

学位論文題名

大腸癌の遺伝子発現プロファイルパターン識別による  
リンパ節転移識別

学位論文内容の要旨

目的

大腸は癌化とその進展の分子機構の研究が進んでいる臓器の1つであるが、各段階での重要な遺伝子が癌の悪性度にどの程度関わるかは十分に理解されていない。予後や癌の悪性度と相関する遺伝子は報告されているが、より多くの遺伝子発現を包括的に判断する cDNA array 研究が注目されている。本研究では大腸癌症例を対象として、臨床病理学的因子を推測する遺伝子発現の同定を試みた。

方法

1. 対象症例と材料

北海道大学附属病院、および道内関連施設(計 33 施設)において 2001 年 6 月から 2001 年 12 月に施行された大腸癌手術症例 89 症例を対象とした。

2. 臨床病理学的因子

大腸癌取り扱い規約第 6 版(大腸癌研究会編)を参考にして、各種の臨床病理学的パラメータを集計した。

3. 遺伝子発現解析

腫瘍組織サンプルから total RNA を抽出し、mRNA を精製した。その後、逆転写反応、Biotin 標識、cDNA の増幅し、癌関連遺伝子 1289 個と 11 個の house keeping gene を搭載した cDNA アレイフィルターにハイブリダイズさせた。続いて、シグナルの検出を行い、画像化しシグナル強度の数値化を行った。

4. アレイデータの統計学的解析

各パラメータを様々な組み合わせで 2 群に群分けし、t 検定で発現に有意差( $p < 0.05$ )の見られた初期遺伝子を抽出した。それらの初期遺伝子から、特徴サブセット選択アルゴリズムを用いて最適診断遺伝子の抽出を試みた。抽出には順次特徴を追加し k-最近隣法により 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前方特徴選択法を用いた。癌の占拠部位については計 6 つのカテゴリ(C, A, T, D, S, R)があるので、これらの段階を追って発現が順次高くなる遺伝子と、逆に発現が低下する遺伝子を抽出するため一般化線形モデルを用いた回帰分析を行った。リンク関数としては logit を採用した。

## 結果

89 症例全体において、各臨床病理学的パラメータで最も多くの初期遺伝子群を得た組み合わせは、癌占拠部位の盲腸 C、上行 A、横行 T、下行結腸 D (CATD 群) 対 S 状結腸 S、直腸 R (SR 群) の 191 個であった。

これら結果から癌占拠部位に注目し、癌占拠部位 C、A、T、D、S、R の 6 つについて一般化線形モデルを用いた回帰解析を行ったところ、170 個の遺伝子(100 個が発現低下し、70 個が発現増加する遺伝子)が抽出された。この 170 個のうち 126 個が CATD 群と SR 群で分けたときの 191 個の遺伝子と共通であった。

また、12 例の CATD 群と 8 例の SR 群の大腸正常粘膜についても検討すると、70 個の遺伝子で発現プロファイルの差が見られた。その内、癌で差が見られた 191 個と共通だったのは 14 個であった。6 つの癌占拠部位の正常粘膜について線形回帰解析で 77 個の遺伝子(33 個が発現低下し、34 個が発現増加する遺伝子)が抽出された。正常粘膜で CATD 群と SR 群で分けたときの 70 個の遺伝子と 31 個が共通であった。癌で抽出された 170 個の遺伝子のうち 9 個が共通であった。

これらの結果より CATD 群と SR 群に分けて臨床パラメータを検討した。その結果、静脈侵襲、リンパ管侵襲、壁深達度 (m, sm, mp, ss 対 se, si)、リンパ節転移の有無、病期分類 (stages I, II 対 III, IV) の 5 つの臨床パラメータでより多くの異なる発現プロファイルが見られた。この 5 つのパラメータに対し最適遺伝子セットの抽出を試みると、各々の minimal error rate は静脈侵襲 (CATD 群: 0%、SR 群: 4.2%)、リンパ管侵襲 (CATD 群: 2.4%、SR 群: 8.3%)、壁深達度 (CATD 群: 0%、SR 群: 8.3%)、リンパ節転移 (CATD 群: 0%、SR 群: 4.2%)、病期分類 (CATD 群: 2.4%、SR 群: 10.4%) であった。

## 考察とまとめ

本研究ではパターン識別に特徴サブセット選択アルゴリズムを用いて大腸癌 89 症例の浸潤性に特徴的な遺伝子発現プロファイルの解析を行った。90%以上の正確さで予後因子の予測するために、どのようにして最適遺伝子セットを抽出するかが問題点であったが、大腸癌を癌占拠部位により CATD 群と SR 群に分けることにより癌の浸潤能を正確に予測することが可能になった。両群の正常大腸粘膜でも異なる遺伝子発現パターンを示したことは、癌占拠部位により特徴的な遺伝子発現パターンを呈したことに多少関連があると考えられ 191 個の異なる遺伝子発現と共通なのは 14 個であった。この結果からも癌の遺伝子発現の差は正常粘膜からの影響よりは癌化により起こる遺伝子変化が強く関係していると考えられた。大腸癌の発癌経路は分子遺伝学に、APC、K-ras、p53、DCC などの遺伝子異常の蓄積により発癌する多段階発癌説、いわゆる adenoma-cancer sequence とミスマッチ修復酵素遺伝子 (mismatch repair enzyme genes : MMR genes) の変異による MSI (Microsatellite Instability) の大きく 2 つに分けられている。CATD 群の癌はほとんど後者のタイプである。本研究では CATD 群と SR 群の癌の遺伝子発現プロファイルから分子病態生理学と癌発生メカニズムの違いが影響していることが明らかとなった。CATD 群と SR 群で異なる遺伝子発現を呈した 191 個の遺伝子には発癌や悪性度との関係が報告されている遺伝子、p53 との関係がある遺伝子、ミスマッチ修復酵素遺伝子 (mismatch repair enzyme genes : MMR genes)、APC-β-catenin と関係がある遺伝子が見られた。

