

科技部補助

大專學生研究計畫研究成果報告

* ***** *
* 計畫名稱：透過空間分析技術探討臺灣行政區域 DNA 甲基化與空氣污染的關係 *
* ***** *

執行計畫學生：何國丞
學生計畫編號：MOST 106-2813-C-040-036-M
研究期間：106年07月01日至107年02月28日止，計8個月
指導教授：廖勇柏

處理方式：本計畫涉及專利或其他智慧財產權，2年後可公開查詢

執行單位：中山醫學大學公共衛生學系（所）

中華民國 107年03月31日

目錄

壹、前言-----	1-5
一、研究動機與研究問題	
二、文獻回顧與探討	
貳、材料與方法-----	6-9
一、研究族群	
二、全基因體 DNA 甲基化	
三、肺癌之抑癌基因的甲基化位點	
四、地區分組	
五、空間變異分布圖繪製	
六、統計方法	
參、結果-----	10-11
一、肺癌的抑癌基因之各區域的甲基化變異程度	
二、肺癌抑癌基因的甲基化位點之空間變異分布圖	
肆、討論-----	12-13
伍、結論-----	14
陸、參考文獻-----	15-16

摘要

本研究之目的在於探討臺灣地區之肺癌抑癌基因之甲基化與空氣污染的相關性，並以空間分析技術來呈現甲基化程度的空間變異。主要利用臺灣人體生物資料庫的甲基化晶片資料作研究，並藉由臺灣癌症地圖與行政院環境保護署的空氣污染資料來做地區分組。使用複線性迴歸分析來探討各 DNA 序列區段之甲基化位點在臺灣不同區域的變異程度，並分析其劑量效應關係，再將其結果以視覺化方式呈現出空間變異圖。結果顯示肺癌抑癌基因啟動子甲基化，北部到南部呈劑量效應的增加趨勢(0.00120、0.00162、0.00192、0.00195)；在空間變異圖上的啟動子部分，可以看到大多數的甲基化程度變異值都呈現增加(>0.00)，且區域在愈南部的地區，其值增加愈多。本研究發現肺癌抑癌基因啟動子甲基化，與空氣污染嚴重程度成正比。

關鍵詞：空氣污染、肺癌、抑癌基因、甲基化、空間變異

壹、前言

一、研究動機與研究問題

近年來，在科技及經濟的快速發展下，空氣污染，已經是全世界必須要重視的嚴重問題。臺灣的工業污染排放、車輛污染排放以及商業污染排放都愈來愈多，加上東北季風的因素，使中國的空氣污染從境外移入臺灣，造成空氣品質每況愈下，成為國人健康上的一大危害。許多文獻都指出空氣中的污染物質，對人體會造成諸多有害的影響，尤其是近幾年特別被關注的 PM2.5，很多有害重金屬物質或是有毒的有機化合物質均會附著在這類細懸浮微粒上，隨其進入人體，導致 DNA 甲基化的異常。然而，這些文獻大部分都是國外專家學者所作的研究，在臺灣鮮少有研究特別針對空氣污染與 DNA 甲基化作分析探討。因此，本研究利用中央研究院針對台灣民眾建置的「臺灣人體生物資料庫」的資料，將 DNA 甲基化資料與空氣污染進行流行病學上的研究並搭配空間變異上的分析，探討 DNA 甲基化與空氣污染在臺灣地區是否也會有其相關性。

精準醫療的時代一步步來臨，人們對於基因定序研究越來越重視，許多研究發現 DNA 甲基化的異常(高度或低度)會影響人體疾病的發生，且有文獻更指出 DNA 甲基化的異常大多都會出現在疾病發生之

前。然而，目前國內外都尚無人將 DNA 甲基化與空氣污染結合空間資訊技術來做分析探討。因此，本研究利用 DNA 甲基化與空氣污染進行空間上的變異分析，預測推算出不同縣市的 DNA 甲基化異常程度，讓民眾清楚知道自己所居住的地區是否位於 DNA 甲基化異常的高危險地區，並作出最適當的疾病預防措施，使 DNA 甲基化異常在早期就被診斷及治療，而不會發展成疾病。

