

学位論文題名

Prediction of lymphatic invasion/lymph node metastasis, recurrence, and survival in patients with gastric cancer by cDNA array-based expression profiling

(cDNA array による発現プロファイルを用いた胃癌患者のリンパ管浸潤, リンパ節転移, 再発および予後の予測)

学位論文内容の要旨

緒言

cDNA array を用い多くの遺伝子の発現を一挙に分析する発現プロファイル解析は、既存の知識に依存することなく、複雑な分子生物学的メカニズムを発見する有望な方法として注目されている。本研究では、胃癌54症例についてcDNAアレイによる遺伝子発現プロファイルを分析し、教師つき学習アルゴリズムを用いて、癌の悪性形質を反映する遺伝子の最適な組み合わせを得る試みを行い、患者の臨床転帰が予測可能かどうかを検討し、leave-one-out cross validation 法を用いて、再発と腫瘍による死亡のリスク予測について、最大の正診率を達成するような遺伝子組み合わせを抽出した。さらに、再発と死亡に関するリスク因子を評価することによって、リンパ節転移とリンパ管浸潤、深達度が最も重要な独立した因子であることを同定し、これらの特性を反映する特徴的な遺伝子を抽出した。これにより、胃癌患者におけるリンパ管浸潤・リンパ節転移、再発、生存についてのcDNAアレイデータによる予測に良好な成績を納めたので、抽出された特徴的遺伝子とともに報告する。

材料と方法

1.胃癌検体と患者の臨床データ

2001年7月から2002年2月までに、胃切除術を施行された計54症例の患者から得られた原発性胃癌の検体を用い、腫瘍の病態分類については、腫瘍の大きさ、腫瘍の局在、肉眼型、組織型、間質量、組織学的浸潤様式、リンパ管浸潤、静脈浸潤、深達度、リンパ節転移、肝転移、腹膜播種、遠隔転移、腹水細胞診の各項目を検討し、他の臨床パラメーターとして、性別、年齢、生存日数、再発までの日数、CEA値、CA19-9値も加えた。

2.RNA抽出とアレイのハイブリダイゼーション

各検体は液体窒素中で凍結状態のまま粉碎。total RNAを抽出し、さらにPoly(A)⁺ mRNAを抽出し逆転写反応を行いcDNAに変換したcDNAにpolyAを付加し、PolyA⁺ cDNAをbiotin-16-dUTPとともにPCRで増幅したBiotin標識cDNAを1289種類の癌関連遺伝子と11個のhouse-keeping geneを搭載したcDNAarrayに、ハイブリダイズし、化学発光を行い、画像化し、画像解析と数量化を行った。

3.統計学的手法

分割表分析を、各パラメーターですべての可能な組み合わせで施行し、それぞれのパラメーターについて、症例を2つのカテゴリー群に分類した。症例のデータの標準化を行ってから、全生存(腫瘍特異的死亡)と無再発生存(再発)

について、一般化 Wilcoxon 検定で予後不良群と予後良好群への 54 症例の分割を permutation 法で検討し、これら 2 つの予後相違群分類において、遺伝子発現値に統計学的に有意差のみられた遺伝子をアレイデータから選択した。続いて、特徴サブセット選択アルゴリズム(逐次前方選択法)を適用して、患者の転帰を最も良く反映する遺伝子の最適組み合わせを探索し、各段階の識別率の評価は、leave-one-out 交差検定による k-最近隣法を使用し、患者の転帰を予測において最大の正診率を得る遺伝子セットを得た。また 1,289 個の各遺伝子について発現中央値を基準として症例を高発現のもの、もしくは低発現のものに分け logrank 検定と一般化 Wilcoxon 検定を遂行し、有意な相関を示す遺伝子を抽出した。さらに、Cox 比例ハザードモデルを用いて回帰分析を行い、それぞれ同定された有意なハザードについて、症例を 2 つのクラスに分類し、そのクラスの間で異なった発現を示す遺伝子を抽出した。これらの遺伝子から患者の転帰を予測するために最適な遺伝子組み合わせを選択した。さらにその遺伝子セットを、上記の選ばれた遺伝子セットと比較検討した。

