nature Protocols* 11, 1757-1774。

【0006】顯微成像允許在其中細胞係用微擾(諸如經設計以破壞細胞內的單一基因同時使其他基因之破壞最小的 siRNA)進行微擾的高通量篩選，且顯微成像係用於量化此類微擾之效應。實際上，此類篩選可用於首先識別出使細胞具有所關注之疾病之特徵的微擾，且進一步用於確定哪些化合物援救藉助於該微擾所誘導之疾病相關特徵。用於此類高通量篩選工作的已知成像技術之缺點為其昂貴、低效、及/或缺乏足夠的通量潛力。

【0007】已知篩選技術(包括細胞繪畫)中出現之缺點為，脫靶效應之量值及盛行率使靶向相同基因的微擾之形態輪廓比靶向不同基因的微擾之形態輪廓看起來更不相似。參見 Singh 等人, 2015, 「Morphological profiles of RNAi-induced gene knockdown are highly reproducible but dominated by seed effects」, *PLoS One* 10, e0131370。在其他多參數檢定中已觀察到此現象，且其並不是使用成像技

術之形態剖析所特有的。此效應阻礙使用諸如短 RNAi 試劑之微擾、或脫靶效應顯著的任何其他微擾的大規模實驗，在該等實驗中實驗設計需要廣泛比較所有樣本。

【0008】 鑒於以上背景，此項技術中需要用於補償脫靶效應以識別出具有對選定靶標之中靶效應的一組微擾的系統及方法。

【發明內容】

【0009】 本揭露連同附屬一組微擾解決此項技術中對允許評估並表徵不完全微擾(諸如 siRNA)之中靶效應的系統及方法的需要。本揭露解決針對試劑之特異性不足而需要此一介入來確定中靶效應的任何疾病的藥物發現之技術。就遺傳性疾病而言，使用若干微擾(諸如 siRNA)來建立人類疾病狀態模型。每一基因使用若干不同的微擾。類似地，可使用來自多個製造商的若干分泌因子或蛋白(諸如細胞介素)以識別出蛋白在生物學方面的中靶本質，而不是與雜訊之獨特產生或其他未識別原因相關的副效應。第一目標為識別出破壞造成所研究疾病的基因之功能的微擾，其被稱為「中靶效應」。不幸的是，此類微擾同樣具有脫靶效應，其中除導致所研究疾病的基因之外的基因之功能同樣經微擾。有利的是，本揭露之系統及方法解決此等脫靶效應。在本揭露中，在用針對給定基因的若干不同微擾(例如，siRNA)進行暴露之後，捕獲健康細胞之影像。在本揭露中，確定具體基因之微擾之共同表型效應，

前提為，具體基因之微擾之此類共同表型效應實際上表示與破壞所關注之基因相關聯的中靶效應。例如，在各微擾為 siRNA 之情況下，因為該複數個 siRNA 中之各 siRNA 經設計成破壞靶基因，所以由此類 siRNA 所致之共同效應表示獨立的破壞基因的中靶效應，且雜訊代表由各單獨的 siRNA 所致之其他脫靶效應。因此，在本揭露中尋求該複數個 siRNA 所共有的訊號。此外，本揭露提供篩選可逆轉 siRNA 之中靶效應的化合物的方式。

【0010】因此，在本揭露中，自暴露於微擾之細胞獲得高維度表型向量，該等微擾靶向一或多個與目標疾病相關聯之基因。各個此類向量包含複數個元素。各個此類元素表示在複數個細胞測試實例中不同特徵在複數個特徵中的量測值之分佈度量，或 (ii) 使用在複數個細胞測試實例中該複數個特徵之量測值計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量。將針對此等微擾中之各者所量測之高維度表型串彼此比較，以識別出共同訊號，即中靶訊號。根據此比較，選擇最代表藉由微擾所表現出之效應(例如，最代表自各自己暴露於靶向一或多個基因的微擾之一的細胞獲得之高維度表型串)的具體微擾(例如，具體 siRNA)作為目標微擾。目標微擾形成高通量篩選之基礎，在高通量篩選中暴露於目標微擾之細胞係用作篩選檢定之基礎，該篩選檢定試圖識別出當暴露於微擾暴露細胞時，能夠逆轉藉由此類微擾暴露細胞所表現出之中靶效應的化合物。

【0011】據此，本揭露之一個態樣提供用於確定一組測試微擾是否相對於使用控制微擾組所形成之虛無分佈區別(例如，區分)對實體(例如細胞)之第一成分(例如基因)的中靶效應的系統及方法。該組測試微擾為第一成分之微擾，且該實體包含複數個成分(例如，複數個基因)。對於該組測試微擾中之各微擾，獲得包含複數個元素的對應向量(例如，高維度表型串)。各元素包含在暴露於各別微擾之後實體實例中特徵之量測值之分佈度量，或(ii)使用微擾暴露之後實體實例中該複數個特徵之量測值所計算的各別維度縮減成分之分佈度量。使用上文所述之向量計算複合度量，並與虛無分佈相比較。當複合度量不同於虛無分佈時，認為該組微擾相對於使用數組控制微擾所形成之虛無分佈區別對第一成分的中靶效應。

【0012】更詳言之，本揭露之一個態樣提供一種用於確定一組測試微擾是否相對於使用數組控制微擾所形成之虛無分佈區別對實體之第一成分的中靶效應的電腦系統。該組測試微擾包含該第一成分之複數個測試微擾，且該實體包含包括該第一成分之複數個成分。該電腦系統包含一或多個處理器、記憶體、及一或多個程式。該一或多個程式存儲於記憶體中且經組態以由該一或多個處理器執行。

【0013】該一或多個程式包括用於針對該組測試微擾中之各個各別測試微擾獲得對應測試向量，從而獲得複數個測試向量的指令。各對應測試向量包含複數個測試元素。該複數個測試元素中之各測試元素包含：(i)在複數個

實體測試實例暴露於各別測試微擾之後在該複數個實體測試實例中不同特徵在複數個特徵中之量測值之分佈度量；或(ii)使用在複數個實體測試實例暴露於各別測試微擾之後在該複數個實體測試實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量。

【0014】 該一或多個程式進一步包括用於使用該複數個測試向量計算複合測試度量的指令。複合測試度量量化該組測試微擾對第一成分之中靶效應。測試度量係藉由第一過程計算。在第一過程中，對於該複數個測試向量中之各個各別測試向量，計算(i)各別測試向量與(ii)在自該複數個測試向量中除去該各別測試向量之情況下，該複數個測試向量之分佈度量之間的測試相似性度量，從而獲得該複數個測試向量之複數個測試相似性度量。該複數個測試相似性度量中之各測試相似性度量獨特地對應於第一組微擾中之微擾。計算複合測試度量，作為該複數個測試相似性度量之集中趨勢量度。計算虛無分佈。虛無分佈包含複數個複合控制度量，該複數個複合控制度量中之各個各別複合控制度量係藉由第二過程計算。

【0015】 在第二過程中，一組各別控制微擾選自該複數個微擾。該組各別控制微擾中之各控制微擾係針對該複數個成分中之不同成分(例如，基因)。在第二過程中，針對該組各別控制微擾中之各個各別控制微擾，獲得對應控制向量，從而獲得複數個各別控制向量。各對應控制向量

包含複數個控制元素。該複數個控制元素中之各控制元素包含：(i)在複數個各別實體控制實例暴露於各別控制微擾之後在該複數個實體控制實例中不同特徵在複數個特徵中之量測值之分佈度量；或(ii)使用在複數個各別實體控制實例暴露於各別控制微擾之後在該複數個各別實體控制實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量。在第二過程中，使用該複數個各別控制向量計算各別複合控制度量。各別控制度量量化該組各別控制微擾對對應不同成分的中靶效應。

【0016】控制度量係藉由第三過程計算。在第三過程中，(1)對於各個各別控制向量，計算(i)各別控制向量與(ii)在自該複數個控制向量除去各別控制向量之情況下該複數個各別控制向量之分佈度量之間的控制相似性度量，從而獲得該複數個各別控制向量之複數個控制向性度量。該複數個控制相似性度量中之各控制相似性度量獨特地對應於該組各別控制微擾中之微擾。計算各別複合控制度量，作為該複數個控制相似性度量之分佈度量。

【0017】該一或多個程式進一步包括用於將測試度量與虛無分佈相比較的指令。當測試度量不同於虛無分佈時，認為該組測試微擾相對於使用數組控制微擾所形成之虛無分佈區別對第一成分的中靶效應。

【0018】在一些實施例中，各不同特徵選自複數個特徵，且該複數個特徵中之各特徵表示在實體暴露於該各別測試微擾或控制微擾之後實體或實體之列舉部分之顏色、

紋理、或大小。

【0019】在一些實施例中，在該複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例中不同特徵之量測值之分佈度量為在該複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例中不同特徵之算術平均數、加權平均數、中程數、中樞紐(midhinge)、三均數(trimean)、溫塞平均數、中位數、或眾數。

【0020】在一些實施例中，在該複數個實體控制實例暴露於各別控制微擾之後該複數個實體控制實例中不同特徵之量測值之分佈度量為在該複數個實體控制實例暴露於各別控制微擾之後該複數個實體控制實例中不同特徵之算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數。

【0021】在一些實施例中，該組測試微擾由2與12之間個不同測試微擾組成。在一些實施例中，該組測試微擾由5與15之間個不同測試微擾組成。在一些實施例中，該組測試微擾由3與300之間個不同測試微擾組成。

【0022】在一些實施例中，該組測試微擾中之各測試微擾進一步具有對該複數個成分中除第一成分之外的一或多個成分的脫靶效應。

【0023】在一些實施例中，該複數個實體測試實例包含500個實體測試實例。在一些實施例中，該複數個實體測試實例包含5000個實體測試實例。

【0024】在一些實施例中，在獲得量測值之前，該複

數個實體測試實例暴露於各別測試微擾達至少一小時或至少一天。

【0025】在一些實施例中，該複數個測試元素由5個測試元素與10,000個測試元素之間組成。在一些實施例中，該複數個測試元素由100個測試元素與2000個測試元素之間組成。

【0026】在一些實施例中，在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量之集中趨勢量度為除該各別測試向量之外的該複數個測試向量中各對應測試元素在該複數個測試元素中之集中趨勢量度。

【0027】在一些實施例中，計算(i)各別測試向量與(ii)在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量之分佈度量之間的測試相似性度量，作為測試向量之對應元素與在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量之分佈度量之間的距離。在一些此類實施例中，距離為角距。例如，在一些此類實施例中，

角距計算為：

$$\frac{\sum_i^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

其中 A_i 為各別測試向量中之測試元素 i ， B_i 為除各別測試向量之外的該複數個測試向量中對應測試元素 i 在該複

數個測試元素中之分佈度量，且 n 為各別測試向量中元素之數目。

【0028】在一些實施例中，該複數個測試相似性度量之該集中趨勢量度為該複數個測試相似性度量之算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數。

【0029】在一些實施例中，該複數個複合控制度量包含100個複合控制度量，且各複合控制度量表示與該複數個微擾不同的控制微擾組合。在一些實施例中，該複數個複合控制度量包含1000個複合控制度量，且各複合控制度量表示與該複數個微擾不同的控制微擾組合。

【0030】在一些實施例中，該組各別控制微擾由3與10之間個不同控制微擾、5與15之間個不同控制微擾、3與300之間個不同控制微擾組成。在一些實施例中，該複數個各別實體控制實例包含1000個實體控制實例或包含5000個實體控制實例。

【0031】在一些實施例中，在獲得量測值之前，該複數個各別實體控制實例暴露於各別控制微擾達至少一小時或至少一天。

【0032】在一些實施例中，該複數個測試元素由5個測試元素與10,000個測試元素之間組成，該複數個控制元素由5個控制元素與10,000個控制元素之間組成，且該複數個測試元素中之各測試元素與該複數個控制元素中之對應控制元素之間存在一對一對應。

【0033】在一些實施例中，該複數個測試元素由100個測試元素與2000個測試元素之間組成，該複數個控制元素由100個控制元素與2000個控制元素之間組成，且該複數個測試元素中之各測試元素與該複數個控制元素中之對應控制元素之間存在一對一對應。

【0034】在一些實施例中，在自該複數個各別控制向量除去各別控制向量之情況下該複數個控制向量之集中趨勢量度為除該各別控制向量之外的該複數個各別控制向量中各對應控制元素在該複數個各別控制元素中之集中趨勢量度。在一些此類實施例中，計算(i)該各別控制向量與(ii)在自該複數個各別控制向量除去該各別控制向量之情況下該複數個各別控制向量之集中趨勢量度之間的控制相似性度量，作為該各別控制向量之對應元素與在自該複數個各別控制向量除去該各別控制向量之情況下該複數個各別控制向量之集中趨勢量度之間的距離。在一些此類實施例中，距離為例如如下計算之角距：

$$\frac{\sum_i^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

其中 A_i 為各別控制向量中之控制元素 i ， B_i 為除各別控制向量之外的該複數個各別控制向量中對應控制元素 i 在該複數個控制元素中之集中趨勢量度，且 n 為各別控制向量中元素之數目。

【0035】 在一些實施例中，該複數個控制相似性度量之該集中趨勢量度為該複數個控制相似性度量之算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數。

【0036】 在一些實施例中，測試度量與虛無分佈之比較包含使用虛無分佈計算測試度量之Z分數，其中當測試度量之Z分數超過臨限值(例如，2.5、6等)時，認為測試度量不同於虛無分佈。

【0037】 在一些實施例中，該複數個維度縮減成分中之各維度縮減成分為藉由主成分分析導出之主成分。

【0038】 在一些實施例中，該複數個特徵中之各特徵為經光學量測之光學特徵。

【0039】 在一些實施例中，第一亞組該複數個特徵為經光學量測之光學特徵，且第二亞組該複數個特徵為非光學特徵。

【0040】 在一些實施例中，該複數個特徵中之各特徵為經非光學量測之特徵。

【0041】 在一些實施例中，對於該複數個實體測試實例中之各測試實例，個別地量測該複數個不同特徵中之不同特徵。

【0042】 在一些實施例中，在該複數個實體測試實例中之至少一亞組測試實例上量測該複數個不同特徵中之不同特徵。

【0043】 在一些實施例中，在該複數個實體測試實例

暴露於各別測試微擾之後該複數個實體測試實例中不同特徵之量測值之分佈度量為在該複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例中不同特徵之多樣性量度(例如，範圍、標準偏差、或變異數)。

【0044】 在一些實施例中，在該複數個實體控制實例暴露於各別控制微擾之後該複數個實體控制實例中不同特徵之量測值之分佈度量為在該複數個實體控制實例暴露於各別控制微擾之後該複數個實體控制實例中不同特徵之多樣性量度(例如，範圍、標準偏差、或變異數)。

【0045】 本揭露之另一態樣提供一種用於確定一組測試微擾是否相對於使用數組控制微擾所形成之虛無分佈區別(例如，區分)對實體(例如細胞)之第一成分(例如基因)的中靶效應的方法。該組測試微擾包含該第一成分之複數個測試微擾，且該實體包含包括該第一成分之複數個成分。該方法包含針對該組測試微擾中之各個各別測試微擾獲得對應測試向量，從而獲得複數個測試向量。各對應測試向量包含複數個測試元素。該複數個測試元素中之各元素包含：(i)在複數個實體測試實例暴露於各別測試微擾之後在該複數個實體測試實例中不同特徵在複數個特徵中之量測值之分佈度量；或(ii)使用在複數個實體測試實例暴露於各別測試微擾之後在該複數個實體測試實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量。在該方法中，使用上文所述之該複數個測試向量計算複合測試度量。複合測試度量量化該

組測試微擾對第一成分的中靶效應。

【0046】藉由第一過程計算測試度量，該第一過程包含對於該複數個測試向量中之各個各別測試向量，計算(i)各別測試向量與(ii)在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量之集中趨勢量度之間的測試相似性度量，從而獲得該複數個測試向量之複數個測試相似性度量，該複數個測試相似性度量中之各測試相似性度量表示獨特地對應於第一組微擾中之微擾。第一過程進一步包含計算複合測試度量，作為該複數個測試相似性度量之分佈度量，諸如集中趨勢量度(算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數)。

【0047】在該方法中，計算虛無分佈。虛無分佈包含複數個複合控制度量。藉由第二過程計算該複數個複合控制度量中之各個各別複合控制度量，該第二過程包含自複數個微擾選擇一組各別控制微擾，其中該組各別控制微擾中之各控制微擾係針對該複數個成分中之不同成分。第二過程進一步包含針對該組各別控制微擾中之各個各別控制微擾，獲得對應控制向量，從而獲得複數個各別控制向量。各對應控制向量包含複數個控制元素。該複數個控制元素中之各控制元素包含：(i)在複數個各別實體控制實例暴露於各別控制微擾之後在該複數個實體控制實例中不同特徵在複數個特徵中之量測值之分佈度量；或(ii)使用在複數個各別實體控制實例暴露於各別控制微擾之後在該複

數個各別實體控制實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量。第二過程進一步包含使用該複數個各別控制向量計算各別複合控制度量。各別控制度量量化該組各別控制微擾對對應不同成分的中靶效應。

【0048】 藉由第三過程計算控制度量，該第三過程包含對於各個各別控制向量，計算(i)該各別控制向量與(ii)在自該複數個各別控制向量除去該各別控制向量之情況下該複數個各別控制向量之集中趨勢量度之間的控制相似性度量，從而獲得該複數個各別控制向量之複數個控制相似性度量，該複數個控制相似性度量中之各控制相似性度量獨特地對應於該組各別控制微擾中之微擾。第三過程進一步包含計算該各別複合控制度量，作為該複數個控制相似性度量之集中趨勢量度。

【0049】 該方法進一步包含將測試度量與虛無分佈相比較。當測試度量不同於虛無分佈時，認為該組測試微擾具有對該第一成分的中靶效應。

【0050】 本揭露之另一態樣提供一種非暫代電腦可讀存儲媒體及嵌入其中的一或多個電腦程式，其確定一組測試微擾是否具有對實體之第一成分的中靶效應，其中該組測試微擾包含複數個對第一成分的測試微擾，且該實體包含複數個成分，包括第一成分。該一或多個電腦程式包含當藉由電腦系統執行時使該電腦系統執行本揭露中所提供之任一方法的指令。

【圖式簡單說明】

【0051】 圖1繪示根據本揭露之實施例之用於確定一組測試微擾是否相對於使用數組控制微擾所形成之虛無分佈區別(例如，區分)對實體之第一成分的中靶效應的示範性工作流程，其中該組測試微擾包含該第一成分之複數個測試微擾，且該實體包含包括該第一成分之複數個成分。

【0052】 圖2繪示根據本揭露之實施例之用於確定一組測試微擾是否相對於使用數組控制微擾所形成之虛無分佈區別(例如，區分)對實體之第一成分的中靶效應的裝置。

【0053】 圖3繪示根據本揭露之實施例之用於確定一組測試微擾是否相對於使用數組控制微擾所形成之虛無分佈區別(例如，區分)對實體之第一成分的中靶效應的資料結構。

【0054】 圖4A、4B、4C、4D、4E、4F、及4G共同提供根據本揭露之各種實施例之用於確定一組測試微擾是否相對於使用數組控制微擾所形成之虛無分佈區別(例如，區分)對實體之第一成分的中靶效應的過程及特徵，其中流程圖之視情況選用之元素係藉由虛線框指示。

【0055】 圖5繪示根據本揭露之實施例，針對一組測試微擾中之各測試微擾，給定測試微擾之經量測特徵如何形成具有N個元素之觀察值之測試向量，且維度N之該組集體向量可經受特徵縮減技術，以基於資料集中N維向量

中個別元素之觀察變異來導出一組維度縮減成分。

【0056】圖6繪示根據本揭露之實施例，圖5之長度N之測試向量如何縮減至複數個長度T之向量，其中T為小於N之正整數。

【0057】在圖式之若干視圖中，相似的參考數字係指對應部分。

【實施方式】

相關申請案之交互參照

【0058】本申請案主張2018年6月1日提交之標題為「SYSTEMS AND METHODS FOR DISCRIMINATING EFFECTS ON TARGETS」之美國專利申請案第15/995,835號之優先權，該案主張2018年3月2日提交之標題為「SYSTEMS AND METHODS FOR EVALUATING WHETHER PERTURBATIONS DISCRIMINATE AN ON TARGET EFFECT」之美國專利第10,146,914號之優先權，該案主張2018年3月1日提交之標題為「SYSTEMS AND METHODS FOR EVALUATING WHETHER PERTURBATIONS DISCRIMINATE AN ON TARGET EFFECT」之美國臨時專利申請案第62/637,121號之優先權，該等案各自特此以引用方式併入。

【0059】參考圖1，本揭露依賴於包含複數個特徵108之量測值的資料集110之獲取。將實體實例鋪板至多孔盤102之各孔中並暴露於微擾。例如，將細胞實例暴露於靶

向細胞中之特定基因的 siRNA。對於暴露於微擾之各實體，量測該複數個特徵，從而形成資料集。在實踐中，此係藉由使用例如落射螢光顯微鏡捕獲多孔盤之影像 106 來實現。然後將影像 106 用作從多孔盤中之各孔提取若干不同特徵從而形成資料集 110 的基礎。將資料集 110 用於確定對實體之具體成分的中靶效應。例如，將資料集 110 用於確定測試組中各 siRNA 對具體細胞型中具體基因的中靶效應。

【0060】 現將詳細參考實施例，該等實施例之實例繪示於隨附圖式中。在以下詳細描述中，闡述許多特定細節以便提供對本揭露之透徹理解。然而，熟習此項技術者將顯而易知，本揭露可在沒有此等特定細節之情況下實踐。在其他情況下，未詳細描述熟知的方法、程序、成分、電路、及網路，以免不必要地模糊實施例之態樣。

【0061】 亦應理解，儘管本文可使用術語第一、第二等來描述各種元素，但此等元素不應受此等術語限制。此等術語僅用於將一個元素與另一元素區分開。例如，在不背離本揭露之範疇之情況下，第一受試者可稱為第二受試者，且類似地，第二受試者可稱為第一受試者。第一受試者及第二受試者均為受試者，但其不為相同的受試者。此外，術語「受試者」、「使用者」、及「患者」可在本文互換使用。術語胰島素筆針意謂合適於施加離散劑量胰島素的注射裝置，其中該注射裝置適於記錄並傳達劑量相關資料。

