

# 知識天地

## 生物資訊、分子印記與肺癌個人化醫療

陳璿宇助研究員 (統計科學研究所)

肺癌在世界上<sup>1</sup>或是臺灣為主要的癌症死因，根據衛生署民國 96 年的生命統計資料顯示，惡性腫瘤占國人死因第一位，其中肺癌為男性癌症死因第二位，女性癌症死因第一位。民國 93 年癌症登記年報資料中指出，男性肺癌發生率為所有癌症的第二位，女性肺癌發生率為第四位。肺癌依組織細胞型態可分為小細胞肺癌 (small cell lung cancer; SCLC)與非小細胞肺癌 (non-small cell lung cancer; NSCLC)，其中以 NSCLC 為最常見的型態。在 NSCLC 中，可以細分成腺癌 (adenocarcinoma)和鱗狀細胞癌 (squamous cell carcinoma)等，以肺腺癌最多。肺癌之另一特徵是早發轉移，早期肺癌症 40%的病人會在術後 5 年內會發生轉移<sup>2</sup>，且目前臨床 TNM 分期<sup>3</sup>方法無法準確預測病人預後。

正常細胞發展至腫瘤細胞的過程相當複雜，主要可以從三種分子層次來探討，基因體組 DNA (genomic DNA)、DNA 轉錄 (transcription)成 RNA 以及 RNA 轉譯 (translation)成蛋白質<sup>4</sup>。在 genomic DNA 層次來說，致癌基因 (oncogene)可能在會有 copy number 增加，抑癌基因 (tumor suppressor gene)可能會有片段缺失 (deletion)的現象。參雜出現兩種現象時，正常細胞可能會逐漸朝向癌化模式，比較性基因體雜合技術 (comparative genomic hybridization; CGH) (包含三十多萬個基因體組探針)與單一核酸多型性 (single nucleotide polymorphism; SNP)晶片 (包含近百萬個 SNP 與拷貝數變異探針)為探討 DNA 變異情形的常用的研究工具。在 RNA 轉譯層次，致癌與抑癌基因在不同組織或是腫瘤細胞會有不同的基因表現強度，可以利用基因表現強度的差異，探討癌症的特徵或是預測病人的存活，此種研究大多利用同時測量大量基因表現的微陣列 (microarray)基因晶片 (包含上萬個基因探針)或 real time RT-PCR 進行 5-8。蛋白質轉譯的層次來說，現有一種新發現的微分子稱為微小核糖核酸 (microRNA)，其能與目標基因的 3'非轉譯區 (3'UTR)結合。結合後能抑制目標基因的 RNA 轉譯成蛋白質<sup>4</sup>。在現有研究中，許多利用 microarray 或是 real time RT-PCR 技術測量研究顯示利用基因表現能預測肺癌<sup>5-9</sup>存活或是復發率或是其他疾病<sup>10-12</sup>，亦有許多研究結果的基因表現與臨床資料已經公開於網路資料庫<sup>7-9, 13</sup>，能讓研究人員使用上述資訊進行癌症研究，並且有利用公開資料庫進行癌症研究的相關結果發表<sup>14-16</sup>。

由此可知，癌症為一個複雜的生物系統，其為何由正常組織發展成癌組織以及後續發生的癌轉移現象，目前仍為科學家努力解決的問題。隨著許多的高通量分析技術 (high throughput technology) 出現，研究人員可以利用這些新穎的生物技術以全基因尺度來檢視其研究，以提出更好且更全面的假說。然而，新技術伴隨的挑戰為更多的傳統生物研究者沒受過訓練來分析這成千上萬的資料。為了突破這層障礙，傳統生物研究者常求助於計算學家以闡明這些研究資料背後所隱藏的真正意義。事實上，最近許多經由這些新穎技術分析的研究報告，其中都含有許多強而有力的計算工具及複雜的統計或數學推論，以發展和完成癌化過程與癌轉移的計算模型，藉由數學模型與實驗生物學的整合，將會為生物學和癌症處理帶來新的思維和方法。整合性數理癌症生物學也因此將許多臨床和基礎癌症研究者與其他諸如數學、物理、資訊科技、影像科技及計算科學研究者，在針對癌症生物學的關鍵問題上共同研究<sup>17</sup>。

在肺癌分子印記與個人化醫療的相關研究中，以微小核糖核酸 (microRNA)為例，其為一類新的小核糖核酸，可以反向調節基因的功能。為研究影響肺癌存活率之相關微小核糖核酸，研究者利用即時反轉錄聚合酶鏈反應技術與生物統計演算法，共分析 172 例非小細胞肺癌患者的微小核糖核酸表現量，發現一組由五個微小核糖核酸組成的肺癌預後分子印記。利用此分子印記估計病人的風險分數，結果顯示高危險評分的肺癌患者，其存活率明顯地比低危險評分的病人短且復發率高，其死亡風險為低風險評分的人的 2.8 倍，發生復發的風險為 2.4 倍，上述研究已在 2008 年癌症醫學領域的頂尖期刊-癌細胞 (Cancer Cell)<sup>18</sup>刊出。此外，在 2007 年新英格蘭醫學期刊 (New England Journal of Medicine)刊出一個五個基因為基礎的肺癌預後預測印記<sup>5</sup>，利用 real time RT-PCR 方法測量 5 個基因表現，可以準確地預測癌症復發和肺癌患者的存活率。此結果不但可以應用在華人族群，套用至歐美族群亦有顯著的成果。該期社論 (editorial)評論上述研究利用現有病人組織以及臨床資料建立基因預測模式，在肺癌個人化治療研究中，完成了第一階段建立預測模式的工作<sup>19</sup>。此外，在 2008 年 Nature Medicine 期刊刊出的大規模多中心的肺癌研究中，也驗