綜觀上述研究動機及背景，本研究之目的以條列方式分述如下：

- (1) 藉由流行病學的探討及空間變異的分析，瞭解臺灣地區的空氣污染與 DNA 甲基化之間的相關性；
- (2) 利用「臺灣人體生物資料庫」以及空間資訊技術，依參與對象的現居地進行分區，繪製臺灣地區 DNA 甲基化之空間變異。

二、文獻回顧與探討

目前流行病學研究已經發現空氣污染與總死亡率的增加有關【1-2】，而暴露到空氣污染物，如：多環芳香烴、二氧化硫、二氧化氮、PM10 及 PM2.5，也會與過敏、呼吸道疾病、心血管疾病、肺部疾病以及癌症的發生率增加有顯著關係【3-7】。而疾病的發生原因有很多，但現在較廣泛被接受的是，由遺傳和表觀遺傳因素，還有環境之間各種暴露相互作用所引起的【8】。因此現今若想要探討各種疾病的發生，不能僅侷限在環境方面的暴露，而是要更進一步去探討基因層面的因素，這樣才能更精準得分析出導致疾病的危險因子。

表觀遺傳是在以不改變基因組序列的前提之下，透過 DNA 和組蛋白的修飾來調控基因表達，目前這種修飾以 DNA 甲基化最為常見，其甲基化的胞嘧啶通常會發生於胞嘧啶與鳥嘌呤相鄰的序列上，即 CpG site【9】。DNA 甲基化是一種重要的表觀遺傳修飾方法，普遍存在於人體基因內，對細胞的正常運作以及發育是不可或缺的，不過有研究指出 DNA 甲基化的異常，例如 CpG island 高度甲基化或 Global DNA 低度甲基化都會與疾病，像是癌症、腫瘤、動脈粥樣硬化、缺血性心臟病有顯著相關，且 DNA 甲基化的異常通常都會出現在疾病發生之前【10-13】。由此可知正常的 DNA 甲基化在人體內是必須的，但如果甲基化出現異常，例如致癌基因低度甲基化或抑癌基因高度甲

基化就會對人體造成危害，因此若能對異常的 DNA 甲基化發生作出準確的預測，或是在發展成疾病前就被檢測出來，這樣不僅能有效地降低疾病的發生也能大幅地節省醫療成本。

近十年(民國九十五年至民國一零五年)的國人十大死因中，惡性腫瘤一直都位居十大死因之首，而在癌症標準化死亡率中，肺癌雖然逐年呈現下降趨勢，但依然位居於主要癌症死因第一位，且肺癌的標準化死亡率都遠遠高於其他癌症【14】。因此本研究針對肺癌相關的基因進行甲基化分析，結合空間分析技術，探討台灣各地區與空氣污染及 DNA 甲基化，三者之間的關係，並依據各地區的甲基化程度，做出最適合的預防措施，以預防肺癌的發生。

近年有研究顯示空氣物污染會造成 DNA 甲基化異常【15-19】。而 PM10 及 PM2.5 都是空氣中的主要污染物，暴露到這些物質，會造成 DNA 甲基化的多種改變【15.20-22】，至於其它污染物，像是砷、鉛、二氧化硫、多環芳香烴也都會造成 DNA 甲基化的異常改變【15】，由此可見空氣污染是造成 DNA 甲基化異常的重要因素。

在許多的文獻中，雖然有指出空氣污染對 DNA 甲基化異常有顯著的影響，但這些文獻大多都是國外學者專家所作的研究，且分析研究的方法尚無人利用空間資訊技術來探討空氣污染與 DNA 甲基化異

常的空間變異關係。因此，本研究主要是透過「臺灣人體生物資料庫」中的 DNA 甲基化資料與空氣污染來作分析探討，並利用空間資訊技術來作空間上的分析探討。

貳、材料與方法

本研究找出肺癌的抑癌基因之甲基化位點，並依空氣污染的程度將台灣地區進行分組，再透過分析各組區域與各個抑癌基因之甲基化位點的相關性，探討各組區域之甲基化程度的劑量效應關係，並探究抑癌基因之甲基化的關係。最後在利用空間資訊處理技術，繪製出抑癌基因之各個甲基化位點與各組區域的空間變異分布圖。

