

主 論 文 要 旨

報告番号	甲 乙 第	号	氏 名	京田 耕司
主 論 文 題 目： Algorithms for Deducing Gene Networks Modeled as Signed Directed Graphs from Gene Expression Profiles of Gene Deletion Mutants (遺伝子欠失株の遺伝子発現データから符号付有向グラフで表現された遺伝子ネットワークを導出するアルゴリズムに関する研究)				
(内容の要旨) 細胞機能の理解のためには、遺伝子ネットワークの特定が必須である。遺伝子欠失株からの大規模遺伝子発現データは、遺伝子ネットワークを特定する上で有用な材料である。遺伝子欠失に対する遺伝子発現の変化を説明できる符号付有向グラフは、遺伝学および細胞生物学で使用されている遺伝子ネットワークの表現法であるが、このような遺伝子発現データに無矛盾な符号付有向グラフで表現された遺伝子ネットワークを導出する実用的な手法は存在しなかった。本研究では、遺伝学および細胞生物学で一般に使用されている遺伝子間制御関係の導出規則を利用して、遺伝子発現データに無矛盾な符号付有向グラフで表現された遺伝子ネットワークを導出する手法を開発した。 最初に、遺伝子発現データに無矛盾で冗長な弧を含まない符号付有向グラフで表現された遺伝子ネットワークを導出する DBRF(Difference-Based Regulation Finding)法を開発した。従来の遺伝子ネットワーク導出法では、遺伝子発現データが離散値である必要があったので、実際の遺伝子発現データに直接適用することが困難だった。DBRF 法では、遺伝学および細胞生物学で共用されている導出規則に基づき、連続値を持つ遺伝子発現データから直接、促進・抑制の制御関係を表現する正・負の弧を、野生型と遺伝子欠失型の遺伝子発現量の差を利用して導出する。次に、それらの弧において間接的な制御関係を表現する冗長な弧を削除する。DBRF 法と二値データを用いた手法の性能の評価を、ネットワークの大きさと最大入力遺伝子数を変化させた人工的な遺伝子ネットワークを構築し、そこから得られた連続値を持つ擬似遺伝子発現データとその二値化データを用いて行った。遺伝子発現データを二値化する手法と比較して、連続値を持つ遺伝子発現データを用いて直接、制御関係を導出する DBRF 法の性能が、感度・特異度の点で優れていることを明らかにした。 次に、遺伝子発現データに対し無矛盾で最少数の符号付弧で構成される符号付有向グラフで表現された遺伝子ネットワークを導出する DBRF-MEGN(Minimum Equivalent Gene Network)法を開発した。DBRF 法は無冗長な符号付有向グラフを多項式時間で導出できるが、導出された符号付有向グラフは最少数の符号付弧では構成されていない。最少数の符号付弧で構成される符号付有向グラフの導出には、指数関数時間の組合せ探索が必要となるため、大規模遺伝子発現データに適用できなかった。DBRF-MEGN 法では、組合せ探索の対象となる要素を削減・分割することで、飛躍的に計算量を低減させることに成功した。これにより、DBRF-MEGN 法の出芽酵母における 265 個の遺伝子欠失株に対するこれら 265 個の遺伝子の発現データへの適用が可能となった。導出されたフェロモン応答経路、アミノ酸制御システム、銅・鉄恒常性システムにおける制御関係を、文献からの制御関係と比較して、DBRF-MEGN 法の性能を評価した。また、導出した遺伝子ネットワークに含まれる制御関係に対して、遺伝学および細胞生物学で従来用いられてきた解釈法を利用することにより、18 個の転写制御遺伝子に対する計 132 個の転写標的遺伝子と転写活性調節遺伝子を予測した。 本研究では、DBRF 法と DBRF-MEGN 法の 2 手法を開発した。実データの解析には DBRF-MEGN 法を利用することが望ましいが、その適用性は遺伝子発現データの規模と複雑性に依存する。適用する遺伝子発現データに応じて 2 手法を使い分けることで、網羅的な細胞機能の解明が可能となる。				