證此五基因分子印記在早期肺癌存活預測的準確性<sup>8</sup>。此發現初步證明利用肺癌預後分子印記可以準確地預測癌症復發和肺癌患者的存活率，此項發現可應用於癌症分子病理學的研究或新標靶性治療的開發，更可進一步發展成癌症檢測試劑，以用於評估病患預後，並可有助於選擇高風險的癌症病人，在早期即進行輔佐性化療，或給予更進一步的治療。

在轉譯醫學的研究中，希望將臨床上遇到的問題，由基礎科學方法研究解決，再將研究成果應用於臨床治療，以達到個人化醫療的目的。在癌症病人的存活或是復發率的研究中，迄今仍未有準確的基因預測模式可以應用於臨床治療，其可能原因為欲分析的資料量太過龐大且複雜，且未有合適的數理計算模式可以套用，此為目前癌症轉譯醫學上的一大瓶頸。在國外已表的研究中，已有初步的成果顯示能利用公開於網路上的基因與臨床資料，去尋找癌症病人的預後基因印記，且結果也發表於有影響力的期刊。然而不同研究團隊找出的基因印記皆不同，且未經過實驗室中生物功能的驗證。因此，發展合適的數理統計方法，以全面性分析探討大規模的癌症基因體資料，並在實驗室中驗證基因印記的生物功能，為目前癌症轉譯醫學上相當重要的工作。以結合數理統計學家以及臨床與基礎的生物醫學家的研究策略，能發展新穎的數理計算方法，以分析龐大且複雜的多功能基因體資料，並以生物醫學的角度，詮釋與驗證發展出的新穎計算模式，才能將基礎研究的成果，在實際臨床上應用。

## 參考文獻

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. *CA Cancer J Clin* 2008;58:71-96.
2. Miller YE. Pathogenesis of lung cancer: 100 year report. *Am J Respir Cell Mol Biol* 2005;33:216-23.
3. Leslie HS. TNM: Evolution and relation to other prognostic factors. *Seminars in Surgical Oncology* 2003;21:3-7.
4. Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* 2006;6:259-69.
5. Chen H-Y, Yu S-L, Chen C-H, et al. A Five-Gene Signature and Clinical Outcome in Non-Small-Cell Lung Cancer. *N Engl J Med* 2007;356:11-20.
6. Potti A, Dressman HK, Bild A, et al. Genomic signatures to guide the use of chemotherapeutics. *Nat Med* 2006;12:1294-300.
7. Beer DG, Kardia SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002;8:816-24.
8. Shedden K, Taylor JMG, Enkemann SA, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008;14:822-7.
9. Potti A, Mukherjee S, Petersen R, et al. A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N Engl J Med* 2006;355:570-80.
10. O'Shaughnessy JA. Molecular Signatures Predict Outcomes of Breast Cancer. *N Engl J Med* 2006;355:615-7.
11. Bull TM, Coldren CD, Geraci MW, Voelkel NF. Gene Expression Profiling in Pulmonary Hypertension. *Proc Am Thorac Soc* 2007;4:117-20.
12. Izuhara K, Saito H. Microarray-based identification of novel biomarkers in asthma. *Allergol Int* 2006;55:361-7.
13. Bhattacharjee A, Richards WG, Staunton J, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 2001;98:13790-5.
14. Sun Z, Wigle DA, Yang P. Non-Overlapping and Non-Cell-Type-Specific Gene Expression Signatures Predict Lung Cancer Survival. *J Clin Oncol* 2008;26:877-83.
15. Lau SK, Boutros PC, Pintilie M, et al. Three-Gene Prognostic Classifier for Early-Stage Non Small-Cell Lung Cancer. *J Clin Oncol* 2007;25:5562-9.
16. Bianchi F, Nuciforo P, Vecchi M, et al. Survival prediction of stage I lung adenocarcinomas by expression of 10 genes. *J Clin Invest* 2007;117:3436-44.
17. Anderson ARA, Quaranta V. Integrative mathematical oncology. *Nat Rev Cancer* 2008;8:227-34.
18. Yu SL, Chen HY, Chang GC, et al. MicroRNA Signature Predicts Survival and Relapse in Lung Cancer. *Cancer Cell* 2008;13:48-57.
19. Herbst RS, Lippman SM. Molecular Signatures of Lung Cancer -- Toward Personalized Therapy. *N Engl J Med* 2007;356:76-8.