一、研究族群

運用臺灣人體生物資料庫的收案資料，採取自願者參與的方式，不設性別限制，其年齡在30至70歲之間，且具有行為能力的國人。但排除不具有臺灣國籍、具有外國血統或是經醫生確診為罹患癌症的民眾。

二、全基因體DNA甲基化

藉由臺灣人體生物資料庫的全基因體甲基化晶片(Methylation microarray)之資料來研究分析DNA甲基化的狀況。其晶片可以直接偵測單一位點的甲基化比例，並將偵測到的CpG位點之甲基化比例分析轉換成一個 β 值，其數值範圍為0~1，代表意義為hypomethylation 到 hypermethylation。因每個人大約可以偵測到850,000個甲基化位點的資訊，而每個位點的甲基化程度又都不相同，所以每個位點都會有一

個 β 值。

三、肺癌之基因的甲基化位點

運用TSGene2.0找出與肺癌相關的抑癌基因，其基因之數據來源均是來自National Library of Medicine(NLM)美國國家醫學圖書館之National Center for Biotechnology Information (NCBI) 國家生技資訊中心製作的資料庫PubMed及Gene Reference Into Function(GeneRIF)【23】。接著與臺灣人體生物資料庫之基因進行查詢比對，找出臺灣人體生物資料庫也有檢測紀錄的基因，最後再從甲基化晶片之資料中找出這些基因各區段的甲基化位點，包含啟動子(promoter)、外顯子(exon)、內含子(intron)、3'非轉譯區(3'UTR)。

四、地區分組

主要依照臺灣癌症地圖(TAIWAN CANCER MAP)之肺癌死亡率的時空變異趨勢圖，並參考搭配行政院環境保護署之臺灣地區空氣污染資料。而為使研究更貼近日常生活，本研究以行政院環境保護署在2006年的細懸浮微粒(PM2.5)濃度值作為參考依據，將臺灣地區由北至南分成五個區域：「新北市、臺北市」，「桃園縣、新竹縣、新竹市、苗栗縣」，「臺中市、南投縣」，「彰化縣、雲林縣」，「嘉義縣、嘉義市、臺南市」，以利進行統計分析及探討。

五、空間變異分布圖繪製

為使空間資訊技術能將統計分析之資料以視覺化的方式呈現，本計畫擬分析臺灣人體生物資料庫，將DNA甲基化資料與臺灣地區由北至南的五個分組區域加以調整分析，透過R語言(R Programming Language x64 3.4.4)之高階繪圖(high-level plotting functions)的指令函數以視覺化方式呈現。並依執行計畫以肺癌的抑癌基因之甲基化全位點、甲基化啟動子位點、甲基化外顯子位點、甲基化內含子位點、甲基化3'非轉譯區位點等五個方式來繪製空間變異分布圖，以探討臺灣各區域間DNA甲基化的變異程度。

六、統計方法

本研究利用SAS 9.4 統計套裝軟 (SAS Institute Inc., Cary, NC, USA)及R語言(R Programming Language x64 3.4.4)進行分析。接著將藉由參考臺灣癌症地圖及環保署空氣污染資料所分成的五個區域設虛擬變項，並使用多變項線性迴歸(multiple linear regression)探討分析其劑量效應關係。而為使能精準地分析出各個甲基化位點的甲基化程度，本計畫將以一個甲基化位點執行一次複線性迴歸的方式來做統計分析，且在每一次複線性迴歸模式中均調整個人健康狀況、飲食習慣、生活習慣、血液生化值。

最後，依甲基化全位點、甲基化啟動子位點、甲基化外顯子位點、甲基化內含子位點、甲基化3'非轉譯區位點等五個分層，計算分析每一個甲基化位點在臺灣五個區域分組中的甲基化程度變異值，再統計轉換成各區域的肺癌之抑癌基因的甲基化程度變異值，並利用R語言之高階繪圖指令函數繪製DNA甲基化的空間變異分布圖。

參、結果

一、肺癌的抑癌基因之各區域的甲基化變異程度

利用多變項線性迴歸(multiple linear regression) 分析各DNA序列區段之甲基化位點在臺灣不同區域的變異程度，研究中考慮了年齡、性別、婚姻、教育程度、抽菸、喝酒、嚼檳榔、飲食習慣、生活習慣、血液生化值等因素。