【0062】本揭露中使用之術語係僅出於描述具體實施例之目的且不意欲限制本發明。除非上下文另外明確指示，否則如本發明之說明書及隨附申請專利範圍中所用，單數形式「一個」、「一種」、及「該」意欲同樣包括複數形式。亦應理解，如本文所用之術語「及/或」係指且涵蓋一或多個相關所列項目之任何及所有可能的組合。應進一步理解，術語「包含(*comprises*及/或*comprising*)」在本說明書中使用時指定所述特徵、整數、步驟、操作、元素及/或成分之存在，而並不排除一或多個其他特徵、整數、步驟、操作、元素、成分及/或其群組之存在或添加。

【0063】如本文所用，術語「若」可解釋為意謂「當...時」、「在...之後」、或「回應於確定」、或「回應於偵測」，其取決於上下文。類似地，片語「若確定」或「若偵測出[所述狀況或事件]」可解釋為意謂「在確定[所述狀況或事件]之後」、或「回應於確定[所述狀況或事件]」、或「在偵測出[所述狀況或事件]之後」、或「回應於偵測出[所述狀況或事件]」，其取決於上下文。

【0064】結合圖1至3描述用於確定一組測試微擾是否在識別出對實體之第一成分的中靶效應中相對於使用數組控制微擾所形成之虛無分佈進行區別的系統250之詳細描述。因此，圖1至3共同繪示根據本揭露之系統之拓撲。在該拓撲中，有用於獲得多孔盤102之影像106的工作流程。

藉由中靶效應確定模組 204 使用影像以呈資料集 110 之形式提取多孔盤之各孔之多個特徵。然後藉由圖 2 及 3 之系統 250 如下文進一步詳述來處理資料集 110，以確定一組測試微擾是否具有對實體之第一成分的中靶效應。

【0065】參考圖 1，為用於確定一組測試微擾是否具有對實體之第一成分的中靶效應的系統 250。為此，系統 250 接收多孔盤 102 之數位影像 250，其中各多孔盤之各孔含有實體(例如細胞)樣本，其已暴露於複數個預定微擾中之一或多者。

【0066】參考圖 2，在典型實施例中，系統 250 包含一或多個電腦。圖 2 中出於說明之目的，系統 250 表示為單一電腦，其包括所有用於確定一組測試微擾是否具有對實體之第一成分的中靶效應的功能性。然而，本揭露不限於此。在一些實施例中，用於確定一組測試微擾是否具有對實體之第一成分的中靶效應的功能性遍佈於任何數目的網路式電腦、及/或存在於若干網路式電腦中之各者、及/或代管於位於遠程位置可通過通信網路 106 訪問的一或多個虛擬機器。熟習此項技術者應理解，任何廣泛不同的電腦拓撲均用於本申請案，且所有此類拓撲均在本揭露之範疇內。

【0067】轉向圖 2 及 3，考慮到前述內容，用於確定一組測試微擾是否具有對實體之第一成分的中靶效應的實例性系統 250 包含一或多個處理單元(CPU) 274、網路或其他通信介面 284、記憶體 192 (例如隨機存取記憶體)、視情況

藉由一或多個控制器 288 存取之一或多個磁碟儲存及/或持續裝置 290、用於互連前述組件的一或多個通信匯流排 213、使用者介面 278 (使用者界面 278 包括顯示器 282 及輸入 280 (例如鍵盤、小鍵盤、觸控螢幕))、及用於向前述組件供能的電源 276。在一些實施例中，使用已知計算技術諸如高速存取將記憶體 192 中之資料與不變性記憶體 290 無縫共享。在一些實施例中，記憶體 192 及/或記憶體 290 包括相對於一或多個中央處理單元 274 遠程定位的大量存儲器。換言之，存儲在記憶體 192 及/或記憶體 290 中的一些資料實際上可代管於在系統 250 外部但可藉由系統 250 使用網路介面 284 經由網際網路、內部網路、或其他形式的網路或電纜(如圖 2 中之元件 292 所繪示)以電子方式訪問的電腦。

【0068】 在一些實施例中，用於確定一組測試微擾是否具有對實體之第一成分的中靶效應的系統 250 之記憶體 192 存儲：

- 作業系統 202，其包括用於處理各種基本系統服務的程序；
- 中靶效應確定模組 204；
- 實體第一成分之一組測試微擾中各測試微擾之各別測試向量 206，各個各別測試向量 206 包含複數個測試元素 208，與不同特徵之分佈度量相關聯的各個各別測試元素 208 對應於各別測試元素，在暴露於測試微擾之後複數個實體測試實例 210 所支持之各測試元素 208 對應於各別測試

向量 206；

- 複合測試度量 212，其中複合測試度量 212 將該組測試微擾對第一成分的中靶效應量化為複數個測試相似性度量 214 之集中趨勢量度，其中各測試相似性度量 214 對應於各別測試向量 206 且因此對應於該組測試微擾中之測試微擾，且表示 (i) 各別測試向量 206 與 (ii) 在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量之集中趨勢量度之間的相似性；及

- 虛無分佈 216，其包含複數個複合控制度量 218，其中各複合測試度量 212 與一組對應控制微擾 220 相關聯，且其中該組對應控制微擾 220 中之各控制微擾 222 係針對該複數個實體成分中之不同成分 224，且其中存在對應於該組控制微擾 220 中之各個各別控制微擾 222 的控制向量 226，其中各控制向量包含複數個控制元素 228，各個各別控制元素 228 包含在複數個各別實體控制實例 230 暴露於各別控制微擾 222 之後該複數個各別實體控制實例中不同特徵在該複數個特徵中之量測值之分佈度量，且其中將各複合控制度量 218 計算為複數個對應控制相似性度量 219 之分佈度量，各控制相似性度量 219 表示對應控制向量 226 與在自複數個各別控制向量除去各別控制向量 226 之情況下控制微擾 222 之複數個各別控制向量之分佈度量之間的相似性。

【0069】在一些實施例中，中靶效應確定模組 204 可用任何瀏覽器(電話、平板電腦、膝上型電腦/桌上型電腦)訪問。在一些實施例中，中靶效應確定模組 204 在本機裝

置框架上運行，且可用於下載至運行作業系統202諸如Android或iOS的系統250上。

【0070】 在一些實施方案中，用於確定一組測試微擾是否具有對實體之第一成分的中靶效應的系統250之一或多個上文所識別之資料元素或模組存儲於一或多個先前所述之記憶裝置中，且對應於用於執行上文所述之功能的一組指令。上文所識別之資料、模組、或程式(例如，數組指令)不必實施為單獨的軟體程式、程序、或模組，且因此各種亞組的此等模組可合併或以其他方式再佈置於各種實施方案中。在一些實施方案中，記憶體192及/或290視情況存儲一亞組上文所識別之模組及資料結構。此外，在一些實施例中，記憶體192及/或290存儲上文未描述之額外模組及資料結構。

【0071】 在一些實施例中，用於確定一組測試微擾是否具有對實體第一成分的中靶效應的裝置250為智慧型手機(例如，iPHONE)、膝上型電腦、平板電腦、桌上型電腦、或其他形式的電子裝置。在一些實施例中，裝置250不為移動的。在一些實施例中，裝置250為移動的。

【0072】 現在揭示用於確定一組測試微擾是否具有對實體第一成分的中靶效應的系統250之細節，根據本揭露之實施例，關於系統之過程及特徵之流程圖的細節參考圖4A至圖4G進行揭示。在一些實施例中，系統之此類過程及特徵係藉由圖2中所繪示之中靶效應確定模組204進行。

【0073】 方塊402。參考圖4A之方塊402，為用於確定

一組測試微擾是否具有對實體之第一成分的中靶效應的電腦系統250。該組測試微擾包含該第一成分之複數個測試微擾，且該實體包含包括該第一成分之複數個成分。該電腦系統包含一或多個處理器274、記憶體192/290、及一或多個程式。該一或多個程式存儲於記憶體中且經組態以由該一或多個處理器執行。該一或多個程式包括用於執行一方法的指令。

【0074】在一些實施例中，實體為細胞培養物，諸如細胞株、初代細胞、或共培養物系統。實體之實例包括但不限於U2OS細胞、A549細胞、MCF-7細胞、3T3細胞、HTB-9細胞、HeLa細胞、HepG2細胞、HEKTE細胞、SH-SY5Y細胞、HUVEC細胞、HMVEC細胞、初代人類纖維母細胞、及初代人類肝細胞/3T3-J2纖維母細胞共培養物。在一些實施例中，實體為人類細胞培養物。

【0075】在一些實施例中，成分為實體基因體中之基因。

【0076】在一些實施例中，測試微擾為特異性識別實體成分的小干擾RNA (siRNA)。各siRNA為雙股RNA分子，長度為20-25個鹼基對，其藉由在轉錄之後降解mRNA從而阻止基因之轉譯來干擾具有互補核苷酸序列之特定基因之表現。siRNA為RNA雙螺旋，其可通過RNA誘導沉默複合物(RISC)所介導之靶mRNA之酶切割來減少基因表現。siRNA具有以接近特異性抑制靶基因的能力。參見Agrawal等人，2003，「RNA interference: biology,

mechanism, and applications,」 *Microbiol Mol Biol Rev.* 67: 657-85 ; 及 Reynolds 等人, 2004, 「 Rational siRNA design for RNA interference,」 *Nature Biotechnology* 22, 326-330, 其各自特此以引用方式併入。在一些此類實施例中, 藉由將 siRNA 轉染至實體中、DNA-載體介導生產、或病毒介導 siRNA 合成來達成微擾。參見例如 Paddison 等人, 2002, 「 Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells 」, *Genes Dev.* 16:948-958 ; Sui 等人, 2002, 「 A DNA vector-based RNAi technology to suppress gene expression in mammalian cells 」, *Proc Natl Acad Sci U S A* 99:5515-5520 ; Brummelkamp 等人, 2002, 「 A system for stable expression of short interfering RNAs in mammalian cells 」, *Science* 296:550-553 ; Paddison 等人, 2004, 「 Short hairpin activated gene silencing in mammalian cells 」, *Methods Mol Biol* 265:85-100 ; Wong 等人, 2003, 「 CIITA regulated plexin-A1 affects T-cell-dendritic cell interactions,」 *Nat Immunol* 2003, 4:891-898 ; Tomar 等人, 2003, 「 Use of adeno-associated viral vector for delivery of small interfering RNA.」 *Oncogene* 22:5712-5715 ; Rubinson 等人, 2003 「 A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference 」, *Nat Genet* 33:401-406 ; Moore 等人, 2005, 「 Stable inhibition of hepatitis B virus proteins by small interfering RNA

expressed from viral vectors」, J Gene Med; 及 Tran 等人, 2003, 「Expressing functional siRNAs in mammalian cells using convergent transcription, BMC Biotechnol 3:21, 其各自特此以引用方式併入。

【0077】在一些實施例中，測試微擾為直接自細胞或自表現出所關注之肌病的患者之體液、組織、或器官(例如，類風濕性關節炎患者之滑液)取得之材料。在一些實施例中，此材料被稱為「條件培養基」。例如，藉由實例，在一些實施例中，材料為滑液組織外植體(參見 Beekhuizen 等人, 2011, 「Osteoarthritic synovial tissue inhibition of proteoglycan production in human osteoarthritic knee cartilage: establishment and characterization of a long-term cartilage-synovium coculture」, Osteoarthritis 63, 1918, 其特此以引用方式併入)，其立即用作測試微擾或在用作微擾之前經培養預定的一段時間。藉由另一實例，在一些實施例中，材料為間葉系幹細胞(MSC)，其自經歷全髖關節置換術之人類患者之肝素化股骨幹骨髓抽出物單離並培養，接種於細胞培養基(例如杜爾貝寇氏改良伊格爾氏培養(Dulbecco's Modified Eagle Medium))中。參見 Buul, 2012, 「Mesenchymal stem cells secrete factors that inhibit inflammatory processes in short-term osteoarthritic synovium and cartilage explant culture」, Osteoarthritis and Cartilage 20, 1186, 其特此以引用方式併入。對於以單離自 BALB/C 小鼠之鼠類 MSC 之形式的條

件培養基之製備之實例，亦參見Kay等人，2017，「Mesenchymal Stem Cell-Conditioned Medium Reduces Disease Severity and Immune Responses in Inflammatory Arthritis」，Nature 7, 18019，其特此以引用方式併入。仍藉由另一實例，在一些實施例中，材料為作為經歷膝蓋置換手術的患者之手術廢棄物材料獲得之人類滑液外植體或軟骨外植體。在此類實施例中，微擾為直接自細胞或自表現出所關注之疾病的患者之體液、組織、或器官提取之材料，其在提取之後立即使用，或在將材料培養一段時間之後使用。在一些實施例中，在旨在刺激材料的因子之存在下培養材料。例如，在材料為間葉系幹細胞之情況下，在一些實施例中，藉由實例，在刺激MSC分泌免疫調節因子的TNF α 及IFN γ 之存在下培養材料。參見，Buil, 2012, Osteoarthritis and Cartilage 20, 1186，其特此以引用方式併入。對於條件培養基之製備之另一實例，參見Martin, 1981，「Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells」，PNAS 78, 7634，其特此以引用方式併入。

【0078】在一些實施例中，測試微擾為小髮夾RNA (shRNA)。參見Taxman等人，2006，「Criteria for effective design, construction, and gene knockdown by shRNA vectors」，BMC Biotechnology 6:7 (2006)，其特此以引用方式併入。在一些此類實施例中，藉由DNA-載體介導生

產、或如上文所引用之參考文獻中針對 siRNA 所大致討論的病毒介導 siRNA 合成來達成微擾。

【0079】在一些實施例中，測試微擾為回文重複序列 (CRISPR) 之情況下使用之單嚮導 RNA (sgRNA)。參見 Sander 及 Young, 2014, 「CRISPR-Cas systems for editing, regulating and targeting genomes」, *Nature Biotechnology* 32, 347-355, 其特此以引用方式併入，其中缺乏核酸內切酶活性之催化死亡 Cas9 (通常表示為 dCas9) 蛋白以 RNA 導引方式調控基因。靶向特異性系藉由單嚮導 RNA (sgRNA) 與基因體基因座之互補鹼基配對來確定。sgRNA 為嵌合非編碼 RNA，其可細分為三個區：20 nt 鹼基配對序列、42 nt dCas9 結合髮夾、及 40 nt 終止子。在一些實施例中，當設計用作微擾之合成 sgRNA 時，僅自整個模板修飾 20 nt 鹼基配對序列。此外，在一些實施例中，考慮二級變項，諸如脫靶效應及 dCas9 結合髮夾結構之維持。在一些實施例中，藉由在編碼 Cas9 之基因之兩個催化殘基 (D10A 及 H840A) 中引入點突變來使 Cas9 催化失活。參見 Jinek 等人, 2012, 「A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity」, *Science* 337, (6096), 816, 其特此以引用方式併入。如此進行時，dCas9 不能切割 dsDNA，但保留靶向 DNA 的能力。在一些此類實施例中，藉由 DNA-載體介導生產、或如上文所引用之參考文獻中針對 siRNA 所大致討論的病毒介導 sgRNA 合成來達成微擾。

【0080】 在一些實施例中，測試微擾為細胞介素或細胞介素之混合物。對於將實體暴露於以細胞介素之形式之微擾的合適檢定(例如，體外檢定，諸如長期培養起始細胞(LTCIC)檢定、卵石區形成細胞(CAFC)檢定、高增殖潛能集落形成細胞(HPP-CFC)檢定、及集落形成單位母細胞(CFU-BI)檢定，以及使用動物模型之體內檢定)，參見Heike及Nakahata, 2002, 「Ex vivo expansion of hematopoietic stem cells by cytokines」, *Biochim Biophys Acta* 1592, 313-321，其特此以引用方式併入。在一些實施例中，實體暴露於以細胞介素(例如，淋巴介質、趨化介素、干擾素、腫瘤壞死因子等)之形式的微擾。在一些實施例中，實體暴露於以淋巴介質之形式(例如，介白素2、介白素3、介白素4、介白素5、介白素6、顆粒球-巨噬細胞集落刺激因子、干擾素 γ 等)的微擾。在一些實施例中，實體暴露於以趨化介素諸如體內恆定趨化介素(例如，CCL14、CCL19、CCL20、CCL21、CCL25、CCL27、CXCL12、CXCL13等)及/或炎性趨化介素(例如，CXCL-8、CCL2、CCL3、CCL4、CCL5、CCL11、CXCL10)之形式的微擾。在一些實施例中，實體暴露於以干擾素(IFN)諸如I型IFN(例如，IFN- α 、IFN- β 、IFN- ϵ 、IFN- κ 、及IFN- ω)、II型IFN(例如，IFN- γ)、或III型IFN之形式的微擾。在一些實施例中，實體暴露於以腫瘤壞死因子諸如TNF α 或TNF阿伐之形式的微擾。

【0081】 在一些實施例中，測試微擾為化合物。在一

些此類實施例中，使用 phosphoflow 技術檢定此一化合物對實體的活性，諸如 Krutzik 等人，2008，「High-content single-cell drug screening with phosphospecific flow cytometry,」 *Nature Chemical Biology* 4, 132-142 中所揭示者，其特此以引用方式併入。在一些實施例中，測試微擾為分子量小於 2000 道耳頓的化合物。在一些實施例中，測試微擾為分子量小於 2000 道耳頓、小於 4000 道耳頓、小於 6000 道耳頓、小於 8000 道耳頓、小於 10000 道耳頓、或小於 20000 道耳頓的任何有機化合物。

【0082】在一些實施例中，測試微擾為滿足里賓斯基五規則 (Lipinski rule of five) 準則的化合物。在一些實施例中，測試微擾為滿足里賓斯基五規則之二或更多條規則、三或更多條規則、或所有四條規則的有機化合物：(i) 不多於五個氫鍵予體 (例如，OH 及 NH 基團)，(ii) 不多於十個氫鍵受體 (例如，N 及 O)，(iii) 分子量低於 500 道耳頓，及 (iv) LogP 低於 5。之所以稱之為「五規則」，是因為四個準則有三個涉及數字五。參見 Lipinski, 1997, 「Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings」, *Adv. Drug Del. Rev.* 23, 3-26，其特此以全文引用之方式併入本文。在一些實施例中，測試微擾滿足一或多個除里賓斯基五規則之外的準則。例如，在一些實施例中，測試微擾為具有五或更少個芳環、四或更少個芳環、三或更少個芳環、或二或更少個芳環的化合物。

【0083】 在一些實施例中，測試微擾為蛋白微擾諸如肽適體。肽適體為以高特異性及強親和力結合靶蛋白的組合蛋試劑。藉由如此進行，其可調節其同源靶標之功能。在一些實施例中，肽適體包含一個(或多個)構形限制短可變肽域，其兩端連接至蛋白支架。因為肽適體引入與治療分子所致之微擾相似之微擾，所以其之使用以高於通常藉由作用於蛋白表現水準之方法所提供之信賴水準的信賴水準。肽適體之組合性質使其能夠『裝飾』許多多態性蛋白表面，其生物學相關性可通過肽適體之表徵來推斷。結合可成藥(druggable)表面的生物活性適體可用於置換篩選檢定，以識別出至表面的小分子命中。參見例如 Baines 及 Colas, 2006, 「Peptide Aptamers as guides for small-molecule drug discovery」, Drug Discovery Today 11, 334-341, 其特此以引用方式併入。在一些實施例中，測試微擾為肽適體，即，經選擇或工程改造以結合特定靶分子的人工蛋白。在一些此類實施例中，肽適體包含一或多個藉由蛋白支架所展示之可變序列之肽環。在一些實施例中，肽適體單離自組合文庫。在一些實施例中，此一組合文庫單離進一步藉由定向突變或數輪可變區誘變及選擇來改良。在一些實施例中，肽適體文庫用作「誘變劑」，其中表現不同肽適體之文庫經引入至實體群體中，以便選擇所要表型，以及識別出引起所要表型的那些適體。

【0084】 在一些實施例中，微擾包含用一或多個功能部分衍生化之肽適體，該一或多個功能部分可引起其靶蛋

白之特定轉譯後修飾，或改變靶標之次細胞定位。參見例如 Colas 等人，2000，「Targeted modification and transportation of cellular proteins」，*Proc. Natl. Acad. Sci. USA.* 97 (25): 13720-13725，其特此以引用方式併入。在一些實施例中，形成適體可變區的肽經合成為與支架相同的多肽鏈之一部分，且藉由至其的鍵聯而限制在其N及C末端。此雙重結構限制降低可變區可採用的構形的多樣性。因此，肽適體可緊密結合其靶標，其中結合親和力與抗體所顯示之結合親和力相當(奈莫耳範圍)。肽適體支架通常是小的、有序的可溶性蛋白。一種此類支架為大腸桿菌硫氧化還原蛋白 *trxA* 基因產物 (TrxA)。參見 Reverdatto 等人，2015，「Peptide aptamers: development and applications」，*Curr. Top. Med. Chem.* 15 (12): 1082-1101，其特此以引用方式併入。在此等分子中，展示可變序列之單一肽而不是 TrxA -Cys-Gly-Pro-Cys-活性位點環中的 Gly-Pro 模體。對 TrxA 之改良包括用側翼半胱胺酸取代絲胺酸，此舉防止在環之鹼基處可能形成二硫鍵、引入 D26A 取代以減少低聚合、以及優化用於在人類細胞中表現的密碼子。Reverdatto 等人進一步揭示可使用的其他支架，如 Škrlec 等人，2015，「Non-immunoglobulin scaffolds: a focus on their targets」，*Trends Biotechnol.* 33 (7): 408-418，其特此以引用方式併入。在一些實施例中，肽適體為所選之酵母雙雜交系統及/或藉由噬菌體文庫及其他表面展示技術諸如 mRNA 展示、核糖體展示、細菌展示、及

酵母展示(例如生物淘選)構築之組合肽文庫。在一些實施例中，微擾為使用MimoDB資料庫中之肽的肽適體。參見Huang等人, 2011, 「MimoDB 2.0: a mimotope database and beyond」, *Nucleic Acids Research*. 40 (1): D271-D277, 其特此以引用方式併入。