大腸癌の浸潤能を予測する遺伝子セットの全ての遺伝子が必ずしも大腸癌の分子生物学的機構に関与しているかは定かではない。本研究で抽出された最適遺伝子セットはヒト遺伝子の一部分(1289 個)から選ばれた遺伝子であり、浸潤の特徴を決める複雑な遺伝子ネットワークの全体を判断するのは困難である。大腸癌の浸潤能の分子生物学的機構を明らかにするため、さらなる研究が必要であると思われた。

大腸癌につき CDNA array 解析を行った結果、占拠部位別にリンパ節転移および壁深達度を予測することによって高い診断率が得られたことから、本法が臨床上の治療指針を立てる上で極めて有用であることが明らかになった。

# 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

## 大腸癌の遺伝子発現プロファイルパターン識別による リンパ節転移識別

大腸は癌化とその進展の分子機構の研究が進んでいる臓器の1つであるが、各段階での重要な遺伝子が癌の悪性度にどの程度関わるかは十分に理解されていない。予後や癌の悪性度と相関する遺伝子は報告されているが、より多くの遺伝子発現を包括的に判断するcDNA array研究が注目されている。申請者はcDNA arrayとパターン識別に特徴サブセット選択アルゴリズムを用いて大腸癌の臨床病理学的因子を予測する遺伝子発現の同定を試みた。

本研究は北海道大学附属病院、および道内関連施設(計33施設)において2001年6月から2001年12月に施行された大腸癌手術症例89症例を対象とした。遺伝子発現解析は癌関連遺伝子1289個を搭載したcDNAアレイフィルターを用いてcDNA arrayを行った。アレイデータの統計学的解析は各パラメータを様々な組み合わせで2群に群分けし、t検定で発現に有意差( $p < 0.05$ )の見られた初期遺伝子を抽出し、それらの初期遺伝子から、特徴サブセット選択アルゴリズムを用いて最適診断遺伝子の抽出を試みた。抽出には順次特徴を追加しk-最近隣法により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前方特徴選択法を用いた。癌の占拠部位については計6つのカテゴリー(C, A, T, D, S, R)があるので、これらの段階を追って発現が順次高くなる遺伝子と、逆に発現が低下する遺伝子を抽出するため一般化線形モデルを用いた回帰分析を行った。

結果として89症例全体において、各臨床病理学的パラメータで最も多くの初期遺伝子群を得た組み合わせは、癌占拠部位の盲腸C、上行A、横行T、下行結腸D(CATD群)対S状結腸S、直腸R(SR群)の191個であった。これら結果から癌占拠部位に注目し、癌占拠部位C、A、T、D、S、Rの6つについて一般化線形モデルを用いた回帰解析を行ったところ、170個の遺伝子(100個が発現低下し、70個が発現増加する遺伝子)が抽出された。この170個のうち126個がCATD群とSR群で分けたときの191個の遺伝子と共通

であった。そこで CATD 群と SR 群に分けて臨床パラメータを検討した。その結果、静脈侵襲、リンパ管侵襲、壁深達度 (m, sm, mp, ss 対 se, si)、リンパ節転移の有無、病期分類 (stages I, II 対 III, IV) の 5 つの臨床パラメータでより多くの異なる発現プロファイルが見られた。この 5 つのパラメータに対し最適遺伝子セットの抽出を試みると、各々の minimal error rate は静脈侵襲 (CATD 群: 0%、SR 群: 4.2%)、リンパ管侵襲 (CATD 群: 2.4%、SR 群: 8.3%)、壁深達度 (CATD 群: 0%、SR 群: 8.3%)、リンパ節転移 (CATD 群: 0%、SR 群: 4.2%)、病期分類 (CATD 群: 2.4%、SR 群: 10.4%) であった。

本研究ではパターン識別に特徴サブセット選択アルゴリズムを用いて大腸癌 89 症例の浸潤性に特徴的な遺伝子発現プロファイルの解析を行った。90%以上の正確さで予後因子の予測するために、どのようにして最適遺伝子セットを抽出するかが問題点であったが、大腸癌を癌占拠部位によって CATD 群と SR 群に分けることにより癌の浸潤能を正確に予測出来ることが示された。

口頭発表において、吉木教授より大腸癌の部位別発生頻度、腫瘍マーカーとリンパ節転移の関係、リンパ節転移の有無に関し CATD 群と SR 群に共通な遺伝子はあるかなどについて質問があった。ついで守内教授より CATD 群と SR 群における P53 や APC との関係、臨床応用について質問があった。最後に加藤教授より予後との関係、正常大腸と癌との遺伝子発現の相関関係について質問があったが、いずれの質問に対しても、申請者は主旨をよく理解し誠意ある回答をしていた。

大腸癌において癌占拠部位別に臨床病理学的因子を識別することによって高い識別率が得られることを明らかにし、オーダーメイド治療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。