表一為肺癌的抑癌基因之各區域的甲基化變異程度。可看到調整上述各因子以後：在抑癌基因的甲基化啟動子平均變異程度方面，區域一到區域四(北部到南部)呈現劑量效應的增加趨勢(0.00120、0.00162、0.00192、0.00195)。

二、肺癌的抑癌基因的甲基化位點之空間變異分布圖

本研究分析台灣人體生物資料庫之甲基化晶片的資料，計算肺癌的抑癌基因之各個甲基化位點程度，以視覺化的方式呈現各區域的空間變異。

圖一到圖四呈現出區域一到區域四之全位點、啟動子位點、外顯子位點、內含子位點、3'非轉譯區位點的甲基化變異程度，每個位點均有與對照組(新北市、臺北市)相比的四個區域之變異值：區域一(青

藍色)、區域二(棕色)、區域三(灰色)、區域四(黑色)。在圖一.b(肺癌的抑癌基因之甲基化啟動子的空間變異圖)，可以看到大多數的甲基化程度變異值都呈現增加(>0.00)，且區域在愈南部的地區，其值增加愈多。

肆、討論

本研究是國內第一篇藉由臺灣區域分組，探討空氣污染與肺癌的抑癌基因之各個甲基化位點關係的研究，再進一步以視覺化方式繪製甲基化的空間變異圖。本研究利用臺灣人體生物資料庫，大量地將肺癌相關的抑癌基因之甲基化位點進行複線性迴歸分析。本研究發現同時考慮了年齡、性別、婚姻、教育程度、抽菸、喝酒、嚼檳榔、飲食習慣、生活習慣、血液生化值等因素後，肺癌抑癌基因在啟動子甲基化平均變異程度方面，由北部至南部呈劑量效應的增加趨勢。

在肺癌方面，根據衛生福利部的調查，肺癌在近十年來一直排名在國人前十大癌症死因的前二名【14】。過去許多專家學者都在研究空氣污染與肺癌的關係，有研究顯示暴露到環境中的空氣污染物質，例如：PM、PM2.5、SO2 等物質，會與肺癌發生及死亡風險的增加有顯著關係【3.4.24.25】。過去研究也指出在人類的癌症中，抑癌基因發生表觀遺傳改變會導致其抑制癌症的功能失活，在抑癌基因 CpG island 高甲基化的轉錄沉默被認為是一個顯著機制，因此與肺癌相關的抑癌基因高度甲基化，例如 P53、RB、PTEN、FHIT 高度甲基化，亦與肺癌的發生有顯著相關【26.27】。有研究更指出暴露到空氣污染物質，例如重複暴露到 PM、PM2.5 等物質，會誘導抑癌基因失去功能，

例如 P53 啟動子高甲基化，啟動子高甲基化會抑制 P53 的表達，導致其抑癌功能失活，增加肺癌發生的風險【28.29】。而本研究也發現，肺癌抑癌基因的甲基化啟動子平均變異程度方面，區域一到區域四(北部到南部)呈現劑量效應的增加趨勢(0.00120、0.00162、0.00192、0.00195)，其與臺灣區域的空氣品質指標呈正比，愈南部空氣污染愈嚴重，甲基化程度愈高。

本研究結果顯示出肺癌之抑癌基因啟動子部分的甲基化平均程度隨空氣污染程度而增加，但卻在甲基化外顯子、內含子和 3'非轉譯區方面看到其甲基化程度的下降，雖然下降幅度沒有呈劑量效應關係，不過都在區域四(空氣汙染最嚴重的區域)下降最多(-0.00260、-0.00349、-0.00126)。

我們認為此結果是符合我們研究的族群(無癌症)以及研究的方法(全抑癌基因)。因此，我們合理推測，當空氣污染濃度愈高時，肺癌抑癌基因之啟動子甲基化會被增加愈多，進而會慢慢導抑癌功能的沉默；而當空污濃度愈高導致肺癌抑癌機制慢慢沉默時，便會產生負回饋來誘發其他區段的甲基化程度下降，以保護其基因能維持部分功能來讓人體正常健康的運作。因此，本研究雖有族群現居於空污嚴重的區域四，但到目前為止均是無肺癌的健康族群。