【0085】在一些實施例中，測試微擾為選擇性地影響實體內蛋白-蛋白相互作用的肽。在一些此類實施例中，此蛋白-蛋白相互作用影響細胞內傳訊事件。參見例如Souroujon及Mochly-Rosen, 1998, 「Peptide modulators of protein-protein interactions in intracellular signaling」, *Nature Biotechnology* 16, 919-924, 其特此以引用方式併入。

【0086】在一些實施例中，測試微擾為核酸微擾，諸如核酸適體。核酸適體為短合成單股寡核苷酸，其特異性結合各種分子靶標，諸如小分子、蛋白、核酸、及甚至細胞及組織。參見Ni等人, 2011, 「Nucleic acid aptamers: clinical applications and promising new horizons」, *Curr Med Chem* 18(27), 4206, 其特此以引用方式併入。在一些情況下，核酸適體選自生物淘選方法，諸如SELEX(指數富集的配體系統進化技術(Systematic Evolution of Ligands by Exponential enrichment))。參見Ellington及Szostak, 1990, 「In vitro selection of RNA molecules that bind specific ligands」, *Nature* 346(6287), 818; 以及Tuerk及Gold, 1990, 「Systematic evolution of ligands by

exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase」, Science 249(4968), 505, 其各自特此以引用方式併入。SELEX篩選方法以 ssDNA 或 ssRNA 之隨機序列文庫開始, 其長度為 20-100 個核苷酸 (nt)。核酸序列之隨機化提供了 4^n 之多樣性, 其中 n 對應於經隨機化鹼基數。通常可在 SELEX 方法中產生並篩選多樣性大約 10^{16} 個適體。各隨機序列區側接恆定序列, 以使用於捕獲或引始 (priming)。為了克服核酸外切酶降解, 適體可經化學合成並用經修飾或反轉核苷酸加帽以防止末端降解。經修飾寡核苷酸亦可在選擇期間或之後併入適體中, 以增強核酸內切酶穩定性。一些經修飾核苷酸三磷酸, 特別是經 2'-O-修飾嘧啶, 可藉由 T7 RNA 聚合酶有效地併入核酸適體轉錄物中。選擇期間所包括之常見化學修飾為 2'-胺基嘧啶及 2'-氟嘧啶。參見 Ni 等人, 2011, 「Nucleic acid aptamers: clinical applications and promising new horizons」, Curr Med Chem 18(27), 4206, 其特此以引用方式併入。

【0087】在一些實施例中, 測試微擾為抗體或其他形式的生物製劑。在一些實施例中, 使用測試微擾文庫, 其中文庫之各成員為不同抗體。在一些此類實施例中, 抗體文庫包含 100 種抗體、1000 種抗體、或一萬種抗體。在一些此類實施例中, 使用噬菌體展示技術產生抗體文庫, 諸如 Wu 等人, 2010, 「Therapeutic antibody targeting of individual Notch receptors」, Nature 464, 1052-1057 中所揭示之技術, 其特此以引用方式併入。在一些實施例中, 使

用測試微擾文庫，其中文庫之各成員為不同生物製劑。在一些此類實施例中，生物製劑文庫包含100種生物製劑、1000種生物製劑、或一萬種生物製劑。在一些此類實施例中，實體暴露於呈抗體之形式的微擾。例如，在一些此類實施例中，此類抗體選擇性結合實體所表現之跨膜蛋白，引起選擇性調控實體內轉錄程式的級聯訊號。例如，如Wu等人(同上)中所揭示，Notch家族中之受體為經廣泛表現之跨膜蛋白，其起關鍵導管的作用，哺乳動物細胞通過其進行通訊，以調控細胞命運及生長。配體結合觸發受體負調控區(NRR)之構形變化，從而使ADAM(解整合素與金屬蛋白酶)蛋白酶在近膜位點進行切割，否則切割位點深埋於靜止NRR內。藉由c-分泌酶(c-secretase)複合物催化之後續膜內蛋白水解釋放細胞內域(ICD)以起始下游Notch轉錄程式。因此，在一些實施例中，測試微擾為抗體，其暴露於實體，從而引起實體內一或多種成分之轉錄的選擇性變化。

【0088】在一些實施例中，測試微擾為鋅指轉錄因子。在一些此類實施例中，將鋅指蛋白轉錄因子編碼至轉型至實體中的載體中，從而引起實體內一或多種靶向成分之表現控制。在一些此類實施例中，藉由呈鋅指蛋白之形式的微擾使用實體中多種(例如，功能相關的)成分所共有的序列，以用呈鋅指轉錄因子之形式的單一微擾控制所有此等成分之微擾。在一些實施例中，呈鋅指轉錄因子之形式的微擾藉由靶向並調節控制實體內相關成分的内源性轉

錄因子之表現來靶向實體內相關成分之家族。參見例如 Doyon, 2008, 「Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases」, Nature Biotechnology 26, 702-708, 其特此以引用方式併入。

【0089】在一些實施例中，各測試微擾圍繞與特定疾病相關的生物訊號或研究中其他形式的生物訊號之特異性建立信賴，例如，藉由獨特地抑制與疾病接近(相關)的生物途徑中之成分而由測試實體所表現出之具體表型(或研究中其他形式的生物訊號)，而各控制微擾對不與研究中生物訊號之成分接近的實體成分具有相似量值的作用。因此，在一些實施例中，該組測試微擾藉由靶向與研究中生物訊號(例如，疾病)相關聯的實體之遺傳成分來提供生物效應，而控制微擾靶向不與研究中該生物訊號接近的實體之遺傳成分。

【0090】在一些實施例中，各測試微擾及各控制微擾為 siRNA、sgRNA、或 shRNA。

【0091】方塊404。參考圖4A之方塊框404，在一些實施例中，該組測試微擾由2與12之間個不同的測試微擾、5與15之間個不同的測試微擾、6與50之間個不同的測試微擾、或3與300之間個不同的測試微擾組成。

【0092】方塊406。參考圖4A之方塊406，在一些實施例中，該組測試微擾中之測試微擾進一步具有對該複數個成分中除第一成分之外的一或多個成分的脫靶效應。

【0093】方塊408。參考圖4A之方塊408，其中針對該

組測試微擾中之各個各別測試微擾獲得對應測試向量 206，從而獲得複數個測試向量。各對應測試向量 206 包含複數個測試元素。

【0094】在一些實施例中，該複數個測試元素中之各測試元素 208 包含在複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例 210 中不同特徵在複數個特徵中之量測值之分佈度量。

【0095】在一些實施例中，此類資料係使用自動化細胞成像系統(例如 ImageXpress Micro, Molecular Devices)獲得，其中在將實體經這一組發射不同離散波長的染料(例如，赫斯特 33342、Alexa Fluor 594 鬼筆環肽等)染色並暴露於微擾之後，將其佈置於多孔盤(例如，384 孔盤)中。在一些實施例中，在 20×放大率、2×分箱下，用藉由所用之標記物染料所確定之暴露(例如，赫斯特為 15 ms，鬼筆環肽為 1000 ms)對實體進行成像。對於各孔，在一些實施例中，使用自對焦於特定染料通道(例如，赫斯特通道)上的雷射來發現最佳焦點。在一些實施例中，然後將自動化顯微鏡程式化以收集 32 個影像之 z 堆疊($z = 0$ 位於最焦平面，16 個影像高於焦平面，16 個影像低於焦平面)，其中斷片(slice)之間為 2 μm 。在一些實施例中，各孔在其中含有數千個實體，且因此藉由照相機捕獲之孔之各數位表示表示若干不同孔之各者中的數千實體。在一些實施例中，使用分段軟體來識別出數位影像中之個別實體，且此外識別出個別實體內的各種成分(例如，細胞成分)。一旦分段

並識別細胞成分，便對此等成分進行數學轉換以獲得特徵之量測值。

【0096】因此，為了說明，在該組測試微擾為五個測試微擾之情況下，將有五個測試向量206，即，該組五個測試微擾中之各測試微擾之測試向量206。此外，各測試向量206將包含複數個測試元素(例如，五或更多個測試元素208、十或更多個測試元素208、二十或更多個測試元素208、100或更多個測試元素208、或一千或更多個測試元素)。

【0097】在一些實施例中，測試向量中之各測試元素208表示在暴露於測試微擾之後從測試實體210之測試實例量測的複數個特徵中的不同特徵。例如，考慮第一測試向量206對應於第一測試微擾的情況，其中第一測試向量包含20個元素。各測試元素表示在暴露於第一次測試微擾之後從實體測試實例量測的不同特徵。第一測試向量206中之第一測試元素208表示在暴露於第一微擾之後複數個實體測試實例210中之各者中第一測試特徵之量測值。更特定言之，各測試元素208表示在暴露於對應於含有測試元素208之測試向量206的測試微擾之後從測試實體之各測試實例210量測的對應測試特徵之分佈度量。

【0098】在複數個實體測試實例暴露於該複數個微擾中之各別微擾之後在該複數個實體測試實例中該複數個特徵之量測值得到N維空間，其中該N維空間中之各整數為該複數個特徵中之不同特徵。

【0099】 在一些替代實施例中，該複數個測試元素中之各測試元素 208 包含使用在複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量。此類實施例為有利的，因為其消除或減少實體之高度相關的特徵(如大小、核面積、及核周長)之間的冗餘。

【0100】 例如，參考圖 5，在一些實施例中，將維度縮減技術(例如，主成分分析、亞組選擇、或緊縮方法)應用於在複數個實體中之各實體上獲得之觀測值 $p_{m,q,n}$ ，其中該複數個實體中之各實體已暴露於複數個控制微擾中之控制微擾。在一些此類實施例中，此複數個控制微擾與下文所述之方塊 448 至 474 中所用的複數個控制微擾相同。即，下文所述之方塊 448 至 474 中所用的控制微擾用於以複數個維度縮減成分之形式獲悉縮減的維度空間。在此類實施例中，圖 3 之控制向量用於識別出維度縮減成分。

【0101】 在一些替代實施例中，用於學習縮減維度空間的該複數個控制微擾與方塊 448 至 474 中所用之一組控制微擾不同。在此類實施例中，以類似於圖 3 中所揭示且下文參考方塊 448 至 474 所述之方式，針對此組控制微擾中之各微擾多次量測複數個特徵，然後使用此資料來確定該複數個維度縮減成分。

【0102】 在一些實施例中，用於維度縮減的該複數個微擾包含 200 或更多個控制微擾、300 或更多個控制微擾、

400或更多個控制微擾、或500或更多個控制微擾。

【0103】在一些實施例中，用於維度縮減的該組控制微擾中之各微擾靶向且抑制實體之該複數個成分中之具體成分。在一些實施例中，該複數個控制微擾包含 siRNA，其中各個此類 siRNA 特異性結合並因此抑制細胞中具體基因之 mRNA。

【0104】因此，參考圖 5，針對該複數個控制微擾所觀察之該複數個特徵中之各特徵 N 有複數個觀察值，其形成維度縮減之基礎，且 N 個特徵可被視為 N 維特徵空間。因此，參考圖 5，各水平線構成具有 N 個元素的觀測值向量，且維度 N 之該組集體向量可經受特徵縮減技術，以基於資料集中 N 維向量中個別元素之觀察變異來導出一組維度縮減成分。因此，維度縮減技術用於識別出複數個維度縮減成分，其中該複數個維度縮減成分共同表示在所有量測之控制實體實例中在 N 維特徵空間之特徵中所觀察之變異數。

【0105】然後將該複數個維度縮減成分用於表示 N 維特徵空間之第一特徵 208-1 至 208- N 組。然而，不一定是該複數個特徵 208 中給定特徵 208 之觀察值對具體維度縮減成分有貢獻的情況。例如，控制資料之維度縮減分析可確定一些特徵對解釋所觀察之變化沒有貢獻，且因此維度縮減分析不將此等特徵併入至任何維度縮減成分中。因此，在此類實施例中，在圖 5 之該複數個實體控制實例之暴露之後，將該複數個實體控制實例中該複數個特徵 (N 個特徵)

首先經受維度縮減，從而計算該複數個維度縮減成分(T個維度縮減成分)，其中T為小於N的正整數。因此，在該組特徵為一千個特徵之情況下，將有少於五百個維度縮減成分。

【0106】參考圖6，一旦使用一組控制微擾識別出維度縮減成分，便將其應用於暴露於測試微擾之測試實體之觀察特徵。在此類情況下，各個各別測試向量將表示該複數個測試微擾中之微擾，且基於在已暴露於對應於各別測試向量的測試微擾的實體中維度縮減成分所表示之測試特徵之值，各個各別測試向量之各個各別元素將包括該複數個維度縮減成分之對應維度縮減成分之值。

【0107】方塊410。參考圖4A之方塊410，用於形成向量206之元素之基礎或用作維度縮減成分之基礎的各特徵選自複數個特徵。該複數個特徵中之各特徵表示在實體暴露於該各別測試微擾或控制微擾之後實體或實體之列舉部分之顏色、紋理、或大小。示例性特徵包括但不限於實體面積、實體周長、實體寬高比、肌動蛋白含量、肌動蛋白紋理、實體堅固性、實體範圍、實體核面積、實體核周長、實體核寬高比等一些此類特徵。在一些實施例中，實例性特徵包括但不限於2017年11月19日訪問之參考文獻Gustafsdottir及Ljosa等人，2013, PLOS Tenth Anniversary, <https://doi.org/10.1371/journal.pone.0080999>之表S2中所見之任何特徵，其特此以引用方式併入。在一些實施例中，使用軟體程式Cellprofiler量測並獲得此類特徵。參見

Carpenter 等人，2006，「CellProfiler: image analysis software for identifying and quantifying cell phenotypes」，*Genome Biol.* 7, R100 PMID: 17076895；Kamentsky 等人，2011，「Improved structure, function, and compatibility for CellProfiler: modular high-throughput image analysis software」，*Bioinformatics* 2011/doi. PMID: 21349861 PMCID: PMC3072555；及 Jones 等人，2008，CellProfiler Analyst: data exploration and analysis software for complex image-based screens, *BMC Bioinformatics* 9(1):482/doi: 10.1186/1471-2105-9-482. PMID: 19014601 PMCID: PMC261443，其各自特此以引用方式併入。

【0108】在一些實施例中，使用全轉錄物組散彈槍測序(RNA-Seq)實驗獲得針對一實體或複數個實體所觀察之一或多個特徵，該實驗以映射至實體成分的轉錄物讀長之計數之形式量化實體(例如，單一實體)之基因表現。因此，在一些實施例中，RNA-Seq實驗目的在於同時自數百萬個短讀長重建成分之所有全長 mRNA 轉錄物。RNA-Seq 有助於查找替代基因剪接轉錄物、轉錄後修飾、基因融合、突變/SNP、及基因表現隨時間的變化、或不同組或治療中基因表現之差異的能力。參見例如 Maher 等人，2009，「Transcriptome sequencing to detect gene fusions in cancer」，*Nature*. 458 (7234): 97-101，其特此以引用方式併入。除 mRNA 轉錄物之外，RNA-Seq 可評估並量化實體內不同 RNA (包括總 RNA、mRNA、miRNA、lncRNA、

snoRNA、或 tRNA) 群體之個別成員。因此，在一些實施例中，針對實體所觀察之一或多個特徵為如使用 RNA-Seq 技術所確定之特定 RNA 物種之個別量。在一些實施例中，RNA-Seq 實驗產生成分之計數(例如，mRNA 讀長之數位計數)，其受生物及技術變異影響。在一些實施例中，RNA-Seq 組裝係使用 Li 等人，2008，「IsoLasso: A LASSO Regression Approach to RNA-Seq Based Transcriptome Assembly,」 Cell 133, 523-536 中所揭示之計數進行，其特此以引用方式併入。

【0109】 在一些實施例中，使用轉錄剖析方法諸如量測一組資訊轉錄物的 L1000 組獲得針對實體或複數個實體所觀察的一或多個特徵。在此一方法中，將連接介導之擴增(LMA)擴展至 1,000-plex 反應，接著在螢光定位之微球珠粒上捕獲擴增產物。例如，溶解在 384 孔盤中生長的細胞，並在寡-dT 塗佈之盤上捕獲 mRNA 轉錄物。自捕獲之錄物合成 cDNA，並使用攜帶獨特的 24 聚體條形碼序列及 5' 生物素標記的基因座特異性寡核苷酸進行 LMA。藉由與具有不同螢光顏色的聚苯乙烯微球(珠粒)雜交來偵測生物素化 LMA 產物，各微球與與條形碼互補的寡核苷酸偶聯，然後用鏈親和素-藻紅素染色。以此方式，可分析各珠粒之顏色(表示界標(landmark)一致性)及藻紅素訊號之螢光強度(表示界標豐度)。參見 Subramanian 等人，「A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles」, Cell 171(6), 1437，其特此以引用方

式併入。在一些實施例中，使用此檢定量測500與1500之間個不同的資訊轉錄物。

【0110】 在一些實施例中，使用微陣列獲得針對實體或複數個實體所觀察的一或多個特徵。微陣列(亦稱為DNA晶片或生物晶片)為附接於固體表面的微觀核酸斑點之集合，其可用於同時量測大量基因之表現水準。各核酸斑點含有皮莫耳特定核酸序列，稱為探針(或報導子或寡核苷酸(oligo))。此等可為一段短基因或其他核酸元件，其用於在高嚴格條件下與cDNA或cRNA(亦稱為反義RNA)樣本(稱為靶標)雜交。例如，藉由非限制性實例，在一些實施例中，使用微陣列，諸如Affymetrix基因晶片微陣列，其為高密度寡核苷酸基因表現陣列。Affymetrix微陣列基因晶片上之各基因通常由探針組表示，該探針組由11對不同的25-bp寡核苷酸組成，覆蓋該基因之經轉錄區之特徵。各對由完美匹配(PM)及錯配(MM)寡核苷酸組成。PM探針與具體標準基因型之序列(通常雜交之一個親本)精確匹配，而MM在中心第13個鹼基具有不同的單一取代。MM探針經設計用於將由非特異性雜交所致之雜訊與特定雜交訊號區別開。參見Jiang, 2008, 「Methods for evaluating gene expression from Affymetrix microarray datasets」, BMC Bioinformatics 9, 284, 其特此以引用方式併入。

【0111】 在一些實施例中，自ChIP-Seq資料獲得針對實體或複數個實體所觀察的一或多個特徵。參見例如

Quigley及Kintner, 2017, 「Rfx2 Stabilizes Foxj1 Binding at Chromatin Loops to Enable Multiciliated Cell Gene Expression」, PLoS Genet 13, e1006538, 其特此以引用方式併入。在一些實施例中, ChIP-seq用於確定轉錄因子及其他染色質相關蛋白如何影響實體(例如, 細胞)中表型影響機制。可藉由染色質免疫沉澱單離出轉錄因子及其他蛋白之直接物理相互作用之特定DNA位點。ChIP產生體內結合所關注之蛋白(成分)的靶DNA位點文庫。然後可結合全基因體序列資料庫使用平行序列分析以分析任何蛋白與DNA之相互作用模式(Johnson等人, 2007, 「Genome-wide mapping of in vivo protein-DNA interactions」, Science. 316: 1497-1502, 其特此以引用方式併入)或任何表觀遺傳染色質修飾之模式。這可應用於可ChIP蛋白及修飾組諸如轉錄因子、聚合酶及轉錄機構、結構蛋白、蛋白修飾、及DNA修飾。

【0112】ChIP選擇性地富集活細胞(實體)中具體蛋白(成分)所結合之DNA序列。ChIP過程使用對所關注之蛋白(成分)的抗體富集特定的交聯DNA-蛋白複合物。然後將寡核苷酸轉接子添加至結合所關注之蛋白的小段DNA, 以實現大規模平行測序。在大小選擇之後, 使用基因體定序儀同時對所有所得ChIP-DNA片段進行測序。單次測序運行可在高解析度之情況下掃描全基因體關聯, 意謂可在染色體上精確定位特徵。可使用各種測序方法。在一些實施例中, 使用轉接子接合ChIP DNA片段在附體流動池基材上

之叢集擴增以產生純系複本來分析序列。藉由基因體分析程式對流動池表面上之所得高密度模板叢集陣列進行測序。各模板叢集使用螢光標記之可逆終止子核苷酸並行地進行合成定序法(sequencing-by-synthesis)。在各讀取期間逐鹼基地對模板進行測序。然後，數據收集及分析軟體將樣本序列與已知基因體序列比對以識別出ChIP-DNA片段。

【0113】在一些實施例中，針對實體或複數個實體所觀察之一或多個特徵獲自ATAC-seq (染色質開放性測序技術 (Assay for Transposase-Accessible Chromatin using sequencing))，其為分子生物學中研究染色質可及性所用之技術。參見 Buenrostro 等人，2013，「Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position」，Nature Methods 10, 1213-1218，其特此以引用方式併入。在一些實施例中，ATAC-seq利用轉位酶Tn5對實體基因體DNA的作用。參見例如 Buenrostro 等人，2015，「ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide」，Current Protocols in Molecular Biology: 21.29.1-21.29.9，其特此以引用方式併入。轉位酶為催化轉位子移動至基因體中其他部分的酶。儘管天然存在之轉位酶之活性水準低，但ATAC-seq採用經突變、極具活性(hyperactive)的轉位酶。高活性允許高效切割經暴露DNA且同時接合特定序列(稱為轉接子)。轉接

子接合之DNA片段然後經單離、藉由PCR進行擴增並用於下一代測序。參見Buenrostro等人, 2013, 「Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position」, Nature Methods 10, 1213-1218, 其特此以引用方式併入。

【0114】儘管不旨在限於任何具體理論, 但據信轉位子一般優先併入不含核小體(無核小體區)或數段暴露DNA之基因體區。因此, 基因體中某些基因座之序列之富集指示該區中不存在DNA結合蛋白或核小體。ATAC-seq實驗通常將產生數百萬下一代測序讀長, 其可成功映射於參考基因體。在消除重複之後, 各測序讀長指向基因體上在實驗期間發生一個轉位(或切割)事件的位置。然後, 可為各基因體位置分配切割計數, 並在鹼基對解析度之情況下產生訊號。將此訊號用作本揭露之一些實施例中的特徵。基因體中在實驗期間DNA為可及的區域將含有顯著更多的測序讀長(因為其為轉位酶優先作用的位置)且形成可用峰處理(peak calling)工具偵測的ATAC-seq訊號峰。在一些實施例中, 將此類峰及其在基因體中的定位用作特徵。在一些實施例中, 藉由整合另外的基因體及表觀基因體資料諸如關於組蛋白修飾之資訊或活性轉錄之證據來將此等區域進一步歸類至各種調控元件類型(例如, 啟動子、增強子、絕緣子等)。在富集ATAC-seq訊號的區域內, 亦可觀察到具有耗盡訊號的亞區。此等亞區通常僅有幾個鹼基對長, 被

