

Seeds

キーワード: 遺伝情報解析、音声解析、感性情報解析、データマイニング、知的処理
大規模遺伝子発現データからの知識抽出法の開発

Yoshifumi Okada



しくみ解明系領域・情報システム学ユニット

おかだ よしふみ

岡田 吉史 教授

Phone:0143-46-5408 Fax:0143-46-5408

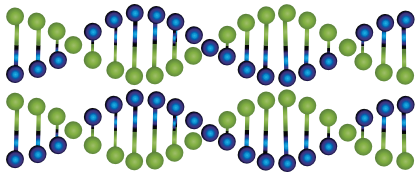
E-mail:okada@csse.muroran-it.ac.jp

URL <http://cb12.csse.muroran-it.ac.jp/>



膨大な生命データから『知』を発掘

研究の目的



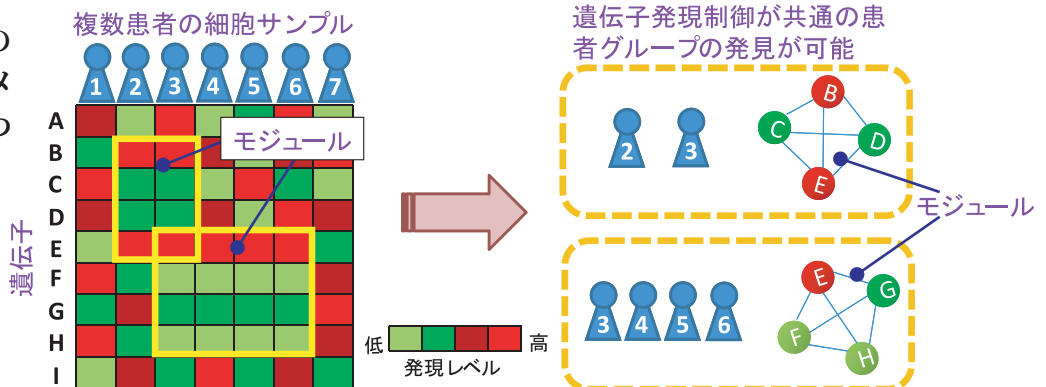
近年、DNAチップ技術の進展により、様々な実験条件下における遺伝子発現データが大量にもたらされ、GEO (Gene Expression Omnibus)などの公共データベースに急速に蓄積されつつある。癌などの難病に関する大規模遺伝子発現データを対象として疾患の原因となっている遺伝子群を高精度で発見する情報解析技術の開発を目指す。

研究の概要

超高速
網羅的
に解析する

従来、大規模な遺伝子発現データを対象としてモジュールを見つけ出すには膨大な組み合わせのチェックが必要であったため現実的な時間内に探索することは不可能であった。本来膨大な商品購買データを対象として複数の顧客から同時に購入される商品群(飽和集合)を見つけ出す方法として用いられる「飽和集合列挙法」と呼ばれる高速データマイニングを置き換え利用することにより超高速かつ網羅的にモジュールを発見することに成功。飽和集合の網羅列挙アルゴリズムとして世界最速のLCM (Linear time Closed itemset Miner)を用い、更に独自のデータ圧縮技術により、高速化と省メモリを実現している。

病気サブタイプの
発見やテーラード
創薬に役立つ



Seeds 大規模遺伝子発現データからの知識抽出法の開発

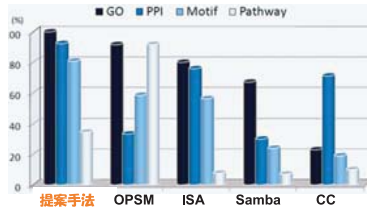
研究(開発)のアピールポイント

◆研究の新規性、独自性

飽和集合列挙法により、現実的な時間内でモジュールの網羅探索を実現。

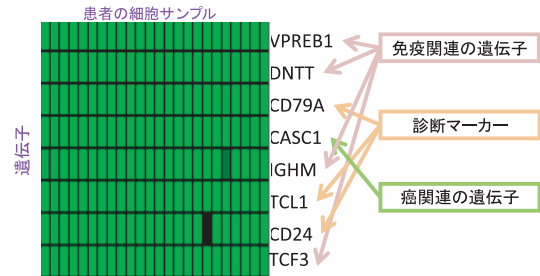
モジュールの遺伝子機能解析結果

生物学的に意味があると判定されたモジュールの割合は、総合的にみて提案手法が最も高い



◆従来研究(技術)と比べての優位性

確率論的またはヒューリスティックに基づく方法は近似解としてのモジュール探索であったが、網羅探索により高精度で生物学的に有用なモジュール抽出が可能となった。



◆研究に関連した特許の出願、登録状況

出願中

研究(開発)のビジョン、ステージ

◆適応分野

病気の原因遺伝子(バイオマーカー)の探索、癌や精神病のサブタイプの発見、再生医療研究における遺伝子スイッチの発見を支援するツールとして利用可能。

◆研究のステージ

基礎研究 **応用段階**

◆製品化、事業化のイメージ

ソフトウェアの開発、情報推奨システムの開発、ウェブアプリケーションの開発。



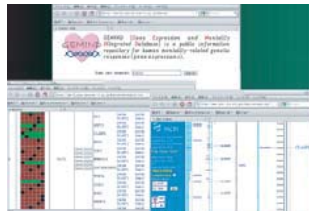
企業等へのご提案、メッセージ

◆研究(開発)に関連して、あるいはそれ以外に関わる業務

生理データ解析、音声識別、情報推薦技術などの分野に特化した新規データマイニング技術の開発。

◆利用可能な設備、装置など

心理状態に関わる遺伝子
モジュールデータベース



乳児の病理音声
診断システム



◆教員からのメッセージ

遺伝子情報、音声、しぐさ、顔色など生き物が発する情報を解析するデータマイニング法を研究しています。データ解析でお悩みの方は、お気軽にお声かけください。また、実用化へ向けた共同研究もできればと思っております。

岡田 吉史

