

学 位 論 文 題 名

cDNA アレイを用いた非小細胞性肺癌における
遺伝子発現解析

学位論文内容の要旨

背景と目的

非小細胞性肺癌は肺癌全体の約 80%を占め、その 5 年生存率は悪性腫瘍の中で最も低いものの一つであるが手術を含む治療法の選択に役立つ情報は貴重である。中でもリンパ節転移の有無は重要な要素となることから、術前にこれらが予測可能であれば、臨床上的有用性は高い。

最近の分子生物学の進歩は目覚ましく、多くの癌遺伝子や癌抑制遺伝子が発見されてきた。しかし、単一あるいは数個の遺伝子により与えられた転移の危険性や予後情報は、時には正しくない結果を招くこともある。

最近の新しい技術の一つであるマイクロアレイの開発によって、多くの遺伝子の発現変動や遺伝子異常を一度に検出することが可能になってきた。本研究では癌関連遺伝子を搭載した cDNA アレイを用いて、非小細胞性肺癌の遺伝子発現プロファイルによる予後因子の予測の可能性を検討した。

方法

(1) 患者と検体：北海道大学付属病院と関連 33 施設の協力により、2001 年 6 月から 2002 年 8 月に手術を受けた非小細胞性肺癌 92 例（扁平上皮癌 37 例、腺癌 55 例）から得た肺癌組織を対象とした。すべての症例につきインフォームドコンセントが得られた。当研究は北海道大学の倫理委員会において承認を得ている。

(2) 病理学的データ：臨床病理学的因子は各施設における病理診断を使用した。pT（原発腫瘍）因子とリンパ節転移の有無について検討した。

(3) cDNA array：摘出臓器より得られた凍結検体から Total RNA を抽出した。次に精製キットにより mRNA を精製した。mRNA 1 μ g から逆転写反応を行い、polyA を付加した。PolyA⁺ cDNA は biotin-16-dUTP とともに PCR で増幅した。biotin 標識 cDNA を 1,289 種類の癌関連遺伝子と 11 個の House keeping gene を搭載した array filter にハイブリダイズさせた。得られた画像から解析ソフトを用いてシグナル強度の数値化を行った。

(4) 発現プロファイル解析：アレイデータはハイブリダイゼーションなどの実験条件の違いにより補正が必要となる。遺伝子発現値の中でも比較的安定した部分を選択し、それらの発現値の平均値で全遺伝子発現値を除算して標準化を行った。

標準化を行った後に、2 群間で発現に有意差を認める遺伝子を選択した(two-sided t-test, $p \leq 0.05$)。これは発現に安定性の得られない遺伝子を除くため施行した。その遺伝子群の

発現値データで構成されるデータ部分空間が、期待される2群に実際に分離されるかを確認するために、EM アルゴリズムを行った。次に発現パターン分類に最適な遺伝子セットを選択するために特徴選択を施行した。パターンの抽出は順次特徴を追加し、Leave-One-Out-Error を評価して最適な遺伝子組み合わせを同定する Sequential Forward Selection 法を用いた。発現値から2群のクラス分けを行う識別器としては k-最近隣法を用いた。この結果、最も誤識別率の低い遺伝子組み合わせ集合からなる最適遺伝子セットを作成した。

一方、アレイデータの標準化を行った後に、腫瘍サイズに関連して発現値が増加あるいは減少する遺伝子を抽出するために、一般化線型モデルによる回帰分析を施行した。

結果

92 例の非小細胞性肺癌について遺伝子発現解析を行った。扁平上皮癌と腺癌において有意差のあった遺伝子は 205 個であった。このことから扁平上皮癌と腺癌は分子病態上大きく異なることが推測されたため、以後の解析は2つの組織型を分離して行った。

(1) リンパ節転移の有無：扁平上皮癌において手術時におけるリンパ節転移症例 16 例と転移なし症例 21 例の発現プロフィール差を検討した。その結果発現に有意差を認めた遺伝子は 45 個であった。EM アルゴリズムでは 2 例を誤って分類した (誤識別率=5.4%)。Sequential Forward Selection 法を用いて特徴選択を施行した結果、最適遺伝子セットとして 12~23 個の遺伝子が選択された。誤識別率は 0%で EM アルゴリズムにて正確に分類できなかった 2 症例も正確に識別できた。