認為是DNA結合蛋白之「足跡」。在一些實施例中，將此類足跡或其不存在或存在用作特徵。

【0115】 在一些實施例中，針對實體或複數個實體所觀察之一或多個特徵係獲自高通量簡化表示表現剖析方法，稱為L1000，其揭示於Subramanian等人，2017，「A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles」，Cell 171, 1437-1452，其特此以引用方式併入。L1000量測人類細胞之的978個「界標」基因之mRNA轉錄物豐度。(L1000中之「L」係指檢定中所量測之界標基因。) 使用500種顏色的Luminex珠粒進行量測，使得兩種轉錄物藉由單一珠粒顏色識別。亦量測80種對照轉錄物之表現，選擇其在細胞狀態上的不變表現。L1000量測轉錄物組之高代表性亞組，與其他表現剖析分析技術相比，以較小部分的成本捕獲大部分的資訊。在一些實施例中，將此等978個「界標基因」之量測值應用於推理算法，以推斷轉錄物組中11,350個額外基因之表現。在一些實施例中，所量測之979個界標基因之表現用作本揭露中之特徵。在一些實施例中，將所推斷之11,350個額外基因之表現用作本揭露中之特徵。

【0116】 在一些實施例中，將使用Luminex珠粒之流動式細胞測量術方法用於獲得針對實體或複數個實體所觀察之一或多個特徵之值。參見例如Süsal 等人，2013，Transfus Med Hemother 40, 190-195，其特此以引用方式併入。例如，Luminex支持之單抗原搏動(Luminex-supported

single antigen bead, L-SAB)測試允許表徵人類白血球抗原 (HLA)抗體特異性。在此一流動式細胞測量方法中，採用塗佈有重組單抗原HLA分子之微珠，以區分兩個反應管中對100種不同的HLA I類對偶基因及100種HLA II類對偶基因的抗體反應性。自平均螢光強度(MFI)導出抗體反應性強度之近似值，且在一些實施例中，此用作本揭露中之特徵。除針對HLA-A、HLA-B、HLA-C、HLA-DR、及HLA-DQB抗原的抗體反應性之外，L-SAB能夠偵測針對HLA-DQA、HLA-DPA、及HLA-DPB抗原的抗體。在一些實施例中，使用其他Luminex套組偵測非HLA抗體，以根據本揭露導出實體之一或多個特徵之值。例如，在一些實施例中，使用主要組織相容性基因複合物I類相關鏈A (MICA)及人類嗜中性球抗體、以及代替重組HLA分子利用親和力純化之匯集人類HLA分子的套組以根據本揭露之實施例確定實體之特徵之值，該等親和力純化之匯集人類HLA分子係來自多個細胞株(在無進一步說明之情況下偵測HLA抗體之存在的篩選測試)或其中各珠粒群體具有衍生自單一實體之細胞株之HLA I類蛋白或HLA II類蛋白的表型小組(小組反應性，PRA測試)。

【0117】 在一些實施例中，使用流動式細胞測量方法諸如螢光細胞條形碼化來獲得針對實體所觀察之一或多個特徵之值。螢光細胞條形碼化(FCB)可例如藉由複用實體樣本之後染色並在細胞儀上獲取來實現高含量流動式細胞測量術。將個別細胞樣本(實體)用螢光染料之獨特標識條

碼化或標記以使得其可混合在一起，染色並作為單一樣本進行分析。藉由在染色前混合樣本，抗體消耗通常減少10至100倍。此外，通過將對照樣本及處理樣本組合，資料穩健性增加，其使移液誤差、染色變化、及對正規化的需要最小。最後，獲取速度增加，使得能夠用標準細胞儀硬體進行大型剖析試驗。參見例如 Krutzik, 2011, 「Fluorescent Cell Barcoding for Multiplex Flow Cytometry」, Curr Protoc Cytom, 第6章: 第6.31單元，其特此以引用方式併入。

【0118】在一些實施例中，使用代謝體學來獲得針對實體或複數個實體所觀察之一或多個特徵之值。代謝體學為為了獲得對疾病途徑的生理解而對小分子進行之系統性評估。在一些實施例中，此類代謝體學包含評估糖尿病 (Newgard 等人, 2009, 「A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance,」 Cell Metab 9: 311-326, 2009)及ESRD (Wang, 2011, 「RE: Metabolite profiles and the risk of developing diabetes」, Nat Med 17: 448-453)之血漿代謝體學。在一些實施例中，使用尿代謝體學來獲得一或多個特徵之值。尿代謝體學提供較大範圍的可量測代謝物,因為腎負責濃縮多種代謝物並將其排泄在尿中。此外，尿代謝體學可提供對與腎功能障礙相關的生化途徑的直接理解。參見例如 Sharma, 2013, 「Metabolomics Reveals Signature of Mitochondrial

Dysfunction in Diabetic Kidney Disease」, J Am Soc Nephrol 24, 1901-12, 其特此以引用方式併入。

【0119】在一些實施例中，使用質譜法來獲得針對實體或複數個實體所觀察之一或多個特徵之值。例如，在一些實施例中，使用蛋白質譜法來獲得針對實體或複數個實體所觀察之一或多個特徵之值。具體而言，在一些實施例中，使用實體內天然巨分子組裝體之生化分餾之後進行串聯質譜法來獲得針對實體或複數個實體所觀察之一或多個特徵之值。參見例如 Wan 等人，2015, 「Panorama of ancient metazoan macromolecular complexes」, Nature 525, 339-344, 其特此以引用方式併入。串聯質譜法亦稱為 MS/MS 或 MS², 涉及多個質譜選擇步驟，其中在各階段之間發生一些形式的碎裂。在串聯質譜儀中，在質譜法之第一階段 (MS¹) 中，離子源中形成離子，且藉由質荷比進行分離。選擇具有具體質荷比之離子 (前驅離子)，且藉由碰撞誘導解離、離子-分子反應、光解離、或其他過程產生碎體離子 (產物離子)。然後在質譜法之第二階段 (MS²) 中將所得離子分離並偵測。在一些實施例中，此類離子之偵測及/或存在用作針對實體或複數個實體所觀察之一或多個特徵之值。

【0120】在一些實施例中，針對實體或複數個實體所觀察之特徵為調節此類實體內蛋白之活性的轉譯後修飾。在一些此類實施例中，使用質譜肽測序及分析技術來偵測並識別此類轉譯後修飾。在一些實施例中，使用同位素標

記策略以及質譜法來研究修飾之動力學，且此用作針對實體或複數個實體所觀察之特徵。參見例如 Mann 及 Jensen, 2003 「Proteomic analysis of post-translational modifications」, Nature Biotechnology 21, 255-261，其特此以引用方式併入。在一些實施例中，使用質譜法來確定實體中之剪接變異體，例如，實體內成分之剪接變異體，且此類剪接變異體及此類剪接變異體之偵測用作一或多個實體之量測特徵。參見例如 Nilsen 及 Graveley, 2010, 「Expansion of the eukaryotic proteome by alternative splicing」, 2010, Nature 463, 457-463，其特此以引用方式併入。

【0121】在一些實施例中，使用成像細胞測量術來獲得針對實體或複數個實體所觀察之一或多個特徵之值。成像流動式細胞測量術將標準流動式細胞測量術之統計功效及螢光靈敏度與數位顯微法之空間解析度及定量形態學相組合。參見例如 Basiji 等人，2007, 「Cellular Image Analysis and Imaging by Flow Cytometry」, Clinics in Laboratory Medicine 27, 653-670，其特此以引用方式併入。

【0122】在一些實施例中，使用電生理學來獲得針對實體或複數個實體所觀察之一或多個特徵之值。參見例如 Dunlop 等人，2008, 「High-throughput electrophysiology: an emerging paradigm for ion-channel screening and physiology」, Nature Reviews Drug Discovery 7, 358-368，

其特此以引用方式併入。

【0123】在一些實施例中，使用蛋白質體學成像/3D 成像來獲得針對實體或複數個實體所觀察之一或多個特徵之值。參見例如美國專利申請案第20170276686 A1號，標題為「Single Molecule Peptide Sequencing」，其特此以引用方式併入。此類方法可用於來自在單分子水準下大規模測序實體或複數個實體之混合物中之單一肽。

【0124】在一些實施例中，特徵表示在實體已暴露於微擾(例如，siRNA)以及一組在不同波長下發射的螢光染色劑之後此類實體之量測值，螢光染色劑諸如伴刀豆球蛋白A/Alexa Fluor 488綴合物(Invitrogen，目錄號C11252)、赫斯特33342(Invitrogen，目錄號H3570)、SYTO 14綠色螢光核酸染色劑(Invitrogen，目錄號S7576)、鬼筆環肽/Alexa Fluor 568綴合(Invitrogen，目錄號A12380)、及MitoTracker深紅(Invitrogen，目錄號M22426)。在一些實施例中，量測特徵包括經標記細胞結構之染色強度、紋理圖案、大小、及形狀，以及跨通道的染色劑之間的相關性、及細胞之間與細胞內結構之間的相鄰關係。在一些實施例中，使用在兩個、三個、四個、五個、六個、七個、或八個通道中成像的兩種、三種、四種、五種、六種、七種、八種、九種、十種、或多於10種螢光染色劑量測實體中之特徵，包括細胞成分或此類實體內之區室。在一些實施例中，自單一實體量測特徵。在一些實施例中，自單一實體之區室或成分(例如，核、內質網、核仁、細胞質

RNA、F-肌動蛋白細胞骨架、高基氏體、質膜、粒線體) 量測特徵。在一些實施例中，此類通道包含(i)激發波長範圍及(ii)濾光片波長範圍，以捕獲在量測之前實體已暴露的該組染料中具體染料之發射。所用之染料及針對五個合適通道量測特徵的實體成分之類型之實例提供於以下表1中，其改寫自Bray等人，2016，「Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes」，Nature Protocols, 11, 第1757-1774頁之表1，其特此以引用方式併入。

表1-用於量測特徵之實例性通道

通道	染料	濾光片 (激發；nm)	濾光片 (發射；nm)	實體成分或區室
1	赫斯特33342	387/11	417-477	核
2	伴刀豆球蛋白A/Alexa Fluor 488綴合物	472/30a	503-538a	內質網
3	SYTO 14綠色螢光核酸 染色劑	531/40	573-613	核仁、細胞質RNA
4	鬼筆環肽/Alexa Fluor 568綴合物、麥胚凝集素 /Alexa Fluor 555綴合物	562/40	622-662c	F-肌動蛋白細胞骨 架、高基氏體、質膜
5	MitoTracker深紅	628/40	672-712	粒線體

【0125】方塊412。參考圖4A之方塊412，在一些實施例中，在該複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例中不同特徵之量測值之分佈度量為在該複數個實體測試實例暴露於各別測試微擾之後該複數

個實體測試實例中不同特徵之算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數。

【0126】方塊414。參考圖4B之方塊414，在一些實施例中，該複數個實體測試實例210包含100個實體測試實例、250個實體測試實例、400個實體測試實例、800個實體測試實例、1000個實體測試實例、2000個實體測試實例、3000個實體測試實例、4000個實體測試實例、5000個實體測試實例、6000個實體測試實例、7000個實體測試實例、或8000個實體測試實例。參考圖5，此意謂在一些實施例中，Q為100或更大、250或更大、400或更大、800或更大、1000或更大、2000或更大、3000或更大、4000或更大、5000或更大、6000或更大、7000或更大、或8000或更大。換言之，對於該組微擾中之各個各別微擾，多於100個實體測試實例、多於250個實體測試實例、多於400個實體測試實例、多於800個實體測試實例、多於1000個實體測試實例、多於2000個實體測試實例、多於3000個實體測試實例、多於4000個實體測試實例、多於5000個實體測試實例、多於6000個實體測試實例、多於7000個實體測試實例、或多於8000個實體測試實例暴露於各別微擾，且在此類微擾暴露之後自此等實體量測各特徵。

【0127】因此，在測試實體為細胞的一些此類實施例中，該複數個實體測試實例210包含100個細胞測試實例、250個細胞測試實例、400個細胞測試實例、800個細胞測

試實例、1000個細胞測試實例、2000個細胞測試實例、3000個細胞測試實例、4000個細胞測試實例、5000個細胞測試實例、6000個細胞測試實例、7000個細胞測試實例、或8000個細胞測試實例。

【0128】方塊416。參考圖4B之方塊416，在一些實施例中，在獲得特徵之量測值之前，該複數個實體測試實例210暴露於各別測試微擾達至少五分鐘、至少20分鐘、至少30分鐘、至少45分鐘、至少一小時、至少五小時、至少10小時、至少12小時、或至少24小時。在一些實施例中，在獲得特徵之量測值之前，該複數個實體測試實例210暴露於各別測試微擾在五分鐘與兩小時之間。在一些實施例中，在獲得特徵之量測值之前，該複數個實體測試實例210暴露於各別測試微擾在30分鐘與五小時之間。在一些實施例中，在獲得特徵之量測值之前，該複數個實體測試實例210暴露於各別測試微擾在一小時與30小時之間。

【0129】方塊418。參考圖4B之方塊418，在一些實施例中，該複數個測試元素208由5個測試元素與10,000個測試元素之間、100個測試元素與5,000個測試元素之間、500個測試元素與20,000個測試元素之間、25個測試元素與8,000個測試元素之間、15個測試元素與10,000個測試元素之間、250個測試元素與7,000個測試元素之間、100個測試元素與20,000個測試元素之間、或100個測試元素與2000個測試元素之間組成。

【0130】方塊420。參考圖4B之方塊420，在該複數個

測試元素中之各測試元素 208 包含使用在複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量的一些實施例中，該複數個維度縮減成分中之各維度縮減成分為藉由主成分分析 (PCA) 所導出之主成分。

【0131】參考圖 5，PCA 藉由將該複數個特徵 208 轉換為概述訓練集之特徵的一組新變項 (主成分) 來縮減所觀察之資料之維度。參見例如 Jolliffe, 1986, *Principal Component Analysis*, Springer, New York，其特此以引用方式併入。PCA 亦描述於 Draghici, 2003, *Data Analysis Tools for DNA Microarrays*, Chapman & Hall/CRC，其特此以引用方式併入。主成分 (PC) 為不相關的且經定序使得在特徵之觀察資料中第 k 個 PC 具有在 PC 之間第 k 大的變異數。第 k 個 PC 可解釋為使資料點之投影之變化最大以使得其正交於前 $k-1$ 個 PC 的方向。前數個 PC 捕獲觀測資料之大部分變化。相比之下，經常假定後數個 PC 僅捕獲觀察資料之其餘「雜訊」。因此，自 PCA 導出之主成分可用作根據本揭露使用的向量之基礎。在此一方法中，圖 6 中之各方塊 602 表示具體測試微擾之選擇特徵之量測且可視為向量。因此，圖 6 可被視為向量矩陣，各向量表示各別微擾且包括自各別測試實體量測之特徵之量測值。

【0132】在該複數個測試元素中之各測試元素 208 包含使用在複數個實體測試實例暴露於各別測試微擾之後該

複數個實體測試實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量的一些實施例中，維度縮減成分係藉由亞組選擇過程導出。亞組選擇方法之實例揭示於Hastie等人，2001, *The Elements of Statistical Learning*, 55-58，其特此以引用方式併入。

【0133】在該複數個測試元素中之各測試元素208包含使用在複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量的一些實施例中，維度縮減成分係藉由緊縮方法、使用一組控制微擾來導出。與在亞組選擇之情況一樣，不是捨棄第一特徵，而是緊縮方法對其係數大小施加懲罰。緊縮方法之實例揭示於Hastie等人，2001, *The Elements of Statistical Learning*, 59-65，其包括lasso方法，其特此以引用方式併入。

【0134】方塊422。參考圖4B之方塊422，在一些實施例中，該複數個特徵中之各特徵為經光學量測之光學特徵。例如，在一些實施例中，特徵提取自使用落射螢光顯微鏡(例如，ImageXpress Micro, Molecular Devices)捕獲之數位影像106。

【0135】在一些實施例中，各個此類特徵為使用電腦視覺技術量測之細胞形態特徵。例如，在一些實施例中，各特徵為使用細胞繪畫技術識別之光學特徵。在此類實施例中，此類特徵之定量資料提取自實體之顯微術影像以基

於此等輪廓識別出實體樣本之間生物學相關的相似性及差異，其中此類實體已暴露於相同或不同微擾。因此，細胞繪畫為複用複數種(例如，六種)螢光染料(在複數個(例如，五個)通道中成像)來顯示複數個(例如，八個)廣泛相關的細胞成分或胞器的形態剖析檢定。將實體鋪板於多孔盤中，用欲測試之微擾進行微擾，染色，固定並在高通量顯微鏡上成像。接著，自動化影像分析軟體識別出個別實體並量測約1,500個形態特徵(例如，大小、形狀、紋理、強度等之各種量度)以產生合適於偵測細微表型的豐富輪廓。參見Bray等人，2016，「Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes」，Nature Protocols 11(9), 1757-1774，其特此以引用方式併入。

【0136】在一些實施例中，使用深度學習來識別特徵。例如，在一些實施例中，對複數個各自暴露於相同或不同微擾的實體進行成像(例如，在複數個不同時間點)，且對於成像實體之各個各別實例(例如，在該複數個時間點中之不同時間點)，取得影像內覆蓋實體之質心體的影像補片(image patch)(例如，二維數位影像內之 27×27 個像素)(例如，在各別時間點)。各個此類影像補片基於與該影像補片相關聯之一或多個量測螢光訊號標記為具體狀態，從而產生針對複數個實體的大量經標記影像補片。在一些實施例中，然後對此等經標記影像補片進行深度學習，諸如卷積神經網路與遞歸神經網路(RNN)組合之架

構，其自動提取局部影像特徵並利用影像補片之資訊(例如，時間資訊、螢光隨微擾及/或時間之變化等)。例如，在一些實施例中，深度學習包含提取影像特徵的三個連接卷積層。參見例如 Buggenthin 等人，2017，「Prospective identification of hematopoietic lineage choice by deep learning」, Nature Methods, 14(4), 403-406，其特此以引用方式併入。

【0137】方塊424。參考圖4B之方塊424，在一些實施例中，第一亞組該複數個特徵為經光學量測之光學特徵。第二亞組該複數個特徵為非光學特徵。

【0138】方塊426。參考圖4B之方塊426，在一些實施例中，該複數個特徵中之各特徵為經光學量測之特徵。例如，在一些實施例中，使用聲共振、電阻抗、微懸臂、奈米線、或微差量熱法量測特徵。參見 Cooper, 2006, 「Non-optical screening platforms: the next wave in label-free screening?」 Drug Discovery Today 11, 1068-1074，其特此以引用方式併入。例如，ACEA Biosciences (<http://www.aceabio.com>)發表一種即時細胞電子感測(RT-CESTM)系統，其係基於整合至標準生物分子科學學會(Society for Biomolecular Sciences) (SBS；<http://www.sbsonline.org>)形式微量盤之底部中的微電子細胞感測器陣列。RT-CES藉由量測感測器上之電阻抗起作用以偵測實體之存在、不存在、或條件變化。對於基於實體之檢定，使實體生長於微量盤之個別、含感測器之孔中並放置於標

準孵育器中。系統可經程式化以藉由在使用者定義之間隔下通過感測器發送標稱電流來以每分鐘之頻率收集特徵資料。電子感測器提供關於阻抗值之資訊，其可用作直接使用本揭露之系統及方法的特徵值或轉化成稱為細胞指數(針對單獨媒介物之阻抗正規化之細胞阻抗)的量度，其繼而可用作本揭露中之特徵。

【0139】可在本揭露中使用的非光學特徵之另一來源可獲自 Applied BioPhysics (<http://www.biophysics.com>) 玻片，其具有八個或96個個別孔以供細胞培養。該裝置之基部具有金膜電極陣列，該等電極將電子細胞-基質阻抗感測(electric cell-substrate impedance sensing, ECIS)電子設備連接至各孔。可用此方法量測範圍為重合流層至稀疏層的細胞密度。電極大小限制可觀察到的錨定細胞之最大數目(通常為100至1000個細胞)。然而，藉由並聯使用多個電極，可覆蓋孔中更大的表面積以量測多達4000個細胞。偵測電子設備之靈敏度足以偵測甚至單一單離細胞反應。以此方式，對於實體為細胞的應用，可獲得表示回應於微擾之細胞行為的非光學特徵，包括細胞增殖、障壁功能、貼附與伸展、遷移、及在靜態及流動條件下的侵入。此外，可使用較高電場自傷口癒合檢定量測用於本揭露中之非光學特徵，且若應用達較短時間，則可用於電穿孔細胞並監測不可滲透膜的分子諸如 dsRNA 之後續進入。該技術亦可用於獲得呈訊號轉導、轉移潛能、及體外毒性之形式的非光學特徵。

【0140】方塊428。參考圖4B之方塊428，在一些實施例中，對於該複數個實體測試實例中之各測試實例，個別地量測該複數個不同特徵中之不同特徵。換言之，自單一實體實例獲得量化特徵之具體實例所需的所有資料。例如，在實體為細胞且特徵為細胞面積的情況下，自單一細胞實例獲得量化細胞之細胞面積所需的所有資料。為了獲得可靠的量測值，且如圖2中所繪示，量測此特徵之多個測試實體實例，然而在根據方塊428之實施例中，自各實體獨立地量測特徵實例。在一些實施例中，如下文參考方塊432所述，將針對測試向量之元素的特徵的此等個別量測值組合成分佈度量。

【0141】方塊430。參考圖4B之方塊430，在一些實施例中，在該複數個實體測試實例之至少一亞組實例中量測該複數個不同特徵中之不同特徵。換言之，在此類實施例中，自多個實體實例，而不是自單一實體獲得量化特徵之具體實例所需的所有資料。例如，在實體為細胞且特徵為肌動蛋白含量的情況下，可在多個細胞實例(例如，多孔盤之具體孔中的所有細胞)中獲得量化細胞之細胞面積所需的所有資料。為了獲得可靠的量測值，且如圖2中所繪示，量測此特徵之多個測試實體實例，然而在根據方塊430之實施例中，在複數個分開的實體中獨立地量測特徵之各實例。在一些實施例中，如下文參考方塊432所述，然後將針對測試向量之元素的特徵的此等個別量測值組合成分佈度量。