結 果

1. 腫瘍特異的死亡の予測

全生存期間において予後良好群と不良群とに分ける際の、発現値に有意差を示す 84 個の遺伝子を抽出し、さらに特徴選択アルゴリズムで最小の誤識別率を示す最適遺伝子セットを選択し、逐次選択した 3,570 の遺伝子組み合わせから、5.6%の最小の誤識別率を与える最適な予測遺伝子セットとして 31 および 32 個の遺伝子の組み合わせを選択した。

2. 腫瘍再発の予測

無再発生存期間において、最も適切な配分を permutation 法にて検索し、最小の p 値を与える 24 例づつの配分を得た。その 2 群間で有意な発現差を示す 96 遺伝子を選択し、特徴選択アルゴリズムにより、leave-one-out 交差検証法で 8.3%の最小の誤識別率を与える 7, 8, 9, 26, 27, 28 個の遺伝子を有する遺伝子セットを選択した。

3. 予後不良の転帰をとる病態生理学的背景の検討

すべての臨床病理学的パラメーターに対して Cox 比例ハザードモデルを用いて回帰分析を施行し、全生存及び無再発生存の双方において、最も重要な危険因子はリンパ管転移と深達度であることが確認された。

次に、リンパ管転移の存在と腫瘍の深達度について遺伝子発現プロファイルの特徴を検討し、深達度において m, sm, mp (40 例)と ss, se si (14 例)の 2 群に分類した時に、発現に有意差のみられた遺伝子が最も多く 113 個の遺伝子が抽出された。特徴サブセット選択にて、5.6%の最小の誤識別率を得る、20~27 個の遺伝子からなる 7 つの遺伝子セットが選ばれた。

同様にして、リンパ管転移に関して有意差のみられた遺伝子は 85 個の遺伝子であり、これらの中より特徴選択法にて選ばれた遺伝子セットを用いて、最小の誤識別率、9.6%を達成した。さらに、リンパ管侵襲に関しては、発現に有意差を認める 135 個の遺伝子を同定した。そして、4~45 個の遺伝子からなる 39 個の遺伝子セットを選択し、1.9%の最小の誤識別率を得た。

全生存において有意差のみられた 84 個の遺伝子は、リンパ管転移の有無において有意差のみられた 85 個の遺伝子とは、13 個の遺伝子を共有し、リンパ管侵襲の有無において有意差のみられた 135 個の遺伝子とは、12 個の遺伝子を共有していた。これらの遺伝子のうち 5 個の遺伝子が両者に共通する遺伝子であった。一方で、再発のリスクとして有意差のみられた 96 個の遺伝子は、リンパ管転移の有無において有意差のみられた 85 個の遺伝子とは、32 個の遺伝子を共有し、リンパ管侵襲の有無において有意差のみられた 135 個の遺伝子とは、35 個の遺伝子を共有していた。これらの遺伝子のうち両者に共通する遺伝子は 21 個の遺伝子であった。これらの結果はともに、全生存と無再発生存の両者における危険度に関連する遺伝子発現プロファイルが、少なくとも一部が 2 つの主要な因子、すなわち腫瘍の進行度とリンパ管組織への侵襲を反映する遺伝子発現プロファイルから構成されることを示している。

Real-time RT-PCR 定量

リンパ管転移の予測に重要な遺伝子の一つである clusterin と survivin の遺伝子発現の差を確かめるために real-time RT-

PCR 定量を行い、両者で発現の差を確認した。

考察と結語

本研究においては、胃癌全 54 症例につき、全生存と無再発生存に関して予後良好群と不良群とに識別する特徴遺伝子を検索し、90%以上の正確さで危険度を予測する事ができるそれぞれ32個と28個の遺伝子を同定することができた。さらに背景因子を分析することによって、関連した因子としてリンパ節転移と深達度、リンパ管浸潤を重要な因子として同定した。

本研究にて同定された遺伝子は、ヒト全体の遺伝子のごく一部分（1289 遺伝子は全体の 3.4%）からしか選ばれていないので、癌の浸潤性や悪性度を決定する、背景にある複雑な遺伝子ネットワークの全体像を評価するのは困難である。従って、本研究にて同定された遺伝子を手がかりとすることによって、胃癌における分子レベルでの浸潤性または悪性度の本態を理解するために、さらなる検討が行われなければならない。

学位論文審査の要旨

主 査 教 授 秋 田 弘 俊
副 査 教 授 守 内 哲 也
副 査 教 授 笠 原 正 典
副 査 教 授 近 藤 哲

学 位 論 文 題 名

Prediction of lymphatic invasion/lymph node metastasis, recurrence, and survival in patients with gastric cancer by cDNA array-based expression profiling