同様に、腺癌において 60 個の遺伝子が有意差をもって選択された (リンパ節転移症例 19 例、転移なし症例 33 例)。EM アルゴリズムにて 6 例が誤って分類された (誤識別率=11.5%)。60 個の遺伝子から Sequential Forward Selection 法を用いて最適遺伝子セットを検討すると 5~43 個の遺伝子が選択され、誤識別率は 5.8%で EM アルゴリズムより識別率が改善された。扁平上皮癌で有意差のあった 45 個の遺伝子と腺癌の 60 個では一致して選択された遺伝子は一つだけであった。

(2) pT (原発腫瘍) 因子：pT 因子において有意差のあった遺伝子のうち扁平上皮癌と腺癌で共通であったものは 5 個であった。

つぎに腫瘍の進展に線形に関係する遺伝子を抽出するために、pT 因子につき一般化線形モデルを用いて回帰分析を行った。その結果、pT 因子に関連して扁平上皮癌で 231 個、腺癌で 75 個の遺伝子が選択された。これらのことから、腫瘍の進展を反映する遺伝子発現プロフィールの変化は、腺癌より扁平上皮癌で大きいことが示唆された。

考察とまとめ

本研究では、cDNA アレイによる非小細胞性肺癌患者の予後因子、特にリンパ節転移の有無における予測の可能性について検討した。今回の研究は腺癌と扁平上皮癌の両者におけるリンパ節転移の予測の可能性を示したはじめての論文である。

分類の特徴に最も関係があって必要なものを抽出するために特徴選択法を用いた。これらの方法により扁平上皮癌と腺癌でそれぞれ非常に重要な遺伝子を 23 個と 43 個選択した。これによりリンパ節転移の有無の分離は、EM アルゴリズムによる分類と比較して扁平上皮癌で 94.6%から 100%へ、腺癌で 88.5%から 94.2%へ増加した。

非小細胞性肺癌において CT によるリンパ節転移の診断率は約 70%程度である。今回の症例中、精密 CT 検査を施行した扁平上皮癌 15 例と腺癌 14 例のリンパ節転移の放射線専

門医による正診率を検討したところ、扁平上皮癌では 73.3%、腺癌では 64.3%だった。以上のことから、今回の研究でほぼ正確にリンパ節転移の予測が可能だったことは、術前の分子生物学的診断への実用化の可能性を示したものと考えられた。

今回の研究では予後の予測のほかに肺癌の分子生物学的進展を理解するために pT 因子の一般化線形モデルによる回帰分析を行ったが、腫瘍の進展に関係する遺伝子の抽出が可能であった。これらの選択された遺伝子を端緒として肺癌の悪性度を理解する研究がさらに進展するものと期待される。

腫瘍の悪性度を正確に予測することが可能であれば、術前の生検組織を用いて病態を予測することにより、リンパ節郭清の程度あるいは適切な手術選択の決定に用い得る。今後臨床の場で実際に使用するためには、大規模な前向き研究にて実際の診断率の確認と一般化のための確認作業が必要と思われる。

学位論文審査の要旨

主 査 教 授 加 藤 紘 之
副 査 教 授 吉 木 敬
副 査 教 授 守 内 哲 也

学 位 論 文 題 名

cDNA アレイを用いた非小細胞性肺癌における 遺伝子発現解析

非小細胞肺癌は肺癌全体の約 80%を占め、その 5 年生存率は悪性腫瘍の中で最も低いものの一つであるが手術を含む治療法の選択に役立つ情報は貴重である。中でもリンパ節転移の有無は重要な要素となることから、術前にこれらが予測可能であれば、臨床上的有用性は高い。今回、申請者は癌関連遺伝子を搭載した cDNA アレイを用いて、非小細胞性肺癌の遺伝子発現プロファイルによる予後因子の予測の可能性を検討した。