【0142】方塊432-434。參考圖4C之方塊432，應注意，暴露於微擾之後的實體測試實例形成測量特徵的基礎。因此，對於各微擾，量測特徵之多個實例，且特徵之此等實例形成測試向量之元素或對形成測試向量之元素的維度縮減成分有貢獻。例如，在一些實施例中，針對此類實體暴露於具體測試微擾之後的具體特徵，量測二或更多個、三或更多個、四或更多個、五或更多個、六或更多個、3與10之間個、2與100之間個、七或更多個、或八或更多個實體測試實例。以此方式，藉由多次測量測特徵，獲得呈特徵之分佈度量之形式的各特徵之更可靠的量測值。因此，在一些實施例中，在該複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例210中不同特徵之量測值之分佈度量為在該複數個實體測試實例暴露於各別測試微擾之後該複數個實體測試實例中不同特徵之多樣性量度。

【0143】在一些此類實施例中，參考圖4C之方塊434，不同特徵在複數個特徵中的多樣性量度為特徵之範圍、標準偏差、或變異數，如針對該複數個測試實例中各實體測試實例所量測。同樣，在對應於具體測試微擾的測試向量之各元素為使用在複數個實體測試實例暴露於具體測試微擾之後該複數個實體測試實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量的實施例中，各別維度縮減成分之多樣性量度為不同維度縮減成分之範圍、標準偏差、或變異數，如使

用對在暴露於具體測試微擾的複數個測試實例中各實體測試實例之不同維度縮減成分有貢獻的成分特徵所計算。

【0144】方塊436-438。參考圖4C之方塊436，使用該複數個測試向量206計算複合測試度量212。複合測試度量212量化該組測試微擾對第一成分的中靶效應。在一些實施例中，藉由第一過程計算複合測試度量，對於該複數個測試向量中之各個各別測試向量206，該第一過程包含計算(i)各別測試向量206與(ii)在自該複數個測試向量中除去該各別測試向量之情況下，該複數個測試向量之分佈度量之間的測試相似性度量214，從而獲得該複數個測試向量之複數個測試相似性度量。該複數個測試相似性度量中之各測試相似性度量214獨特地對應於第一組微擾中之微擾。因此，為了計算第一測試向量(其表示一組測試微擾中之測試微擾)之測試相似性度量214，將第一測試向量之第一元素與在所有其他測試向量(不包括第一測試向量)之第一元素中所計算的分佈度量相比較，將第一測試向量之第二元素與在所有其他測試向量(不包括第一測試向量)之第二元素中所結算的分佈度量相比較，等等。換言之，計算各別測試向量與參考測試向量之間的距離，其中參考測試向量之各個各別元素為在不包括該各別測試向量的該複數個測試向量中該各別元素之分佈度量。所以，對於第一測試向量，計算第一測試向量與參考測試向量之間的距離，其中參考測試向量之各個各別元素為在不包括該第一測試向量的該複數個測試向量中該各別元素之分佈度量。

參考圖 4C 之方塊 438，在一些實施例中，分佈度量為在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量 206 之集中趨勢量度(算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數)。

【0145】方塊 440-442。參考圖 4C 之方塊 440，在一些實施例中，計算 (i) 各別測試向量 206 與 (ii) 在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量之分佈度量(例如，集中趨勢量度)之間的測試相似性度量 214，作為測試向量 206 之對應元素 208 與在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量之分佈度量之間的距離。例如，在一些實施例中，距離為角距。參考方塊 442，在一些實施例中，距離為例如如下計算之角距：

$$\frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

其中 A_i 為各別測試向量中之測試元素 i 208， B_i 為除各別測試向量之外的該複數個測試向量 206 中對應測試元素 i 在該複數個測試元素 208 中之分佈度量，且 n 為各別測試向量中元素之數目(例如，如圖 2 中所繪示之 S 個元素)。

【0146】可用於計算 (i) 各別測試向量 206 與 (ii) 在自該複數個測試向量中除去該各別測試向量之情況下，該複數

個測試向量之分佈度量之間的距離的其他相似性量度討論於Duda及Hart, *Pattern Classification and Scene Analysis*, 1973, John Wiley & Sons, Inc., New York, (後文為「Duda 1973」)之第6.7節。在一些實施例中，不是使用相似性度量計算(i)各別測試向量206與(ii)在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量之分佈度量之間的距離，而是使用非度量相似性函數 $s(x, x')$ 來比較兩個向量 x (各別測試向量206)及 x' (其中向量之各個各別元素為在自該複數個測試向量除去各別測試向量之情況下該複數個測試向量中各別元素之分佈度量的向量)。習知， $s(x, x')$ 為對稱函數，當 x 及 x' 以某種方式「相似」時，其值為大的。非度量相似性函數 $s(x, x')$ 之實例提供於Duda 1973第216頁。出於本揭露之目的，認為術語「測試相似性度量」涵蓋相似性度量，諸如方塊442之角距函數，以及諸如Duda 1973中所述之非相似性度量。

【0147】方塊444。參考圖4D之方塊444，圖4C之方塊436中所指出之第一過程進一步包含計算複合測試度量212，作為該複數個測試相似性度量214之分佈度量，諸如集中趨勢量度(例如，算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數)。以此方式，計算複合測試度量436之值，其表示整組測試微擾之效應。

【0148】方塊446。參考圖4E之方塊444，在一些實施例中，該複數個測試相似性度量214之該集中趨勢量度為

該複數個測試相似性度量之算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數。

【0149】方塊448。參考圖4D之方塊448，計算包含複數個複合控制度量218的虛無分佈216，以計算複合測試度量。儘管最終用於計算複合測試度量212的各測試向量係基於目標第一微擾之微擾，但是虛無分佈係建立於旨在靶向考慮中之實體之目標隨機成分之微擾上。以此方式，虛無分佈提供為確定複合測試度量212之計算值之相關性提供統計學基礎。

【0150】在一些實施例中，藉由第二過程計算該複數個複合控制度量中之各個各別複合控制度量218，該第二過程包含自複數個微擾選擇一組各別控制微擾220。該組各別控制微擾220中之各控制微擾222係針對該複數個成分中之不同成分224。

【0151】方塊450。參考圖4D之方塊450，在一些實施例中，該複數個複合控制度量218包含50個複合控制度量218、100個複合控制度量218、500個複合控制度量218、1000個複合控制度量、或2000個複合控制度量，其中各複合控制度量表示與該複數個微擾不同的一組控制微擾。

【0152】方塊452。參考圖4D之方塊452，在一些實施例中，該複數個複合控制度量中具體複合控制度量218之該組各別控制微擾220由2與12之間個不同控制微擾222、3與10之間個不同控制微擾222、5與15之間個不同控制微擾222、6與60之間個不同控制微擾、或3與300之間個不同控

制微擾222組成。在一些實施例中，控制微擾為特異性識別實體成分的小干擾RNA (siRNA)。各siRNA為雙股RNA分子，長度為20-25個鹼基對，其藉由在轉錄之後降解mRNA從而阻止基因之轉譯來干擾具有互補核苷酸序列之特定基因之表現。參見Agrawal等人，2003，「RNA interference: biology, mechanism, and applications」，Microbiol Mol Biol Rev. 67: 657，其特此以引用方式併入。

【0153】方塊454。參考圖4E之方塊454，在一些實施例中，第二過程進一步包含針對該組各別控制微擾220中之各個各別控制微擾222，獲得對應控制向量226，從而獲得複數個各別控制向量。各對應控制向量226包含複數個控制元素228。

【0154】在一些實施例中，該複數個控制元素228中之各控制元素228包含在複數個實體控制實例暴露於各別控制微擾222之後該複數個實體控制實例230中不同特徵在複數個特徵中之量測值之分佈度量。

【0155】在一些實施例中，此類資料係使用自動化細胞成像系統(例如ImageXpress Micro, Molecular Devices)獲得，其中在將實體經這一組發射不同離散波長的染料(例如，赫斯特33342、Alexa Fluor 594鬼筆環肽等)染色及控制微擾之後，將其佈置於多孔盤(例如，384孔盤)中。在一些實施例中，在20×放大率、2×分箱下，用藉由所用之標記物染料所確定之暴露(例如，赫斯特為15 ms，鬼筆環

肽為1000 ms)對實體進行成像。對於各孔，在一些實施例中，使用自對焦於特定染料通道(例如，赫斯特通道)上的雷射來發現最佳焦點。在一些實施例中，然後將自動化顯微鏡程式化以收集32個影像之z堆疊(z = 0位於最焦平面，16個影像高於焦平面，16個影像低於焦平面)，其中斷片(slice)之間為2 μm 。在一些實施例中，各孔在其中含有數千個實體，且因此藉由照相機捕獲之孔之各數位表示表示若干不同孔之各者中的數千實體。在一些實施例中，使用分段軟體來識別出數位影像中之個別實體，且此外識別出個別實體內的各種成分(例如，細胞成分)。一旦分段並識別細胞成分，便對此等成分進行數學轉換以獲得特徵之量測值。

【0156】因此，為了說明，在給定複合控制度量218之一組控制微擾220為五個控制微擾222的情況下，將存在五個控制向量226，即，給定複合控制度量之該組五個控制微擾220中各控制微擾之控制向量226。此外，各控制向量226將包含複數個控制元素228(例如，五或更多個控制元素228、十或更多個控制元素228、二十或更多個控制元素228、100或更多個控制元素228、或一千或更多個控制元素228)。控制向量226中之各控制元素228表示在暴露於控制微擾之後從控制實體230之控制實例量測的複數個特徵中的不同特徵。在典型實施例中，針對各控制向量226量測之該組特徵為與針對各測試向量206量測之特徵相同的特徵。

【0157】為了進一步說明，考慮第一控制向量226對應於第一控制微擾222的情況，其中第一控制向量226包含20個控制元素228。各控制元素228表示在暴露於第一次控制微擾222之後從實體控制實例量測的不同特徵。例如，第一控制向量226中之第一控制元素228表示在暴露於第一微擾之後複數個實體控制實例230中之各者中第一特徵之量測值。更特定言之，各控制元素228表示在暴露於對應於含有控制元素228之控制向量226的控制微擾222之後從控制實體之各控制實例230量測的對應特徵之分佈度量。

【0158】在典型實施例中，暴露於測試微擾以形成測試向量206的實體為與暴露於控制微擾222以形成控制向量226的實體類型相同的實體。

【0159】對於複合控制度量218而言，在複數個實體控制實例暴露於該複數個微擾中之各別控制微擾222之後在該複數個實體控制實例中該複數個特徵之量測值得到N維空間，其中該N維空間中之各整數為該複數個特徵中之不同特徵。

【0160】在一些替代實施例中，控制向量226之該複數個控制元素中之各控制元素228包含各別維度縮減成分在複數個維度縮減成分中之分佈度量，其中如方塊408中所述，使用在複數個實體測試實例暴露於各別控制微擾之後該複數個實體控制實例中該複數個特徵之量測值識別維度縮減成分。此類實施例為有利的，因為其消除或減少實體之高度相關的特徵(如大小、核面積、及核周長)之間的

冗餘。因此，在此類實施例中，使用在複數個實體控制實例暴露於控制微擾之後該複數個實體控制實例 230 中之該複數個特徵 226 (N 個特徵) 填充控制向量 226 之給定控制元素 228 之該複數個維度縮減成分。因此，在該組特徵為一千個特徵之情況下，控制向量中將有少於五百個維度縮減成分。在此類情況下，各個各別控制向量 226 將表示複合控制度量 218 之該組控制微擾 220 中之對照微擾 222，且基於在已暴露於對應於各別控制向量 226 的控制微擾 222 的控制實體中維度縮減成分所表示之特徵之值，各個各別控制向量 226 之各個各別控制元素 228 將包括對應維度縮減成分之值。

【0161】方塊 456。參考圖 4E 之方塊 456，在一些實施例中，該複數個各別實體控制實例 230 包含 100 個實體控制實例、250 個實體控制實例、400 個實體控制實例、800 個實體控制實例、1000 個實體控制實例、2000 個實體控制實例、3000 個實體控制實例、4000 個實體控制實例、5000 個實體控制實例、6000 個實體控制實例、7000 個實體控制實例、或 8000 個實體控制實例。換言之，對於一組控制微擾 220 中之各個各別控制微擾 222，多於 100 個實體控制實例、多於 250 個實體控制實例、多於 400 個實體控制實例、多於 800 個實體控制實例、多於 1000 個實體控制實例、多於 2000 個實體控制實例、多於 3000 個實體控制實例、多於 4000 個實體控制實例、多於 5000 個實體控制實例、多於 6000 個實體控制實例、多於 7000 個實體控制實例、或多於

8000個實體控制實例暴露於各別控制微擾，且在此類微擾暴露之後自此等實體量測各特徵。

【0162】因此，在控制實體為細胞的一些此類實施例中，該複數個實體控制實例230包含100個細胞控制實例、250個細胞控制實例、400個細胞控制實例、800個細胞控制實例、1000個細胞控制實例、2000個細胞控制實例、3000個細胞控制實例、4000個細胞控制實例、5000個細胞控制實例、6000個細胞控制實例、7000個細胞控制實例、或8000個細胞控制實例。

【0163】方塊458。參考圖4E之方塊458，在較佳實施例中，該複數個各別實體控制實例230暴露於各別控制微擾222達與方塊416中測試實例210暴露於測試微擾之時間量相同的時間。據此，在一些實施例中，在獲得特徵之量測值之前，該複數個實體控制實例230暴露於各別控制微擾222達至少五分鐘、至少20分鐘、至少30分鐘、至少45分鐘、至少一小時、至少五小時、至少10小時、至少12小時、或至少24小時。在一些實施例中，在獲得特徵之量測值之前，該複數個實體控制實例230暴露於各別控制微擾222在五分鐘與兩小時之間。在一些實施例中，在獲得特徵之量測值之前，該複數個實體控制實例230暴露於各別控制微擾222在30分鐘與五小時之間。在一些實施例中，在獲得特徵之量測值之前，該複數個實體控制實例230暴露於各別控制微擾222在一小時與30小時之間。

【0164】方塊460-461。應注意，在暴露於微擾之後

實體之控制實例形成量測各複合控制度量218之特徵的基礎。因此，對於各控制微擾222，量測特徵之多個實例，且特徵之此等實例直接形成控制向量226之元素228或對形成測控制向量的維度縮減成分有貢獻。例如，在一些實施例中，針對此類實體暴露於具體控制微擾之後的具體特徵，量測二或更多個、三或更多個、四或更多個、五或更多個、六或更多個、3與10之間個、2與100之間個、七或更多個、或八或更多個實體測試實例。以此方式，藉由多次測量測特徵，獲得呈特徵之分佈度量之形式的各特徵之更可靠的量測值。在一些實施例中，分佈度量為在複數個實體控制實例230暴露於各別控制微擾222之後在該複數個實體控制實例中不同特徵之算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數。

【0165】在一些實施例中，參考圖4E之方塊461，在該複數個實體控制實例暴露於各別控制微擾222之後該複數個實體控制實例230中不同特徵之量測值為在該複數個實體控制實例暴露於各別控制微擾之後該複數個實體控制實例中不同特徵之多樣性量度。在一些實施例中，此多樣性量度為特徵之範圍、標準偏差、或變異數，如針對具體控制微擾222的在該複數個控制實例中之各實體控制實例230所量測。同樣，在對應於具體控制微擾222的控制向量226之各元素228為使用在暴露於具體控制微擾之後複數個實體控制實例中該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之分佈度量的實施例

中，多樣性量度為不同維度縮減成分之範圍、標準偏差、或變異數，如使用對在暴露於具體控制微擾的複數個控制實例中各實體控制實例之不同維度縮減成分有貢獻的成分特徵所計算。

【0166】方塊462-464。參考圖4F之方塊462，在一些實施例中，該複數個測試元素208由5個測試元素與10,000個測試元素之間組成，該複數個控制元素228由5個控制元素與10,000個控制元素之間組成，且該複數個測試元素中之各測試元素208與該複數個控制元素中之對應控制元素228之間存在一對一對應。即，在此類實施例中，各測試向量具有與各控制向量相同數目的元素。例如，在此類實施例中，若各測試向量中有100個測試元素，則各控制向量中有100個控制元素。然而，測試元素為使用影響實體之第一成分的測試微擾的量測結果，而控制向量為使用影響實體之除第一成分之外的成分的控制微擾的量測結果。參考圖4F之方塊464，在一些實施例中，該複數個測試元素208由100個測試元素與2000個測試元素之間組成，該複數個控制元素228由100個控制元素與2000個控制元素之間組成，且該複數個測試元素中之各測試元素208與該複數個控制元素中之對應控制元素228之間存在一對一對應。

【0167】方塊466-468。參考圖4F之方塊466，第二程序進一步包含使用如圖3中所繪示對應於控制度量218的該複數個各別控制向量226計算各別複合控制度量218。各別複合控制度量218量化該組各別控制微擾220對與該複合控

制度量相關聯的一組控制微擾 220 所靶向之對應不同成分 224 的中靶效應。

【0168】 在一些實施例中，藉由第三過程計算各別複合控制度量 218，對於各別複合控制度量 218 之各個各別控制向量 226，該第三過程包含計算 (i) 各別控制向量 226 與 (ii) 在自該複數個各別控制向量中除去該各別控制向量 226 之情況下，該複數個各別控制向量之分佈度量之間的控制相似性度量 219，從而獲得該複數個各別控制向量之複數個控制相似性度量 219。該複數個控制相似性度量中之各控制相似性度量 219 獨特地對應於該組各別控制微擾 220 中之微擾 222。因此，為了計算第一控制向量 (其表示針對各別複合控制度量 218 的一組控制微擾 220 中之控制微擾 222) 之控制相似性度量 219，將第一控制向量 226 之第一元素 228 與在所有與各別複合控制度量 218 相關聯的其他控制向量 (不包括第一控制向量) 之第一元素中所計算的分佈度量相比較，將第一控制向量之第二元素與在所有與各別複合控制度量 228 相關聯的其他控制向量 (不包括第一控制向量) 之第二元素中所結算的分佈度量相比較，等等。換言之，計算各別控制向量與參考控制向量之間的距離，其中參考控制向量之各個各別元素為在不包括各別控制向量的與各別複合控制度量 218 相關聯的該複數個控制向量中該各別元素之分佈度量。所以，對於第一控制向量，計算第一控制向量與參考控制向量之間的距離，其中參考控制向量之各個各別元素為在不包括第一控制向量的與各別複合

控制度量 218 相關聯的該複數個控制向量中該各別元素之分佈度量。參考圖 4F 之方塊 468，在一些實施例中，分佈度量為在自複數個控制向量除去各別控制向量 226 之情況下在與各別控制向量 218 相關聯的除各別控制向量之外的複數個各別控制向量 226 中各對應控制元素 228 在複數個各別控制元素中的集中趨勢量度 (算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數)。

【0169】方塊 470-472。參考圖 4F 之方塊 470，在一些實施例中，計算 (i) 各別控制向量 226 與 (ii) 在自複數個各別控制向量除去各別控制向量之情況下與各別複合控制度量 218 相關聯的複數個各別控制向量之分佈度量之間的控制相似性度量 219，作為各別控制向量之對應元素與在自複數個各別控制向量除去各別控制向量之情況下複數個各別控制向量之分佈度量之間的距離。在一些實施例中，距離為角距。例如，參考圖 4G 之方塊 472，在一些實施例中，距離為如下計算之角距：

$$\frac{\sum_i^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

其中 A_i 為各別控制向量中之控制元素 i ， B_i 為除各別控制向量之外的該複數個各別控制向量中對應控制元素 i 在該複數個控制元素中之分佈度量，且 n 為各別控制向量中

元素之數目。

【0170】可用於計算(i)各別控制向量226與(ii)在除去各別控制向量之情況下，該複數個控制向量之分佈度量之間的距離的其他相似性量度討論於Duda及Hart, *Pattern Classification and Scene Analysis*, 1973, John Wiley & Sons, Inc., New York, (後文為「Duda 1973」)之第6.7節。在一些實施例中，不是使用相似性度量計算(i)各別控制向量226與(ii)在除去各別測試向量之情況下該複數個測試向量之分佈度量之間的距離，而是使用非度量相似性函數 $s(x, x')$ 來比較兩個向量 x (各別控制向量226)及 x' (其中向量之各個各別元素為在除去各別控制向量之情況下該複數個控制向量中各別元素之分佈度量的向量)。習知， $s(x, x')$ 為對稱函數，當 x 及 x' 以某種方式「相似」時，其值為大的。非度量相似性函數 $s(x, x')$ 之實例提供於Duda 1973第216頁。出於本揭露之目的，認為術語「測試相似性度量」涵蓋相似性度量，諸如方塊472之角距函數，以及諸如Duda 1973中所述之非相似性度量。

【0171】方塊474。參考圖4G之方塊474，在一些實施例中，第三過程進一步包含計算各別複合控制度量218，作為與各別複合控制度量相關聯之該複數個控制測試相似性度量之分佈度量，諸如集中趨勢量度(例如，算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數)。

【0172】方塊476-478。參考圖4G之方塊476，在一些

實施例中，將複合測試度量212與虛無分佈216相比較。當複合度量212不同於虛無分佈時，認為該組測試微擾相對於虛無分佈區別對第一成分的中靶效應。參考圖4G之方塊478，在一些實施例中，複合測試度量212與虛無分佈216之比較包含使用虛無分佈216計算複合測試度量之Z分數。當複合測試度量之Z分數超過臨限值(例如，2.5、6等)時，認為複合測試度量不同於虛無分佈。z分數為自平均值之標準偏差數，資料點為平均值。因此，其為複合測試度量212超過群體平均值的標準偏差數之量度，如藉由虛無分佈216所表示。z分數亦稱為標準分數，且其可放置於常態分佈曲線上。為了計算z分數，計算複合控制度量之平均值 μ 及複合控制度量218之標準偏差 σ 。在一些實施例中，複合測試度量之z分數公式為：

$$z = (x - \mu) / \sigma$$

其中x為複合測試度量， μ 為複數個複合測試度量218之平均值，且 σ 為複數個複合控制度量218之標準偏差。

引用及替代實施例之參考

【0173】本文引用之所有參考以全文引用方式且出於所有目的併入本文，其程度如同明確且個別地指示欲出於所有目的將各個別公開案或專利或專利申請案以全文引用之方式併入一般。

