

Title	Classification of Cancer and Inference of Cancer-specific Gene Regulatory Networks Based on Gene Expression Profiles Using Soft Computing Rules(Abstract_要旨)
Author(s)	Wang, Xiaosheng
Citation	Kyoto University (京都大学)
Issue Date	2010-03-23
URL	http://hdl.handle.net/2433/120370
Right	
Type	Thesis or Dissertation
Textversion	none

(続紙 1)

京都大学	博士 (情報学)	氏名	王 小晟
論文題目	Classification of cancer and inference of cancer-specific gene regulatory networks based on gene expression profiles using soft computing rules 遺伝子発現プロファイルの柔計算解析に基づく腫瘍分類および腫瘍特異的遺伝子ネットワークの推定		
(論文内容の要旨)			
<p>本論文は遺伝子発現プロファイルに基づくがんの分類とがん特異的遺伝子制御ネットワーク推定のための機械学習とデータマイニングの技法について述べており、全5章と付録3部により構成される。</p> <p>第1章は序章であり、研究の背景と本論文の概要を示している。</p> <p>第2～4章が本論文の主要な内容を占め、以下のように要約される。</p> <p>第2章では、マイクロアレイ技術により得られる遺伝子発現量データに基づくがん分類のための特徴量選択について広範な調査を行っている。極めて多数の特徴量、この場合は個々の遺伝子の発現量からの分類のために最適なものを少数選択する手法の確立は、高い自由度と信頼性をもつ分類を行うために必須である。この目的のために従来様々な方法が提案されてきた。ここでは、ラフセット理論で用いられてきた「依存度」の概念を拡張した「α-依存度」を用いる方法を新たに提案している。特徴量選択のために従来用いられてきた他の5つの手法と比較した実験の結果、提案手法は他と同等かそれ以上の安定性と応用性を持つことが示されている。また、最適なα値は、その分類がもつ本来的な困難さと高い相関をもつことも示されている。</p> <p>第3章は、遺伝子発現量に基づくがんと正常、あるいはがんの種別間の簡略な規則分類手法について述べている。この目的のために、従来、多数の遺伝子の発現量を用いた複雑な分類手法が開発されてきた。これらの手法の多くは手続きがブラックボックス化しているため、仮によい分類結果が得られたとしても、それを解釈することがしばしば困難であった。第2章で述べたα-依存度を利用する手法を用いて、一つかたかだか二つの遺伝子を選択し、それに簡単な決定規則を適用する分類手法を提案している。公開されている遺伝子発現量データに対してその手法を適用した場合、従来法に比べて少なくとも同等の分類能力をもつことが示されている。さらに、そのようにして選択された遺伝子は、がんの病理と強く関連していることも示されている。</p> <p>第4章は、遺伝子発現プロファイルに基づくがん特異的遺伝子制御ネットワークの推定について述べている。遺伝子ネットワークの擾乱が発がん過程に関与することは広く認められているため、そのようなネットワーク推定は重要な研究課題である。ここでは、ラフセット理論を応用した柔計算規則に基づく有向な遺伝子制御ネットワークを推定する方法を提案している。その方法の適用例として、特定のがん(大腸がん)に関連する重要な遺伝子間の遺伝子制御ネットワークの再構築を試みている。その結果、(i)がんが発現が昂進している遺伝子は発現が抑制されている遺伝子よりもより多くの他の遺伝</p>			

子による制御を受ける、(ii)がん抑制遺伝子はがん遺伝子を抑制し、一方、他のがん抑制遺伝子を強く活性化させる、(iii)がん遺伝子は他のがん遺伝子を活性化させ、がん抑制遺伝子を弱く抑制する、などの知見が得られた。これらは、がん化に伴う遺伝子発現の変化が生物学的な系として安定であることを示唆し、がんの病理に有用な洞察を提供している。

最後に、第5章は全体のまとめと今後取り組むべき課題についての議論に当てられている。なお、付録A～Cはそれぞれ2～4章の内容に関する付随的な結果をまとめたものである。

(論文審査の結果の要旨)

本論文は、遺伝子発現プロファイルに基づくがんの分類とがん特異的遺伝子制御ネットワークの推定のために、申請者が開発した情報解析手法について記述している。主な成果は以下の3点である。

(1) ラフセット理論に基づく、新たな特徴量選択手法の開発

アレイ技術により発現量が測定された極めて多数の遺伝子の中から少数の適切なものを選択する手法は、信頼性の高いがん分類のための前処理として必須である。ここでは、従来ラフセット理論で用いられてきた「依存度」の概念を拡張した「 α -依存度」を新たに導入し、様々な特徴量選択手法と分類手法とを組み合わせた多くの実験によって提案手法の有効性を示している。

(2) 遺伝子発現量に基づくがんの単純な規則分類手法の開発

がんと正常、あるいはがんの異なる型を遺伝子発現プロファイルに基づき分類する方法は従来数多く開発されてきたが、その結果の解釈はしばしば困難であった。申請者は、(1)で述べた手法を利用して得られたごく少数の遺伝子に簡単な規則を適用することにより、従来法に劣らぬ分類能力が得られることを示している。また、選択された遺伝子がバイオマーカーとして有効である可能性を示している。

(3) 遺伝子発現量に基づくがん特異的遺伝子制御ネットワーク推定法の開発

ここでは、因果関係を示す有向な遺伝子制御ネットワークをラフセット計算規則に基づき推定する方法を提案している。この方法を実際の大腸がんのデータに適用し、がん化に伴い発現が昂進する遺伝子と抑制される遺伝子のグループの間に特徴的な制御関係が存在するという知見を得ている。

以上、本論文はがんの分子分類にラフセット理論を適用した新たな手法を提案し、また、同様の理論を応用した遺伝子制御ネットワーク推定の新たな枠組みを示すものである。これらは、機械学習・データマイニング手法の一応用にとどまらず、それを実データに適用することにより、がんの診断、治療、予後予測に関わる、医学的に重要な知見をもたらす可能性を秘めている。バイオインフォマティクス・システムバイオロジーの実用的な応用として当該分野に十分なインパクトを与えることが、公表論文のアクセス数の多さからも推測されている。よって、本論文は博士(情報学)の学位論文として価値あるものと認める。

また、平成22年2月22日実施した論文内容とそれに関連した試問の結果合格と認めた。