

氏 名 盛 啓太

学位(専攻分野) 博士(統計科学)

学位記番号 総研大甲第 1719 号

学位授与の日付 平成26年9月29日

学位授与の要件 複合科学研究科 統計科学専攻  
学位規則第6条第1項該当

学位論文題目 Cancer Outlier Analysis Based on Mixture Modeling of Gene  
Expression Data

論文審査委員 主 査 准教授 間野 修平  
教授 江口 真透  
助教 野間 久志  
教授 松浦 正明 帝京大学  
教授 松井 茂之 名古屋大学

論文内容の要旨

Summary of thesis contents

がんの発生には分子レベルの異常が関与していると考えられる。表現型が同様ながん疾患でも、分子レベルでは全く別の素因をもつという異質性が報告されている。近年、マイクロアレイとよばれる、多数の遺伝子について、同時に発現量を調べる手段が確立されている。がん細胞から採られた遺伝子発現量の標本において、一部の検体の発現量が特異的に高い、もしくは低いけれども、他については正常な標本と変わらないことがあることが知られている。この現象を遺伝的異質性といい、発現量が異常な検体の集合を cancer outlier とよぶ。このような異質性を前提として、がんに関連する遺伝子を選抜する方法が求められている。

従来の方法は 2 標本検定と多重比較の枠組みによるものが多い。2 標本検定は一様な差を検出することを想定しており、上記のような異質性の下では検出力が低くなるので、従来の方法は、2 標本検定を頑健にすることで異質性に対応する方向に発展してきた。ただし、2 標本検定には、選抜された遺伝子について、どの検体が cancer outlier であるかという医学的に重要な情報が得られにくいという短所がある。

本論文では、むしろ異質性を積極的に捉えるために、3 成分（それぞれの成分は正常な発現をする検体、高く発現する検体、低く発現する検体に対応）からなる正規混合モデルを導入し、がんに関連する遺伝子の選抜と、選抜された遺伝子について、各検体が cancer outlier とみなされるべき確率を与えることを提案している。

まず、最近の研究動向を参考にした典型的な正規混合モデルを仮定したシミュレーションを行い、2 標本検定の枠組みで提案された 5 つの従来法を合わせて提案法の性能を検証している。現実的な設定において、ほぼいずれの場合でも、提案法は所与の偽陽性率の水準に対して真陽性率が最も高いことが確認された。t-分布の混合モデルを仮定したシミュレーションでも同様な結論が導かれている。

さらに、骨髄異形成症候群の研究において得られたマイクロアレイ実験のデータを解析している。139 例の罹患者のサンプルと白血病でない 69 例についての遺伝子数 54,675 のデータに対して、5 つの従来法と提案法を適用し、それぞれの方法により選ばれた上位 200 遺伝子を比較した結果、提案法によって選択された遺伝子が 5 つの従来法により選択されたものと共通している場合が多いことが分かった。それにとどまらず、提案法で選択されるにも関わらず、従来法では選択されない遺伝子がいくつかあり、それらには既報により骨髄異形成症候群との関連が示唆されているものが含まれていたことが報告されている。

以上のように、本論文では、がん細胞の遺伝的異質性という最近のがん研究において重要な現象を対象とする統計学的課題について、従来の 2 標本検定のロバスト化とは異なる接近をしている。提案法によれば、通常の臨床研究において一般的な 2 標本検定の枠組みでは検出できない遺伝子を検出できる可能性があり、がんの生物学、がんを標的とする新薬開発への寄与も期待できると考えられる。

博士論文の審査結果の要旨

Summary of the results of the doctoral thesis screening

審査委員の全員の出席の下に、盛啓太氏の論文発表会、引き続いて審査委員会が開かれた。その結果、提出された博士申請論文は学位授与の水準に達していると判断され、本審査の合格を推薦する結論に達した。

原稿は7章から成り、1章の要旨、2章の序論では、がんを標的とする新薬開発の最近の研究動向が説明され、特に本研究の背景が準備されている。3章では cancer outlier の定義が与えられ、その検出の従来法について、次章の主要な結果と比較・検討するための詳細なレビューがされている。ここで、cancer outlier とは、がん細胞から採られた遺伝子発現のサンプルにおいて、一部のグループが特異的に高発現、または低発現しているが、残りのサンプルは正常細胞と同程度の発現をしているというがんサンプルにみられる遺伝的異質性をいう。特に、cancer outlier profile analysis に代表されるような 2 標本検定のロバスト化を図った一連の方法について紹介している。4章では、主要な結果である方法の提案と、その統計的性質について考察している。まず、従来法の限界として、2 標本検定の枠組みでは有用な情報が捉えきれていないことが指摘され、混合モデルの適用が提案されている。提案法は、正常細胞の遺伝子発現のサンプルよりがん細胞の遺伝子発現を標準化し、3成分の正規混合モデルを導入することで、がんサンプルの異質性を積極的に捉えることを狙うものである。ここで、3成分は、正常な発現、高発現、低発現のグループである。Cancer outlier の影響を受けないことを目指したロバスト 2 標本検定では、がんサンプルの異質性を利用できないことが主張されている。さらに、事後確率に基づく遺伝子選択のための統計量を提案している。5節では、シミュレーションの結果が説明されている。シミュレーションでは、2章で紹介された最近の研究動向を参考にした典型的な正規混合モデルが仮定され、3章で紹介された5つの従来法を合わせて提案法の性能を検証している。現実的な設定において、ほぼいずれの場合においても、提案法は所与の偽陽性率の水準に対して真陽性率が最も高いことが確認された。t-分布の混合モデルを仮定したシミュレーションでも同様な結論が導かれた。6章では、骨髄異形成症候群の研究において得られたマイクロアレイ実験の実データ解析について説明されている。139例の罹患者のサンプルと白血病でない69例についての遺伝子数54,675のデータに対して、5つの従来法と提案法を適用した。それぞれの方法により選ばれた上位200遺伝子の比較を行った結果、提案法によって選択された遺伝子が5つの従来法により選択されたものと共通している場合が多いことが分かった。それにとどまらず、提案法で選択されるにも関わらず従来法では選択されない遺伝子がいくつあることが指摘されている。8章では全体の考察にあてられており、今後の発展の方向性、提案法の改善の余地などが検討されている。

以上に説明したように、本博士申請論文では、がん細胞の遺伝的異質性という最近のがん研究において重要な現象を対象とする統計学的課題について、従来のロバスト化の考えを超えた興味深い考察がなされている。混合モデルに基づく提案法によれば、通常の臨床研究において一般的な2標本検定の枠組みでは検出できない遺伝子を検出できる可能性がある。実際、本研究ではそのような遺伝子の検出に成功しており、がんの生物学、がんを標的とする新薬開発への寄与も期待できる。本研究の内容は、査読付き学術雑誌に掲載されている。これらの理由により、審査委員会は、提出された博士申請論文は学位授与の水準に達していると判断した。