【0174】本發明可實施為電腦程式產品，其包含嵌入非暫態電腦可讀存儲媒體中的電腦程式機制。例如，電腦程式產品和含有如1、2、3之任一組合中所示及/或圖4中所述之程式模組。此等程式模組可存儲於CD-ROM、DVD、磁碟儲存產品、USB密鑰、或任何其他非暫態電腦可讀資料或程式存儲產品。

【0175】如熟習此項技術者將顯而易見，可在不背離本發明之精神及範疇之情況下對本發明進行許多修改及變化。本文所述之特定實施例僅藉由實例來提供。可對實施例進行選擇及描述以最好地說明本發明之原理及其實際應用，從而使得其他熟習此項技術者能夠在合適於所實施之具體用途時用各種修改最好地利用本發明及各種實施例。本發明僅藉由隨附申請專利範圍之各項以及此類申請專利範圍享有權利的等同物之全部範疇進行限制。

【符號說明】

【0176】

102：多孔盤

104：落射螢光顯微鏡

106：影像/通訊網路/數位影像

108：特徵

110：資料集

192：記憶體

202：作業系統

- 204：中靶效應確定模組
- 206：測試向量/第一測試向量/向量
- 208：測試元素/第一測試元素/特徵
- 210：測試實例
- 212：複合測試度量
- 213：通信匯流排
- 214：測試相似性度量
- 216：虛無分佈
- 218：複合控制度量
- 219：控制相似性度量
- 220：一組控制微擾
- 222：控制微擾
- 224：成分
- 226：控制向量
- 228：控制元素
- 230：控制實例
- 250：系統/數位影像/裝置/電腦系統
- 274：處理單元/中央處理單元/處理器
- 276：電源
- 278：使用者介面
- 280：輸入
- 282：顯示器
- 284：網路或其他通信介面
- 288：控制器

290：磁碟存儲及/或持續裝置/記憶體

292：元件

402：方塊

404：方塊

406：方塊

408：方塊

410：方塊

412：方塊

414：方塊

416：方塊

418：方塊

420：方塊

422：方塊

424：方塊

426：方塊

428：方塊

430：方塊

432：方塊

434：方塊

436：方塊

438：方塊

440：方塊

442：方塊

444：方塊

446：方塊

450：方塊

452：方塊

454：方塊

456：方塊

458：方塊

460：方塊

461：方塊

462：方塊

464：方塊

448：方塊

474：方塊

466：方塊

468：方塊

470：方塊

472：方塊

476：方塊

478：方塊

602：方塊



202000919

【發明摘要】

【中文發明名稱】

用於區別對靶標之效應的系統及方法

【英文發明名稱】

SYSTEMS AND METHODS FOR DISCRIMINATING EFFECTS ON
TARGETS

【中文】

揭示用於確定一組測試微擾是否相對於虛無分佈區別對實體之第一成分的中靶效應的系統及方法。該等微擾為該第一成分之微擾，且該實體包含複數個成分。對於該組中之各微擾，獲得包含複數個元素的對應向量。各元素包含在暴露於該各別微擾之後該實體實例中特徵之量測值之分佈度量，或(ii)使用在微擾暴露之後該實體實例之該複數個特徵之該量測值所計算的各別維度縮減成分之分佈度量。使用該等向量計算複合度量，並與虛無分佈相比較。當該複合度量不同於該虛無分佈時，認為該組微擾相對於該虛無分佈區別對該第一成分的該中靶效應。

【 英文 】

Systems and methods for determining whether a set of test perturbations discriminates over a null distribution for an on target effect against a first component of an entity are disclosed. The perturbations are perturbations of the first component and the entity comprises a plurality of components. For each perturbation in the set, a corresponding vector comprising a plurality of elements, is obtained. Each element comprises a distribution metric of measurements of a feature across instances of the entity upon exposure to the respective perturbation or (ii) a distribution metric of a respective dimension reduction component computed using the measurement of the plurality of features across instances of the entity upon the perturbation exposure. A composite metric is computed, using the vectors, and compared to a null distribution. When the composite metric is differentiated from the null distribution, the set of perturbations is deemed to discriminate the on target effect against the first component over the null distribution.

【指定代表圖】第(2)圖。

【代表圖之符號簡單說明】

- 192：記憶體
- 202：作業系統
- 204：中靶效應確定模組
- 206：測試向量/第一測試向量/向量
- 208：測試元素/第一測試元素/特徵
- 210：測試實例
- 212：複合測試度量
- 213：通信匯流排
- 214：測試相似性度量
- 216：虛無分佈
- 218：複合控制度量
- 219：控制相似性度量
- 220：一組控制微擾
- 222：控制微擾
- 224：成分
- 226：控制向量
- 228：控制元素
- 230：控制實例
- 250：系統/數位影像/裝置/電腦系統
- 274：處理單元/中央處理單元/處理器
- 276：電源
- 278：使用者介面
- 280：輸入

282：顯示器

284：網路或其他通信介面

288：控制器

290：磁碟存儲及/或持續裝置/記憶體

292：元件

【特徵化學式】無

【發明申請專利範圍】

【第1項】

一種用於確定一組測試微擾是否相對於虛無分佈區別對細胞之一或多個第一遺傳成分的中靶效應的方法，其中該組測試微擾包含複數個對該一或多個第一遺傳成分的測試微擾，且該細胞包含複數個遺傳成分，包括該一或多個第一遺傳成分，該方法包含：

在包含處理器及記憶體之電腦系統處：

(A) 針對該組測試微擾中之各個各別測試微擾，獲得對應測試向量，從而獲得複數個測試向量，其中各對應測試向量包含複數個測試元素，該複數個測試元素中之各測試元素包含使用在複數個細胞實例暴露於各別測試微擾之後該複數個細胞實例之複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之測試分佈度量，其中該測試分佈度量之該複數個維度縮減向量小於該複數個特徵；

(B) 藉由第一過程，使用該複數個測試向量，計算複合測試度量，該複合測試度量量化該組測試微擾對該一或多個第一遺傳成分的該中靶效應，該第一過程包含：

(a) 對於該複數個測試向量中之各個各別測試向量，計算(i)該各別測試向量與(ii)在自該複數個測試向量除去該各別測試向量之情況下該複數個測試向量之測試分佈度量之間的相似性度量，從而獲得該複數個測試向量之複數個測試相似性度量，該複數個相似性度量中之各測試相似

性度量表示該第一組微擾中之微擾，且

(b) 自該複數個相似性度量計算該複合測試度量；

(C) 計算該虛無分佈，其中該虛無分佈包含複數個複合控制度量，該複數個複合控制度量中之各個各別複合控制度量係藉由第二過程計算，該第二過程包含：

(a) 自複數個微擾選擇一組各別控制微擾，其中該組各別控制微擾中之各控制微擾係針對該複數個遺傳成分中除該一或多個第一遺傳成分之外的一或多個不同遺傳成分；

(b) 針對該組各別控制微擾中之各個各別控制微擾，獲得對應控制向量，從而獲得複數個各別控制向量，其中各對應控制向量包含複數個控制元素，該複數個控制元素中之各控制元素包含使用在該複數個各別細胞控制實例暴露於該各別控制微擾之後該複數個各別細胞控制實例之該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之控制分佈度量，其中該控制分佈度量之該複數個維度縮減成分小於該複數個特徵；

(c) 使用該複數個各別控制向量計算該各別複合控制度量，其中該各別控制度量量化該組各別控制微擾對該一或多個對應的不同遺傳成分的該中靶效應，且其中該控制度量係藉由第三過程計算，該第三過程包含：

(1) 對於各個各別控制向量，計算(i)該各別控制向量與(ii)在自該複數個各別控制向量除去該各別控制向量之情況下該複數個各別控制向量之分佈度量之間的控制相似

性度量，從而獲得該複數個各別控制向量之複數個控制相似性度量，該複數個控制相似性度量中之各控制相似性度量獨特地對應於該組各別控制微擾中之微擾，且

(2) 計算該各別複合控制度量，作為該複數個控制相似性度量之集中趨勢量度；且

(D) 將該複合測試度量與該虛無分佈相比較，其中當該複合測試度量不同於該虛無分佈時，認為該組測試微擾相對於該虛無分佈區別對該一或多個第一遺傳成分的該中靶效應。

【第2項】

如申請專利範圍第1項之方法，其中該複數個特徵中之特徵表示在細胞暴露於該各別測試微擾或控制微擾之後細胞或細胞之列舉部分之顏色、紋理、或大小。

【第3項】

如申請專利範圍第1項或第2項之方法，其中該組測試微擾由3與300之間個不同測試化學成分組成。

【第4項】

如申請專利範圍第1項或第2項之方法，其中該一或多個第一遺傳成分為第一基因，且該組測試微擾之各測試微擾進一步具有對該細胞中除該第一基因之外的一或多個基因的脫靶效應。

【第5項】

如申請專利範圍第1項或第2項之方法，其中該複數個細胞測試實例包含500個細胞測試實例。

【第6項】

如申請專利範圍第1項或第2項之方法，其中在獲得該量測值之前，該複數個細胞實例暴露於該各別測試微擾達至少一小時。

【第7項】

如申請專利範圍第1項或第2項之方法，其中該複數個測試元素由5個測試元素與10,000個測試元素之間組成。

【第8項】

如申請專利範圍第1項或第2項之方法，其中在自該複數個測試向量除去該各別測試向量之情況下該複數個測試向量之該測試分佈度量為除該各別測試向量之外的該複數個測試向量中各對應測試元素在該複數個測試元素中之集中趨勢量度。

【第9項】

如申請專利範圍第1項或第2項之方法，其中該複數個測試相似性度量之該集中趨勢量度為該複數個測試相似性度量之算術平均數、加權平均數、中程數、中樞紐、三均數、溫塞平均數、中位數、或眾數。

【第10項】

如申請專利範圍第1項或第2項之方法，其中該複數個複合控制度量包含100個複合控制度量，各複合控制度量表示與該複數個微擾不同的控制微擾組合。

【第11項】

如申請專利範圍第1項或第2項之方法，其中該組各別

控制微擾由3與300之間個不同控制微擾組成。

【第12項】

如申請專利範圍第1項或第2項之方法，其中該複數個各別細胞控制實例包含1000個細胞控制實例。

【第13項】

如申請專利範圍第1項或第2項之方法，其中在獲得該量測值之前，該複數個各別細胞控制實例暴露於該各別控制微擾達至少一小時。

【第14項】

如申請專利範圍第1項或第2項之方法，其中在自該複數個各別控制向量除去該各別控制向量之情況下該複數個控制向量之該分佈度量為除該各別控制向量之外的該複數個各別控制向量中各對應控制元素在該複數個各別控制元素中之集中趨勢量度。

【第15項】

如申請專利範圍第1項或第2項之方法，其中計算(i)該各別控制向量與(ii)在自該複數個各別控制向量除去該各別控制向量之情況下該複數個各別控制向量之該分佈度量之間的該控制相似性度量，作為該各別控制向量之對應元素與在自該複數個各別控制向量除去該各別控制向量之情況下該複數個各別控制向量之該分佈度量之間的距離。

【第16項】

如申請專利範圍第1項或第2項之方法，其中該複合測試度量與該虛無分佈之該比較包含使用該虛

無分佈計算該複合測試度量之Z分數，且

當該複合測試度量之Z分數超過臨限值時，認為該複合測試度量不同於該虛無分佈。

【第17項】

如申請專利範圍第1項或第2項之方法，其中
第一亞組該複數個特徵為經光學量測之光學特徵；且
第二亞組該複數個特徵為非光學特徵。

【第18項】

如申請專利範圍第1項或第2項之方法，其中該組測試微擾包括第一小干擾RNA (siRNA)，且一組各別控制微擾包括第二siRNA。

【第19項】

如申請專利範圍第1項或第2項之方法，其中該組測試微擾包括條件培養基，其係以直接自細胞或自一或多個表現出所關注之疾病的患者之體液、組織、或器官取得之材料之形式。

【第20項】

如申請專利範圍第1項或第2項之方法，其中該組測試微擾包括小髮夾RNA、單嚮導RNA、細胞介素、分子量小於4000道耳頓之化合物、滿足里賓斯基五規則準則之化合物、肽適體、衍生化肽適體、肽、核酸適體、抗體、鋅指轉錄因子、siRNA、或小髮夾RNA。

【第21項】

如申請專利範圍第1項或第2項之方法，其中該組測試

微擾包括單嚮導RNA，且在複數個細胞實例暴露於該各別測試微擾之後該複數個細胞實例之該複數個特徵之該量測係以CRISPR實驗之形式進行。

【第22項】

如申請專利範圍第1項或第2項之方法，其中在該複數個細胞實例暴露於該各別測試微擾之後該複數個細胞實例之該複數個特徵之該量測係在將該複數個細胞實例佈置於一或多個多孔盤中且用在不同離散波長下發射的一組染料染色之後進行。

【第23項】

如申請專利範圍第1項或第2項之方法，其中在該複數個細胞實例暴露於該各別測試微擾之後該複數個細胞實例之至少一亞組該複數個特徵之該量測為全基因體轉錄物組散彈槍測序實驗之一部分，該實驗以映射至該等細胞中之基因的轉錄物讀長之計數之形式量化細胞之基因表現。

【第24項】

如申請專利範圍第1項或第2項之方法，其中在該複數個細胞實例暴露於該各別測試微擾之後該複數個細胞實例之至少一亞組該複數個特徵之該量測為染色質開放性測序技術之一部分。

【第25項】

如申請專利範圍第1項或第2項之方法，其中在該複數個細胞實例暴露於該各別測試微擾之後該複數個細胞實例之至少一亞組該複數個特徵之該量測為L1000簡化表現剖

析實驗之一部分。

【第26項】

如申請專利範圍第1項或第2項之方法，其中在該複數個細胞實例暴露於該各別測試微擾之後該複數個細胞實例之至少一亞組該複數個特徵之該量測係使用(i)利用珠粒之流動式細胞測量術或(ii)成像細胞測量術獲得。

【第27項】

如申請專利範圍第1項或第2項之方法，其中在該複數個細胞實例暴露於該各別測試微擾之後該複數個細胞實例之至少一亞組該複數個特徵之該量測係通過質譜法獲得。

【第28項】

如申請專利範圍第1項或第2項之方法，其中在該複數個細胞實例暴露於該各別測試微擾之後該複數個細胞實例之至少一亞組該複數個特徵之該量測係使用電生理學或蛋白質體學成像獲得。

【第29項】

一種非暫態電腦可讀存儲媒體及嵌入其中的一或多個電腦程式，其用於確定一組測試微擾是否相對於虛無分佈區別對細胞之一或多個第一遺傳成分的中靶效應，其中該組測試微擾包含複數個對該一或多個第一遺傳成分的測試微擾，且該細胞包含複數個遺傳成分，包括該一或多個第一遺傳成分，該一或多個電腦程式包含當藉由電腦系統執行時使該電腦系統執行方法的指令，該方法包含：

(A) 針對該組測試微擾中之各個各別測試微擾，獲得

對應測試向量，從而獲得複數個測試向量，其中各對應測試向量包含複數個測試元素，該複數個測試元素中之各測試元素包含使用在複數個細胞實例暴露於各別測試微擾之後該複數個細胞實例之複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之測試分佈度量，其中該測試分佈度量之該複數個維度縮減向量小於該複數個特徵；

(B) 藉由第一過程，使用該複數個測試向量，計算複合測試度量，該複合測試度量量化該組測試微擾對該一或多個第一遺傳成分的該中靶效應，該第一過程包含：

(a) 對於該複數個測試向量中之各個各別測試向量，計算(i)該各別測試向量與(ii)在自該複數個測試向量除去該各別測試向量之情況下該複數個測試向量之測試分佈度量之間的相似性度量，從而獲得該複數個測試向量之複數個測試相似性度量，該複數個相似性度量中之各測試相似性度量表示該第一組微擾中之微擾，且

(b) 自該複數個相似性度量計算該複合測試度量；

(C) 計算該虛無分佈，其中該虛無分佈包含複數個複合控制度量，該複數個複合控制度量中之各個各別複合控制度量係藉由第二過程計算，該第二過程包含：

(a) 自複數個微擾選擇一組各別控制微擾，其中該組各別控制微擾中之各控制微擾係針對該複數個遺傳成分中除該一或多個第一遺傳成分之外的一或多個不同遺傳成分；

(b) 針對該組各別控制微擾中之各個各別控制微擾，獲得對應控制向量，從而獲得複數個各別控制向量，其中各對應控制向量包含複數個控制元素，該複數個控制元素中之各控制元素包含使用在該複數個各別細胞控制實例暴露於該各別控制微擾之後該複數個各別細胞控制實例之該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之控制分佈度量，其中該控制分佈度量之該複數個維度縮減成分小於該複數個特徵；

(c) 使用該複數個各別控制向量計算該各別複合控制度量，其中該各別控制度量量化該組各別控制微擾對該一或多個對應的不同遺傳成分的該中靶效應，且其中該控制度量係藉由第三過程計算，該第三過程包含：

(1) 對於各個各別控制向量，計算(i)該各別控制向量與(ii)在自該複數個各別控制向量除去該各別控制向量之情況下該複數個各別控制向量之分佈度量之間的控制相似性度量，從而獲得該複數個各別控制向量之複數個控制相似性度量，該複數個控制相似性度量中之各控制相似性度量獨特地對應於該組各別控制微擾中之微擾，且

(2) 計算該各別複合控制度量，作為該複數個控制相似性度量之集中趨勢量度；且

(D) 將該複合測試度量與該虛無分佈相比較，其中當該複合測試度量不同於該虛無分佈時，認為該組測試微擾相對於該虛無分佈區別對該一或多個第一遺傳成分的該中靶效應。

【第30項】

一種用於確定一組測試微擾是否相對於虛無分佈區別對細胞之一或多個第一遺傳成分的中靶效應的電腦系統，其中該組測試微擾包含複數個對該一或多個第一遺傳成分的測試微擾，且該細胞包含複數個遺傳成分，包括一或多個該第一遺傳成分，該電腦系統包含：

一或多個處理器；

記憶體；及

一或多個程式，其中該一或多個程式存儲於該記憶體中且經組態以由該一或多個處理器執行，該一或多個程式包括以下指令：

(A) 針對該組測試微擾中之各個各別測試微擾，獲得對應測試向量，從而獲得複數個測試向量，其中各對應測試向量包含複數個測試元素，該複數個測試元素中之各測試元素包含使用在複數個細胞實例暴露於各別測試微擾之後該複數個細胞實例之複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之測試分佈度量，其中該測試分佈度量之該複數個維度縮減向量小於該複數個特徵；

(B) 藉由第一過程，使用該複數個測試向量，計算複合測試度量，該複合測試度量量化該組測試微擾對該一或多個第一遺傳成分的該中靶效應，該第一過程包含：

(a) 對於該複數個測試向量中之各個各別測試向量，計算(i)該各別測試向量與(ii)在自該複數個測試向量除去

該各別測試向量之情況下該複數個測試向量之測試分佈度量之間的相似性度量，從而獲得該複數個測試向量之複數個測試相似性度量，該複數個相似性度量中之各測試相似性度量表示該第一組微擾中之微擾，且

(b) 自該複數個相似性度量計算該複合測試度量；

(C) 計算該虛無分佈，其中該虛無分佈包含複數個複合控制度量，該複數個複合控制度量中之各個各別複合控制度量係藉由第二過程計算，該第二過程包含：

(a) 自複數個微擾選擇一組各別控制微擾，其中該組各別控制微擾中之各控制微擾係針對該複數個遺傳成分中除該一或多個第一遺傳成分之外的一或多個不同遺傳成分；

(b) 針對該組各別控制微擾中之各個各別控制微擾，獲得對應控制向量，從而獲得複數個各別控制向量，其中各對應控制向量包含複數個控制元素，該複數個控制元素中之各控制元素包含使用在該複數個各別細胞控制實例暴露於該各別控制微擾之後該複數個各別細胞控制實例之該複數個特徵之量測值所計算的各別維度縮減成分在複數個維度縮減成分中之控制分佈度量，其中該控制分佈度量之該複數個維度縮減成分小於該複數個特徵；