伍、結論

本研究發現肺癌抑癌基因啟動子的甲基化程度，會隨各區域的空氣污染嚴重程度上升而增加，與過往的研究一致。

因此，我們建議盡可能不要直接暴露在空氣污染嚴重的地方，而暴露於毒性物質甚多的環境時一定要做好安全防護措施，最重要的是要維持正常作息、均衡飲食以及規律運動，才能維持抑癌基因正常的運作。

陸、參考文獻

1. Brook, Robert D., et al. "Particulate matter air pollution and cardiovascular disease." *Circulation* 121.21 (2010): 2331-2378.
2. Samet, Jonathan M., et al. "Fine particulate air pollution and mortality in 20 US cities, 1987–1994." *New England journal of medicine* 343.24 (2000): 1742-1749.
3. Demetriou, Christiana A., and Paolo Vineis. "Carcinogenicity of ambient air pollution: use of biomarkers, lessons learnt and future directions." *Journal of thoracic disease* 7.1 (2015): 67.
4. Cohen, Aaron J., et al. "The global burden of disease due to outdoor air pollution." *Journal of Toxicology and Environmental Health, Part A* 68.13-14 (2005): 1301-1307.
5. Franchini, Massimo, and Pier Mannuccio Mannucci. "Air pollution and cardiovascular disease." *Thrombosis research* 129.3 (2012): 230-234.
6. Kelly, F. J., and J. C. Fussell. "Air pollution and airway disease." *Clinical & Experimental Allergy* 41.8 (2011): 1059-1071.
7. Takizawa, Hajime. "Impact of air pollution on allergic diseases." *The Korean journal of internal medicine* 26.3 (2011): 262-273.
8. Fleischhacker, Michael, et al. "The role of DNA methylation as biomarkers in the clinical management of lung cancer." *Expert review of respiratory medicine* 7.4 (2013): 363-383.
9. 顾婷婷, 张忠明, and 郑鹏生. "DNA 甲基化研究方法的回顾与评价." *中国妇幼保健研究* 17.6 (2006): 555-560.
10. HUANG, Zhen-ming, et al. "Advances in association between DNA methylation with environmental exposure in lung cancer." *TUMOR* 36.11 (2016): 1272-1279.
11. Baccarelli, Andrea, et al. "Ischemic heart disease and stroke in relation to blood DNA methylation." *Epidemiology (Cambridge, Mass.)* 21.6 (2010): 819.
12. Kim, Myungjin, et al. "DNA methylation as a biomarker for cardiovascular disease risk." *PloS one* 5.3 (2010): e9692.
13. Chi, Gloria C., et al. "Long-term outdoor air pollution and DNA methylation in circulating monocytes: results from the Multi-Ethnic Study of Atherosclerosis (MESA)." *Environmental Health* 15.1 (2016): 119.
14. 衛生福利部 <https://www.mohw.gov.tw/mp-1.html>
15. Cao, Yi. "Environmental pollution and DNA methylation: carcinogenesis, clinical significance, and practical applications." *Frontiers of medicine* 9.3 (2015): 261-274.
16. Byun, Hyang-Min, et al. "Effects of Air Pollution and Blood Mitochondrial DNA Methylation on Markers of Heart Rate Variability." *Journal of the American Heart Association* 5.4 (2016): e003218.