北海道大学病院と関連 33 施設の協力により、2001 年 6 月から 2002 年 8 月に手術を受けた非小細胞性肺癌 92 例（扁平上皮癌 37 例、腺癌 55 例）から得た肺癌組織を対象とした。すべての症例につきインフォームドコンセントが得られた。pT（原発腫瘍）因子とリンパ節転移の有無について検討した。摘出臓器より得られた凍結検体から mRNA を精製した。biotin 標識 cDNA を 1,289 種類の癌関連遺伝子と 11 個の House keeping gene を搭載した array filter にハイブリダイズさせた。得られた画像から解析ソフトを用いてシグナル強度の数値化を行った。標準化を行った後に、2 群間で発現に有意差を認める遺伝子を選択した(two-sided t-test, $p \leq 0.05$)。次に発現パターン分類に最適な遺伝子セットを選択するために特徴選択を施行した。パターンの抽出は順次特徴を追加し、Leave-One-Out-Error を評価して最適な遺伝子組み合わせを同定する Sequential Forward Selection 法を用いた。発現値から 2 群のクラス分けを行う識別器としては k-最近隣法を用いた。この結果、最も誤識別率の低い遺伝子組み合わせ集合からなる最適遺伝子セットを作成した。

一方、アレイデータの標準化を行った後に、腫瘍サイズに関連して発現値が増加あるいは減少する遺伝子を抽出するために、一般化線型モデルによる回帰分析を施行した。

結果として扁平上皮癌と腺癌において有意差のあった遺伝子は 205 個であった。

このことから以後の解析は 2 つの組織型を分離して行った。扁平上皮癌において手術時におけるリンパ節転移症例 16 例と転移なし症例 21 例の発現プロファイル差を検討した。その結果発現に有意差を認めた遺伝子は 45 個であった。EM アルゴリズムでは 2 例を誤って分類した (誤識別率=5.4%)。Sequential Forward Selection 法を用いて特徴選択を施行した結果、最適遺伝子セットとして 12~23 個の遺伝子が選択された。誤識別率は 0% で EM アルゴリズムにて正確に分類できなかった 2 症例も正確に識別できた。同様に、腺癌において 60 個の遺伝子が有意差をもって選択された (リンパ節転移症例 19 例、転移なし症例 33 例)。EM アルゴリズムにて 6 例が誤って分類された (誤識別率=11.5%)。60 個の遺伝子から Sequential Forward Selection 法を用いて最適遺伝子セットを検討すると 5~43 個の遺伝子が選択され、誤識別率は 5.8% で EM アルゴリズムより識別率が改善された。

つぎに腫瘍の進展に線形に関係する遺伝子を抽出するために、pT 因子につき一般化線形モデルを用いて回帰分析を行った。その結果、pT 因子に関連して扁平上皮癌で 231 個、腺癌で 75 個の遺伝子が選択された。これらのことから、腫瘍の進展を反映する遺伝子発現プロファイルの変化は、腺癌より扁平上皮癌で大きいことが示唆された。

以上の結果から扁平上皮癌と腺癌において遺伝子発現プロファイルの違いがみられたことから分子病態上大きく異なることが示唆された。非小細胞肺癌を扁平上皮癌と腺癌にわけて解析することで有力な予後因子の一つであるリンパ節転移の有無を高い精度で診断することが可能であった。このことは術前の分子生物学的診断への実用化の可能性を示していたものと考えられた。

口頭発表において、吉木教授より診断に用いたリンパ節の部位、扁平上皮癌・腺癌に特徴的な遺伝子、実用化に向けての課題について質問があった。ついで守内教授より臨床上の問題点、小細胞癌について質問があった。最後に加藤教授より腺癌と扁平上皮癌の分子生物学的な違いに対する考察、環境や喫煙との関連性について質問があったが、申請者はおおむね妥当な回答をした。

非小細胞肺癌における遺伝子発現プロファイルを解析し、リンパ節転移有無の予測、分子生物学的診断への実用化の可能性を示唆した本研究の意義は大きく、審査員一同協議の結果、本論文は博士 (医学) の学位授与に値するものと判定した。