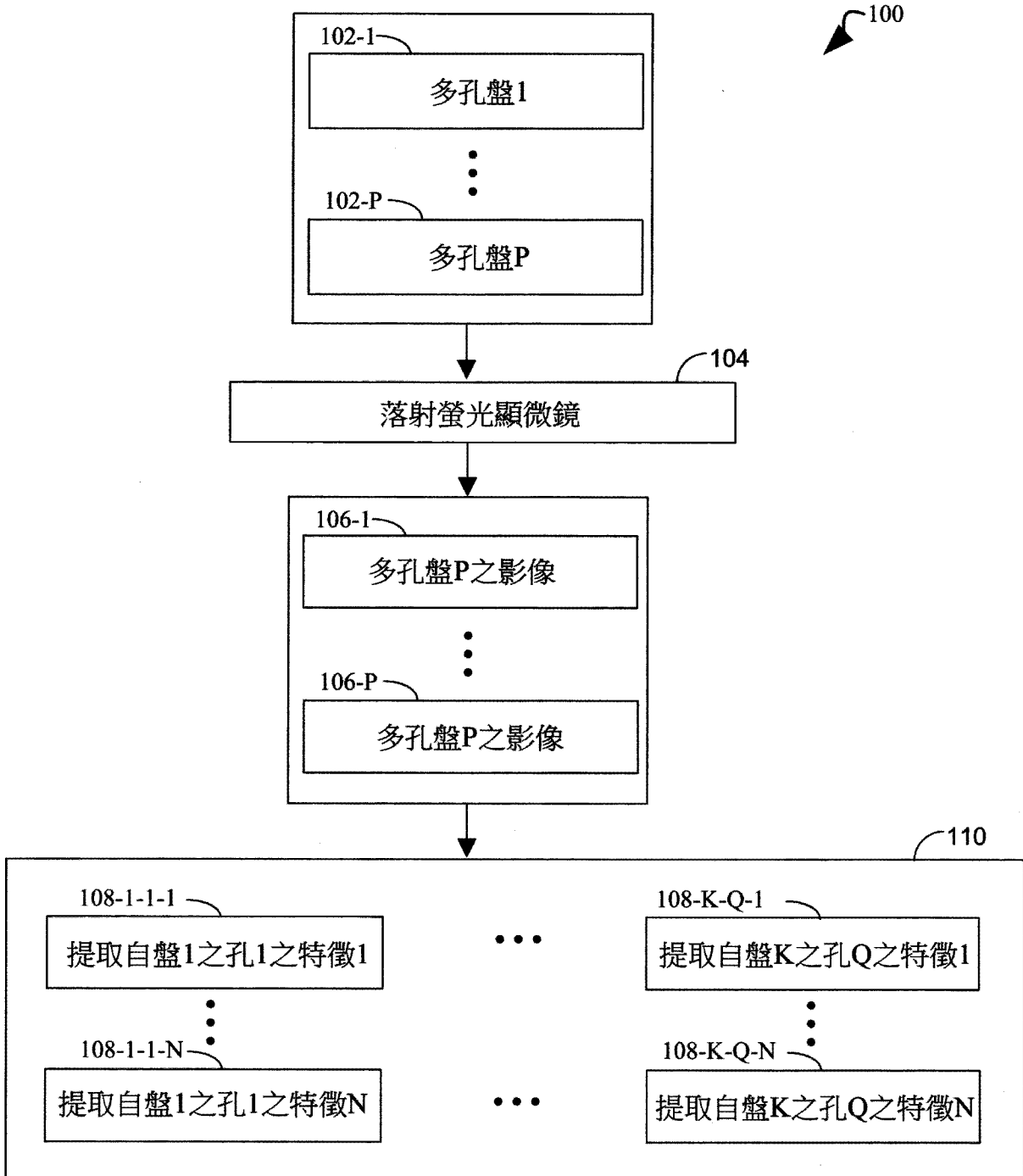
(c) 使用該複數個各別控制向量計算該各別複合控制度量，其中該各別控制度量量化該組各別控制微擾對該一或多個對應的不同遺傳成分的該中靶效應，且其中該控制度量係藉由第三過程計算，該第三過程包含：

(1) 對於各個各別控制向量，計算(i)該各別控制向量與(ii)在自該複數個各別控制向量除去該各別控制向量之情況下該複數個各別控制向量之分佈度量之間的控制相似性度量，從而獲得該複數個各別控制向量之複數個控制相似性度量，該複數個控制相似性度量中之各控制相似性度量獨特地對應於該組各別控制微擾中之微擾，且

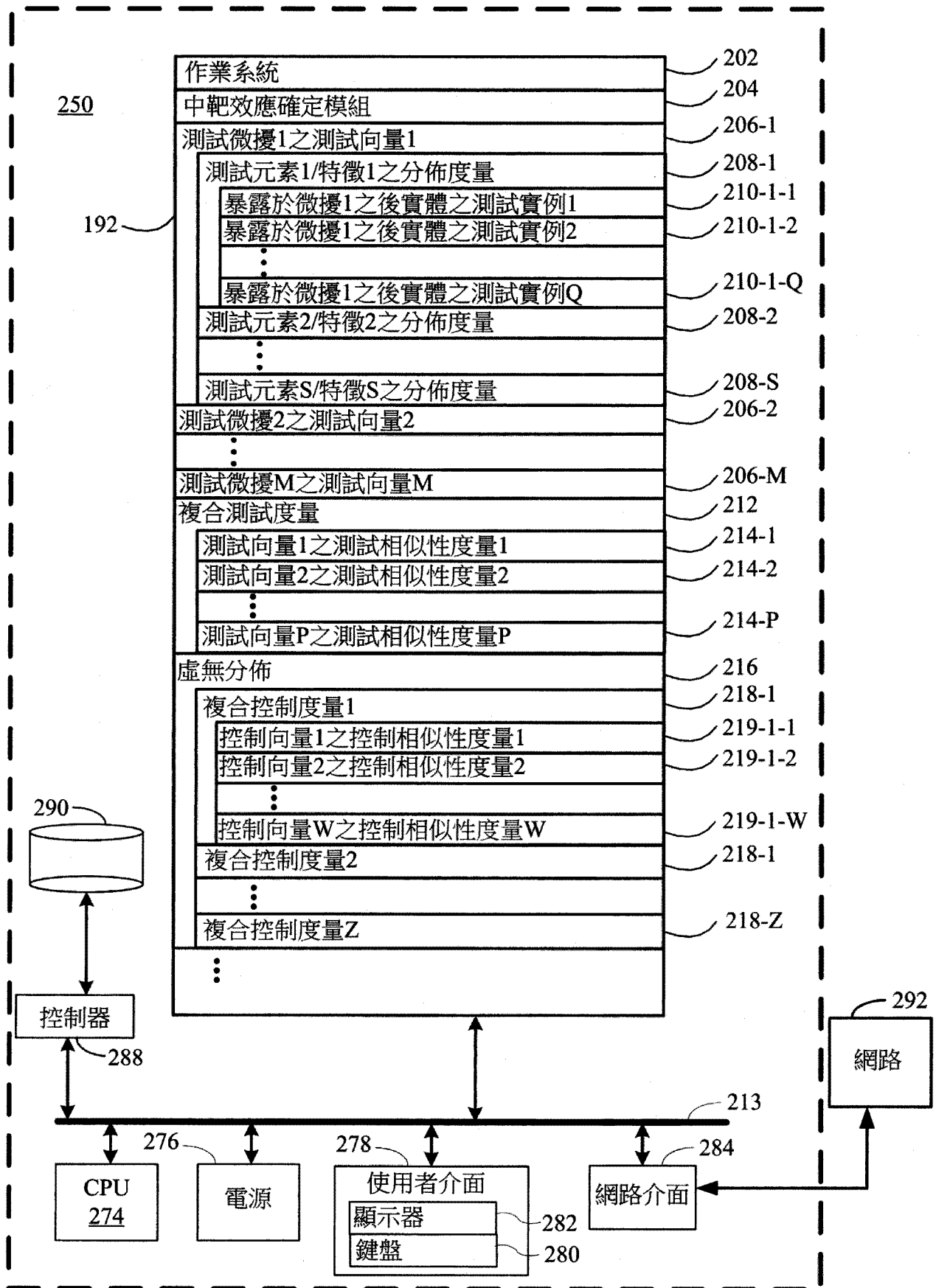
(2) 計算該各別複合控制度量，作為該複數個控制相似性度量之集中趨勢量度；且

(D) 將該複合測試度量與該虛無分佈相比較，其中當該複合測試度量不同於該虛無分佈時，認為該組測試微擾相對於該虛無分佈區別對該一或多個第一遺傳成分的該中靶效應。

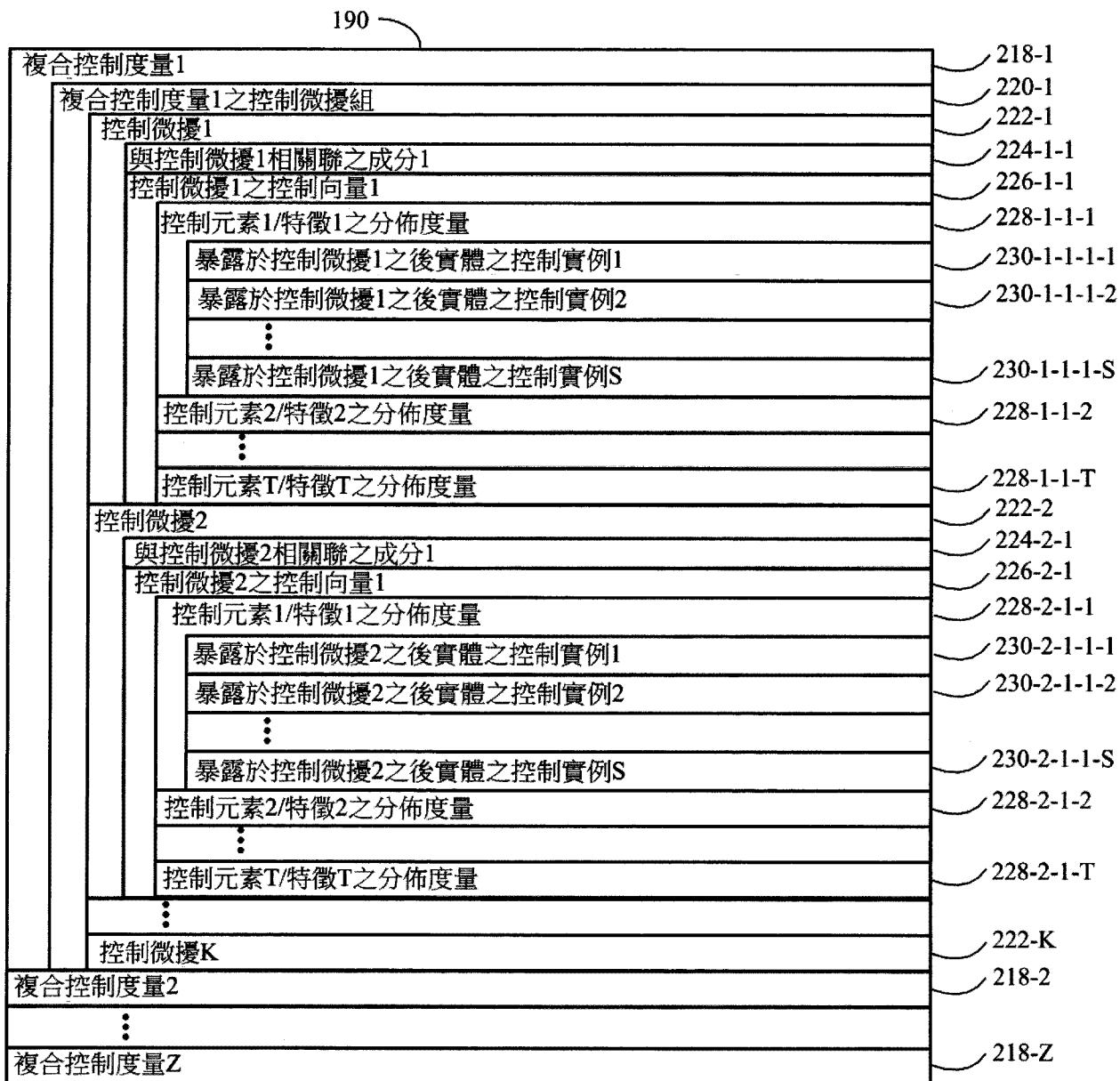
【發明圖式】



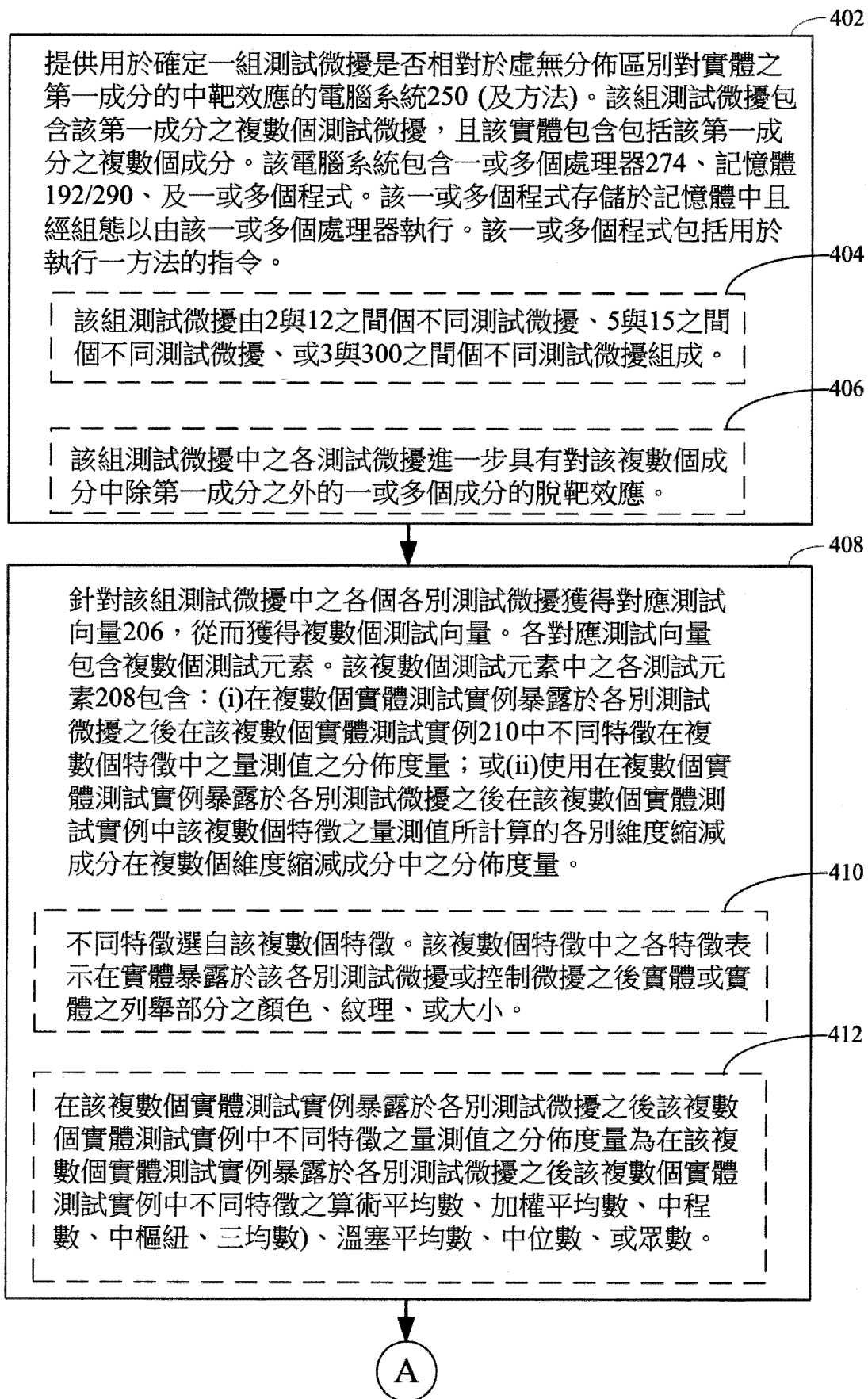
【圖 1】



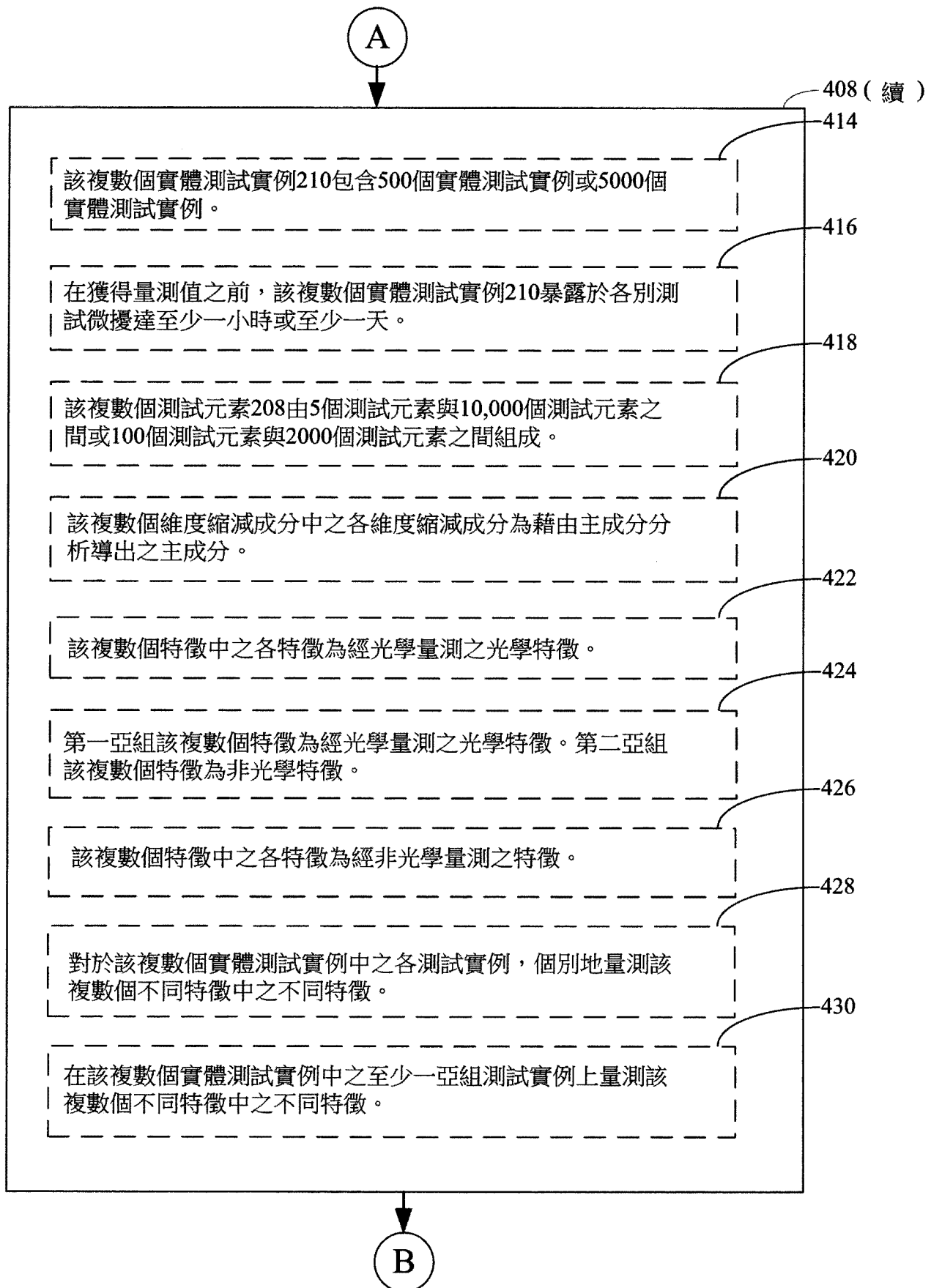
【圖 2】



【圖 3】

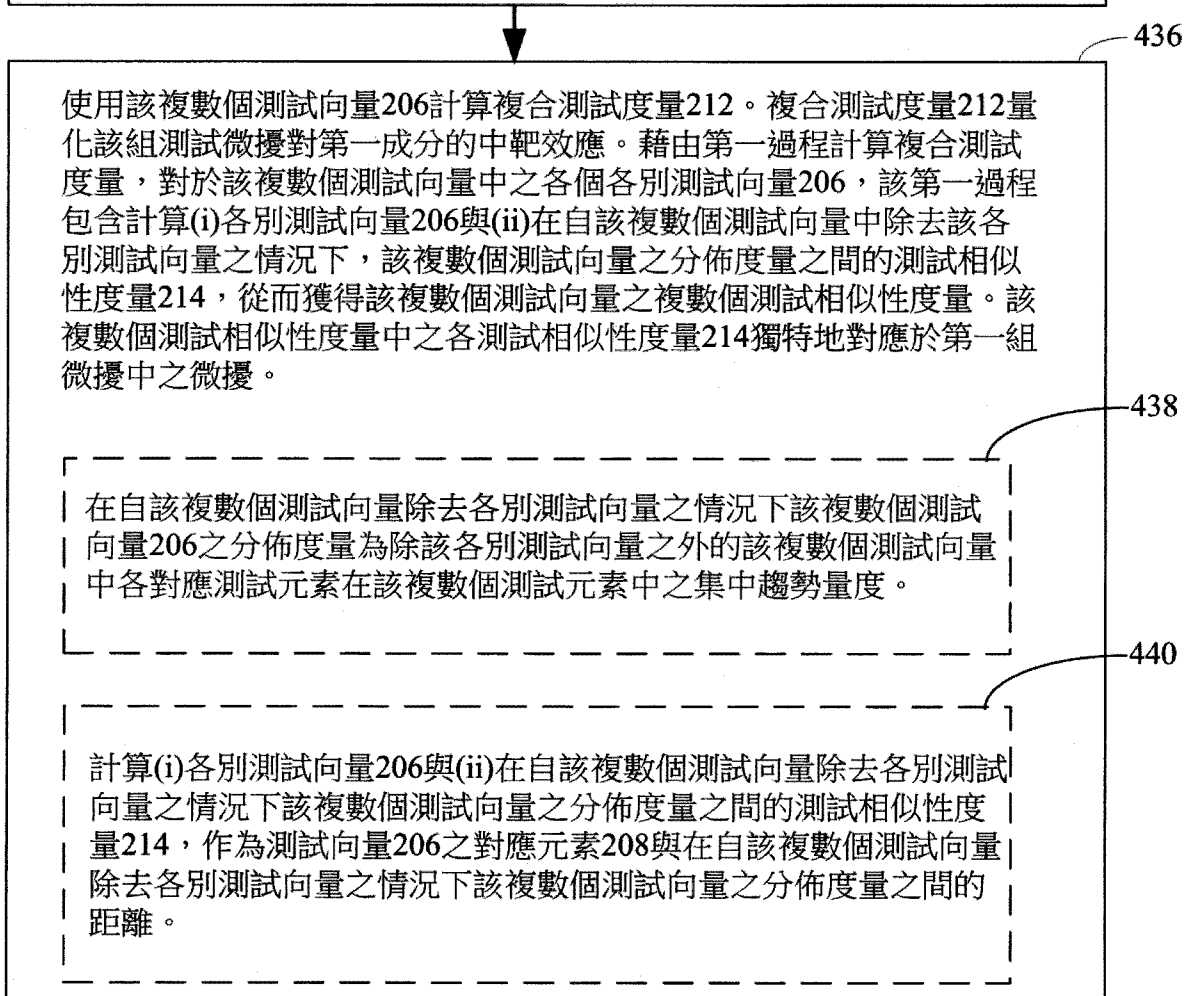
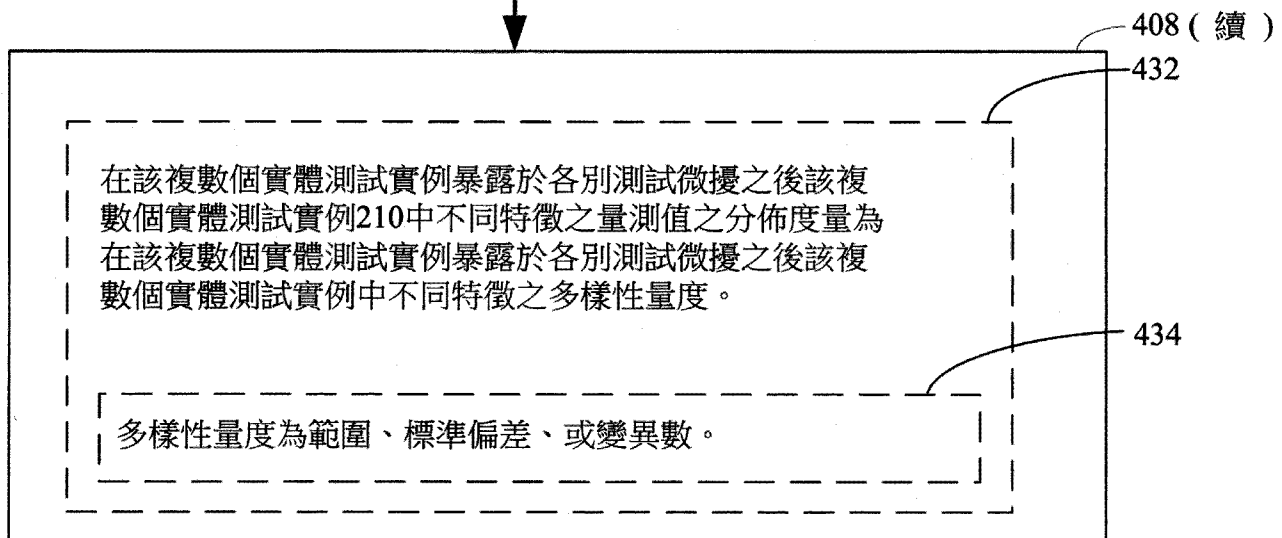