17. De Prins, Sofie, et al. "Influence of ambient air pollution on global DNA methylation in healthy adults: a seasonal follow-up." *Environment international* 59 (2013): 418-424.
18. Baccarelli, Andrea, et al. "Rapid DNA methylation changes after exposure to traffic particles." *American journal of respiratory and critical care medicine* 179.7 (2009): 572-578.
19. Breton, Carrie V., and Amy N. Marutani. "Air pollution and epigenetics: recent findings." *Current environmental health reports* 1.1 (2014): 35-45.
20. Hou, Lifang, et al. "Altered methylation in tandem repeat element and elemental component levels in inhalable air particles." *Environmental and molecular mutagenesis* 55.3 (2014): 256-265.
21. Bind, Marie-Abele, et al. "Air pollution and gene-specific methylation in the Normative Aging Study: association, effect modification, and mediation analysis." *Epigenetics* 9.3 (2014): 448-458.
22. Lepeule, Johanna, et al. "Epigenetic influences on associations between air pollutants and lung function in elderly men: the normative aging study." (2014).
23. Zhao, Min, et al. "TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes." *Nucleic acids research* 44.D1 (2015): D1023-D1031.
24. Pope III, C. Arden, et al. "Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution." *Jama* 287.9 (2002): 1132-1141.
25. Turner, Michelle C., et al. "Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers." *American journal of respiratory and critical care medicine* 184.12 (2011): 1374-1381.
26. Kohno, Takashi, and Jun Yokota. "How many tumor suppressor genes are involved in human lung carcinogenesis?" *Carcinogenesis* 20.8 (1999): 1403-1410.
27. Osada, Hirotaka, and Takashi Takahashi. "Genetic alterations of multiple tumor suppressors and oncogenes in the carcinogenesis and progression of lung cancer." *Oncogene* 21.48 (2002): 7421.
28. Zhou, Wei, et al. "Repeated PM_{2.5} exposure inhibits BEAS-2B cell P53 expression through ROS-Akt-DNMT3B pathway-mediated promoter hypermethylation." *Oncotarget* 7.15 (2016): 20691.
29. Li, R., Zhou, R., Zhang, J. "Function of PM_{2.5} in the pathogenesis of lung cancer and chronic airway inflammatory diseases (Review)". *Oncology Letters* 15.5 (2018): 7506-7514.

表一、肺癌的抑癌基因之各區域的甲基化變異程度

	地區	抑癌基因之甲基化迴歸係數				總和
		個數	最小值	最大值	平均值	
全位點						
	區域一	2340	-0.04350	0.08309	-0.00179	-4.17745
	區域二	2340	-0.03553	0.04614	-0.00150	-3.51155
	區域三	2340	-0.04349	0.08245	-0.00031	-0.72643
	區域四	2340	-0.05359	0.10105	-0.00242	-5.65208
啟動子						
	區域一	323	-0.03388	0.04140	0.00120*	0.38668
	區域二	323	-0.02436	0.03765	0.00162*	0.52439
	區域三	323	-0.03711	0.04632	0.00192*	0.62122
	區域四	323	-0.05028	0.05694	0.00195*	0.63144
外顯子						
	區域一	323	-0.02836	0.03099	-0.00193	-0.62434
	區域二	323	-0.01936	0.02158	-0.00137	-0.44205
	區域三	323	-0.02791	0.03068	-0.00075	-0.24253
	區域四	323	-0.04313	0.04043	-0.00260	-0.83931
內含子						
	區域一	1482	-0.04350	0.08309	-0.00254	-3.75996
	區域二	1482	-0.03553	0.04614	-0.00235	-3.47851
	區域三	1482	-0.03536	0.08245	-0.00070	-1.03106
	區域四	1482	-0.04750	0.10105	-0.00349	-5.17722
3'非轉譯區						
	區域一	212	-0.04257	0.04121	-0.00085	-0.17982
	區域二	212	-0.02766	0.03032	-0.00054	-0.11538
	區域三	212	-0.04349	0.08208	-0.00035	-0.07406
	區域四	212	-0.05359	0.07411	-0.00126	-0.26699