(cDNA array による発現プロファイルを用いた胃癌患者のリンパ管浸潤, リンパ節転移, 再発および予後の予測)

本研究では、胃癌 54 症例について cDNA アレイによる遺伝子発現プロファイルを分析し、教師つき学習アルゴリズムを用いて、癌の悪性形質を反映する遺伝子の最適な組み合わせを得る試みを行い、患者の臨床転帰が予測可能かどうかを検討し、leave-one-out cross validation 法を用いて、再発と腫瘍による死亡のリスク予測について、最大の正診率を達成するような遺伝子組み合わせを抽出した。その結果は全生存期間において予後良好群と不良群とに分ける際の、発現値に有意差を示す 84 個の遺伝子を抽出し、さらに特徴選択アルゴリズムで最小の誤識別率を示す最適遺伝子セットを選択し、5.6%の最小の誤識別率を与える最適な予測遺伝子セットとして 31 および 32 個の遺伝子の組み合わせを選択した。同様にして無再発生存期間においては、2 群間にて有意な発現差を示す 96 遺伝子を選択し、8.3%の最小の誤識別率を与える 7, 8, 9, 26, 27, 28 個の遺伝子を有する遺伝子セットを選択した。リンパ節転移に関して有意差のみられた遺伝子は 85 個の遺伝子であり、これらの中より特徴選択法にて選ばれた遺伝子セットを用いて、最小の誤識別率、9.6%を達成した。さらに、リンパ管侵襲に関しては、発現に有意差を認める 135 個の遺伝子を同定し、1.9%の最小の誤識別率を得た。すべての臨床病理学的パラメーターに対して Cox 比例ハザードモデルを用いて回帰分析を施行し、全生存及び無再発生存の双方において、最も重要な危険因子はリンパ節転移と深達度であることが確認された。

口頭発表において笠原正典教授より①施設間のデータ統合の標準化について②多様な

病態の胃癌を一括して扱うことの危険性についての質問があった。これらの質問に対して申請者は、胃癌に対する術式ならびにリンパ節郭清度はほぼ一致しており統一化できている。またデータの標準化は house keeping gene を使わずに global normalization にて行った。また②については今回は症例数が限られており、組織型に分けて検討するまでにはいたらず、今後の課題であると回答した。続いて守内哲也教授より①遺伝子の数が 30~40 個では多すぎるので減らすことは可能か②アレイの種類によって異なる遺伝子が抽出されているため、今後、アレイ以外の方法を使っていく必要があるのではないか③間質で重要な因子なのか腫瘍本体で重要な因子なのかを遺伝子でわけられないのかとの質問があった。これらの質問に対しては、今後、Real-time PCR などで遺伝子発現の定量を行い、発現に差のある遺伝子の絞り込みを行っていきたい。また、標準化については Real-time PCR による方法が考えられるが、今後の課題である。腫瘍細胞の近傍の間質や随伴炎症部、血管内皮や線維間質には腫瘍・間質相互に関与する遺伝子が発現しており、癌の浸潤増殖に関与すると考えられるため、本研究ではそれらも含めた検討が重要であると考えた。また血管新生因子などは間質で重要な遺伝子であるような印象があると回答した。続いて近藤 哲教授からは①今回 54 症例で 90%以上の予測結果が得られたが、別の症例群での確認はどうか②生検材料でのアレイ予測はできるのかとの質問があった。これについては、今回の検討は 54 症例のなかで学習と検証を行った結果 90%以上の予測が得られたにすぎず、それ以外の新たな胃癌症例を用いての検討が必要である。また 5 mm 角で切り出した検体ですら、mRNA を安定して採取するのは難しく、バイオプシーでは技術的に難しいのではないかと考えていると回答した。最後に秋田弘俊教授より、今回抽出された特徴遺伝子のなかで、興味深い遺伝子はなにかとの質問があったが、血管新生因子である VEGF や細胞増殖因子である Ki-67 に興味があると回答した。いずれの質問に対しても、申請者は主旨をよく理解し自らの研究データと文献的考察を混じえ誠意ある回答をした。

審査員一同は、これらの成果を高く評価し、大学院課程における研鑽や取得単位なども併せ申請者が博士(医学)の学位を受けるのに十分な資格を有すると判定した。