【圖 4A】



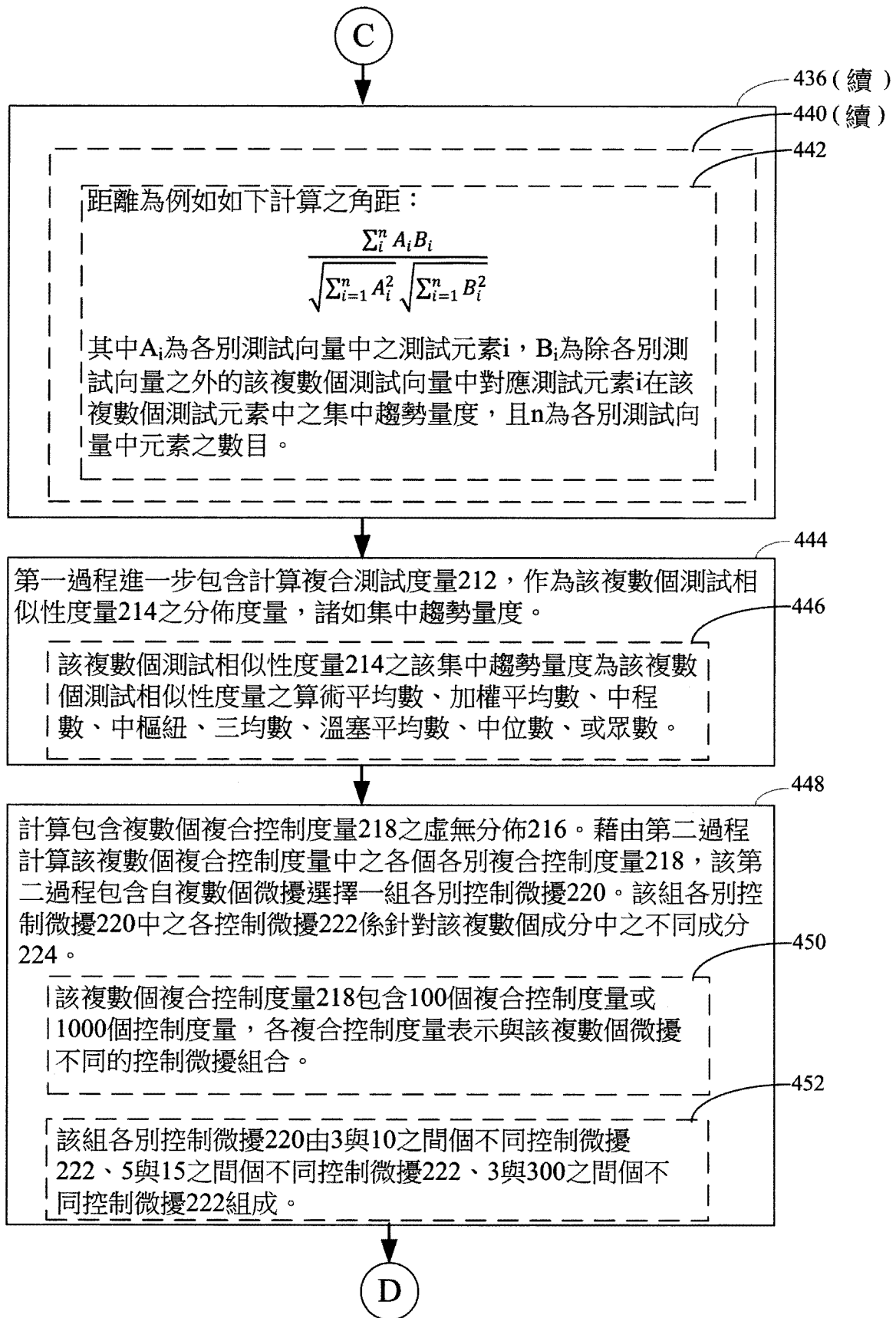
【圖 4B】

(B)

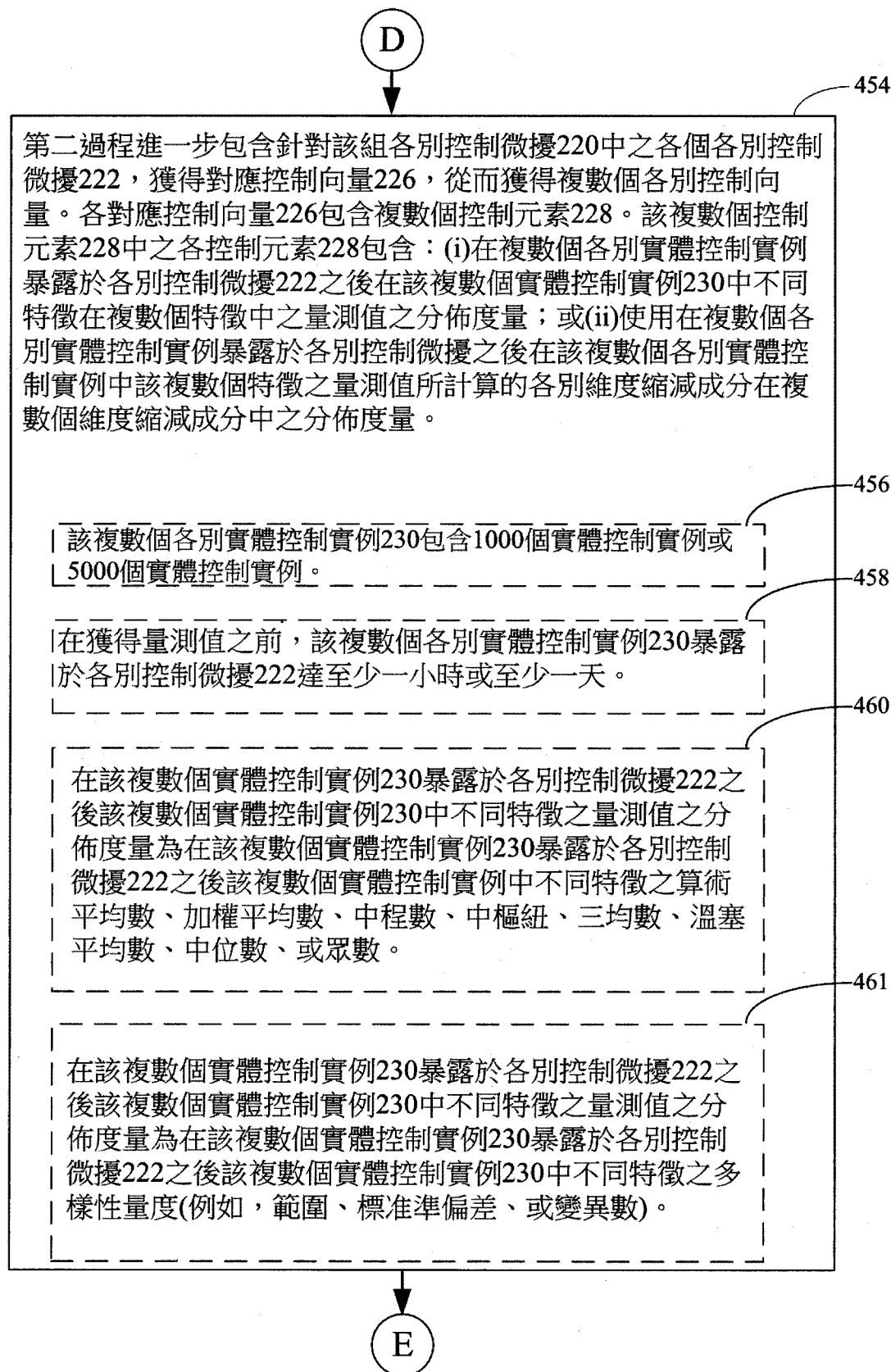


(C)

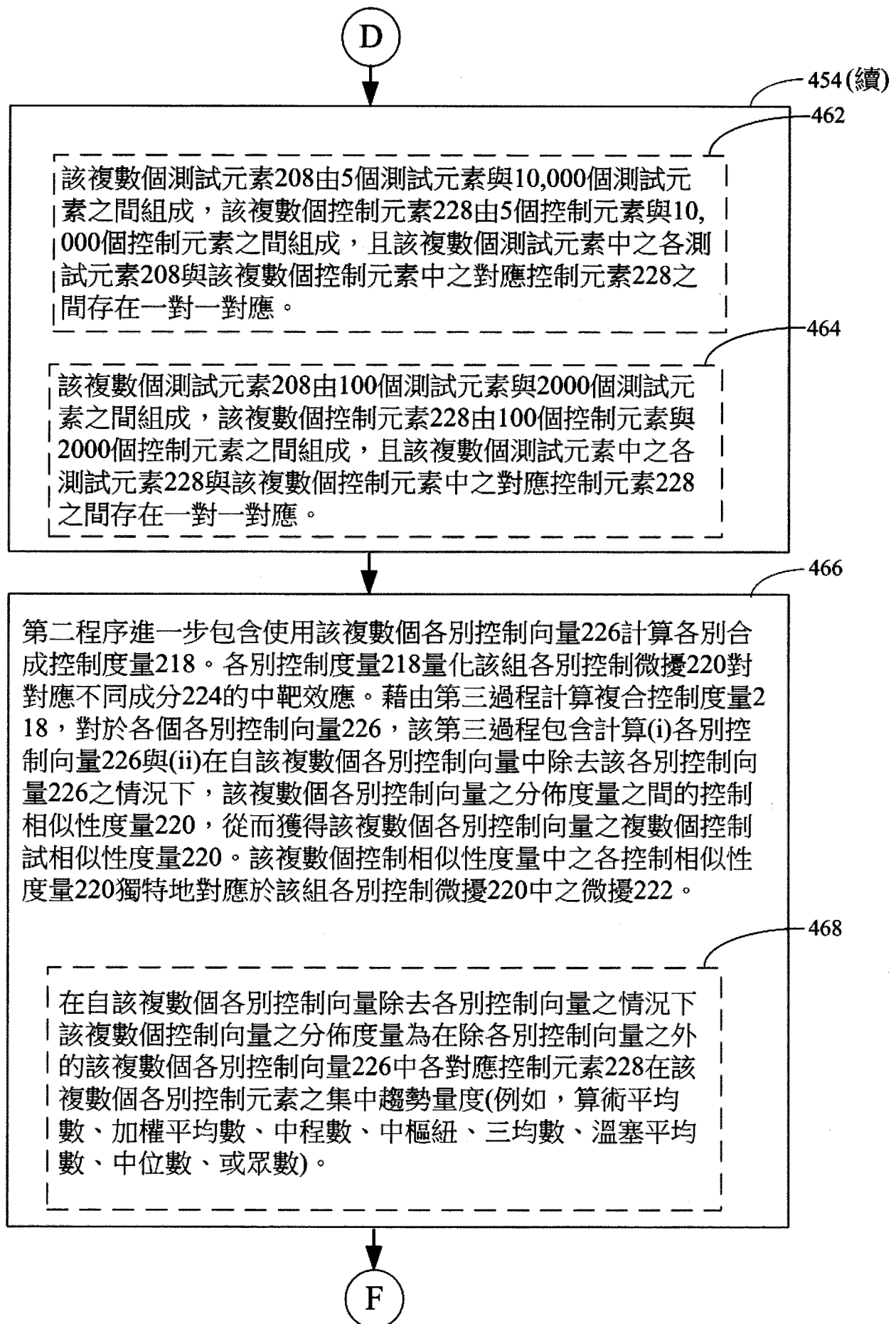
【圖 4C】



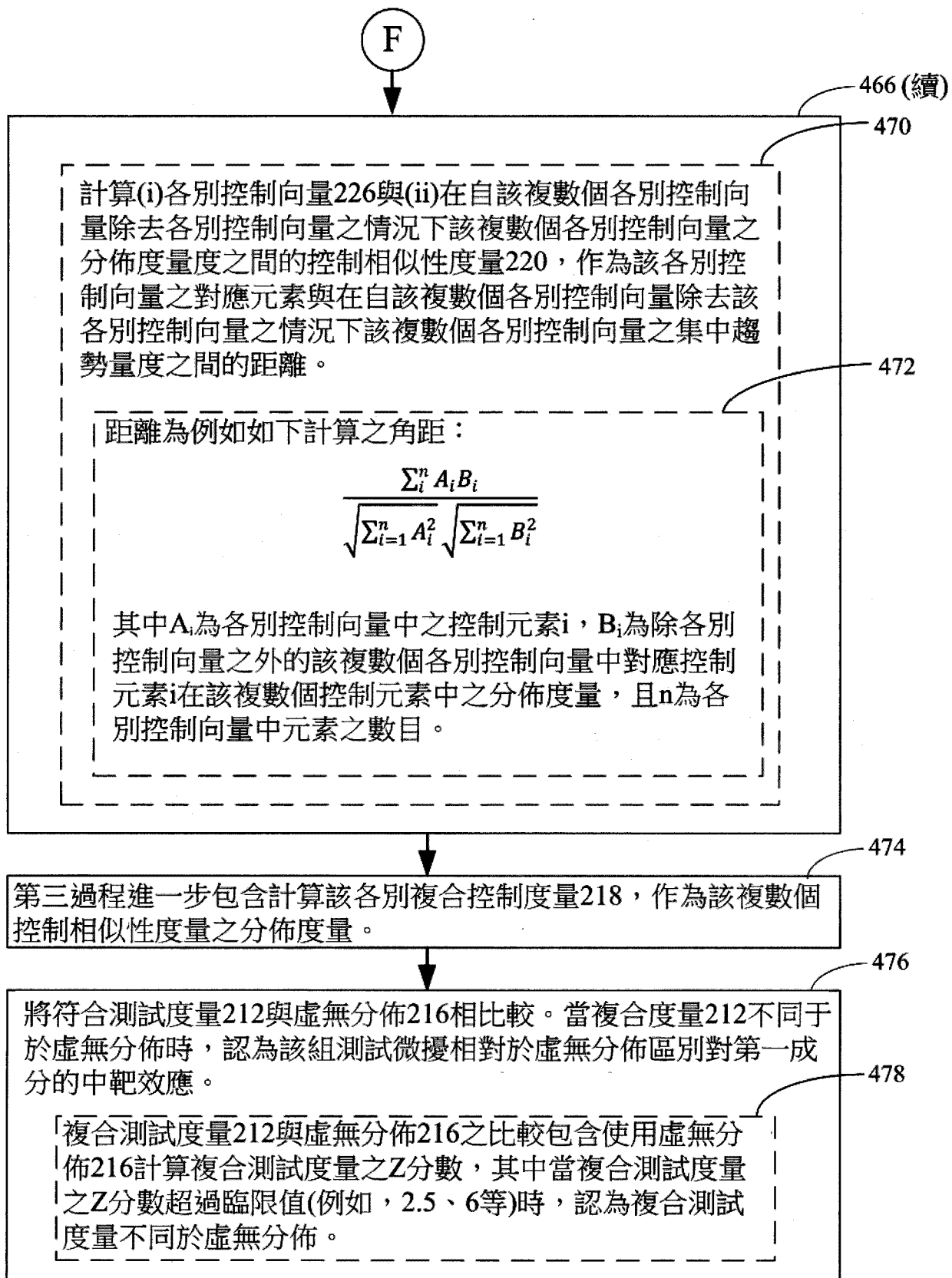
【圖 4D】



【圖 4E】



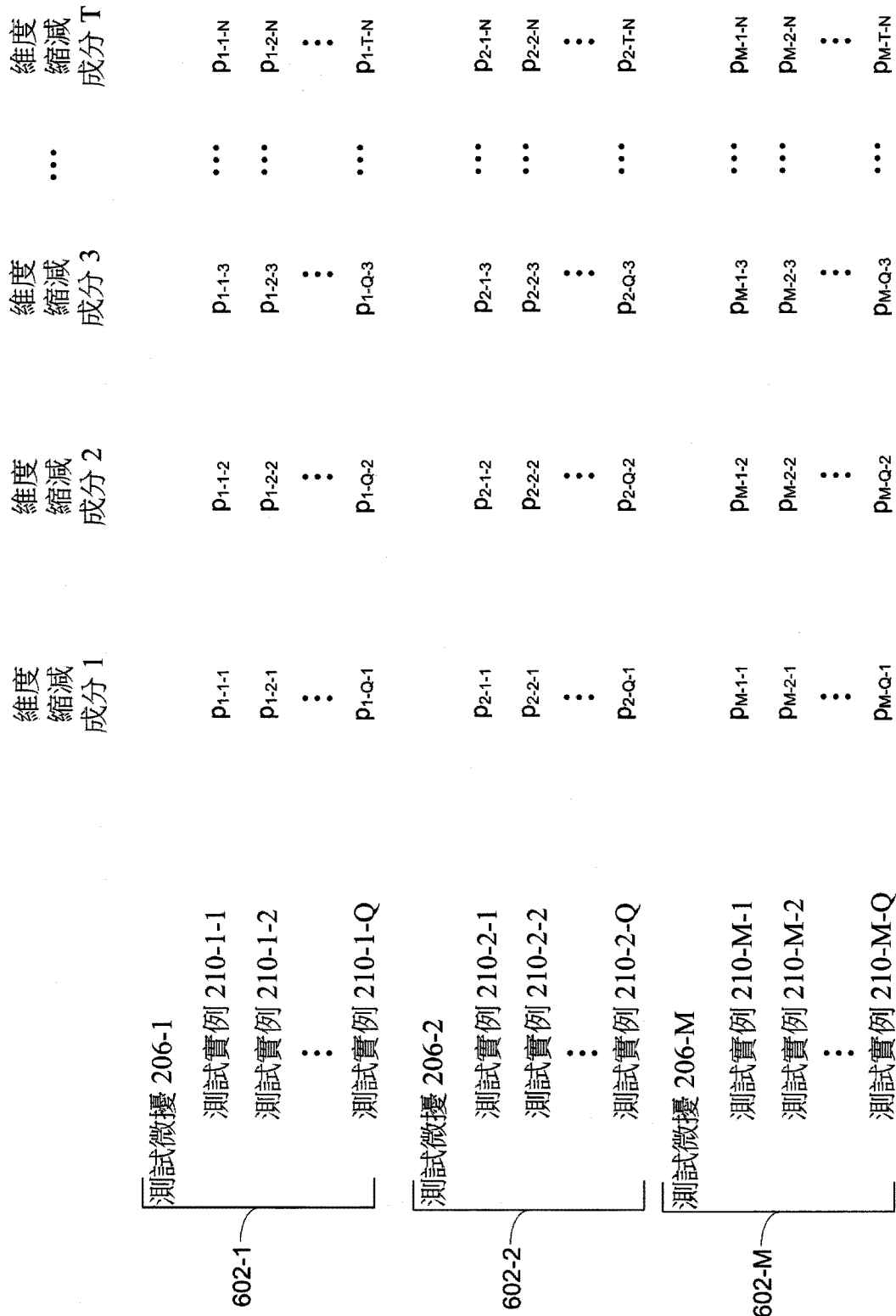
【圖 4F】



【圖 4G】

	特徵 1	特徵 2	特徵 3	...	特徵 N
控制微擾 1					
控制實例 1-1	p_{1-1-1}	p_{1-1-2}	p_{1-1-3}	...	p_{1-1-N}
控制實例 1-2	p_{1-2-1}	p_{1-2-2}	p_{1-2-3}	...	p_{1-2-N}
⋮	⋮	⋮	⋮		⋮
控制實例 1-Q	p_{1-Q-1}	p_{1-Q-2}	p_{1-Q-3}	...	p_{1-Q-N}
控制微擾 2					
控制實例 2-1	p_{2-1-1}	p_{2-1-2}	p_{2-1-3}	...	p_{2-1-N}
控制實例 2-2	p_{2-2-1}	p_{2-2-2}	p_{2-2-3}	...	p_{2-2-N}
⋮	⋮	⋮	⋮		⋮
控制實例 2-Q	p_{2-Q-1}	p_{2-Q-2}	p_{2-Q-3}	...	p_{2-Q-N}
控制微擾 M					
控制實例 M-1	p_{M-1-1}	p_{M-1-2}	p_{M-1-3}	...	p_{M-1-N}
控制實例 M-2	p_{M-2-1}	p_{M-2-2}	p_{M-2-3}	...	p_{M-2-N}
⋮	⋮	⋮	⋮		⋮
控制實例 M-Q	p_{M-Q-1}	p_{M-Q-2}	p_{M-Q-3}	...	p_{M-Q-N}

【圖 5】



【圖 6】