區域一：桃園縣、新竹縣、新竹市、苗栗縣

區域二：臺中市、南投縣

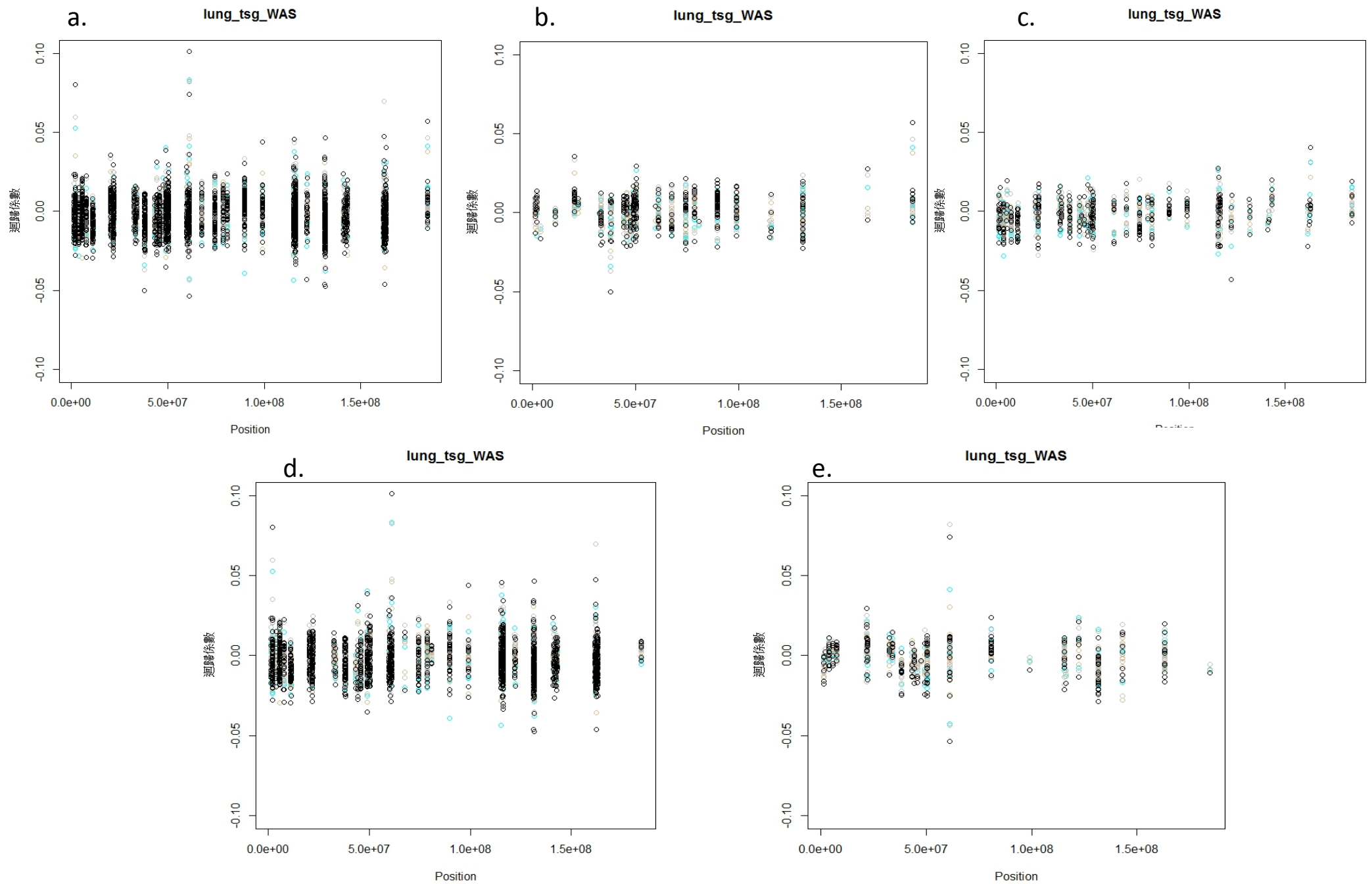
區域三：彰化縣、雲林縣

區域四：嘉義縣、嘉義市、臺南市

每個區域的甲基化迴歸係數均是與對照區域(新北市、臺北市)比較後所得的結果

每個甲基化迴歸係數均已調整年齡、性別、教育程度、婚姻、菸、酒、檳榔、素食、咖啡、茶、煮食、線香、運動、血糖、總膽固醇、三酸甘油脂、高密度脂蛋白膽固醇、低密度脂蛋白膽固醇、血中尿素氮、肌酸酐、尿酸、血壓、BMI

*呈劑量效應關係



圖一、肺癌的抑癌基因之甲基化的空間變異圖

a. 全位點 b. 啟動子 c. 外顯子 d. 內含子 e. 3'非轉譯區

青藍色：區域一(桃園縣、新竹縣、新竹市、苗栗縣) 棕色：區域二(臺中市、南投縣) 灰色：區域三(彰化縣、雲林縣)

黑色：區域四(嘉義縣、嘉義市、臺南市)