

氏 名	さ とう てつ や 佐 藤 哲 也
学位(専攻分野)	博 士 (理 学)
学 位 記 番 号	理 博 第 3176 号
学位授与の日付	平 成 19 年 3 月 23 日
学位授与の要件	学 位 規 則 第 4 条 第 1 項 該 当
研 究 科 ・ 専 攻	理 学 研 究 科 生 物 科 学 専 攻
学 位 論 文 題 目	共進化情報を利用したタンパク質間相互作用の解析

論文調査委員	(主 査) 教 授 金 久 實 教 授 藤 吉 好 則 教 授 七 田 芳 則
--------	--

論 文 内 容 の 要 旨

本論文において申請者は、タンパク質間相互作用の予測法であるミラーツリー法を改良した新規予測法の提案を行い、予測精度の向上を目指している。一方で、新規予測法が示す擬陰性（実際には相互作用するにもかかわらず相互作用しないと予測すること）にも注目し、それが生じる原因について詳細にまとめている。

本論文の前半部において、まず先行研究であるミラーツリー法の問題点を分析している。ミラーツリー法では、相互作用しているタンパク質ペアは共進化するという仮定に基づき、共進化の強度から相互作用を予測する。この際、共通の生物種集合から得られるオーソログ配列間の遺伝的距離を要素とした距離行列の間の相関係数を、予測したい2つのタンパク質間の共進化の強度とみなす。このミラーツリー法の長所としては、アミノ酸配列の情報だけを使って相互作用予測を行える点が挙げられるが、一方、短所として、予測結果に擬陽性（実際には相互作用しないにもかかわらず相互作用すると予測すること）が多く含まれることが指摘されている。申請者は、この擬陽性の原因を、オーソログ配列に由来する生物種の進化的関係の情報が距離行列間で共有されていることにより、相関係数が高く評価される傾向があるためではないかと考え、距離行列からオーソログ配列に由来する生物種の進化的関係の情報を除去した後の残差情報のみを利用して、共進化の強度を評価することを試みた。本論文では、この作業仮説は、(1)射影演算子を用いて残差情報を求め、得られた残差間の相関係数を共進化の強度として予測を行う手法、(2)多重回帰モデルの残差間の相関係数である偏相関係数を共進化の強度として予測を行う手法、の2つの手法として実装している。これらの予測法の精度に関して、相互作用タンパク質ペアの情報を蓄積している DIP データベースに登録されている、大腸菌由来の相互作用タンパク質821ペアを用いて評価を行った結果、いずれの手法においても、ミラーツリー法に比べて擬陽性を大幅に減少できることが確認された。

一方、本論文で提案された手法はいずれも、ミラーツリー法に比べ、擬陰性、すなわち低い相関係数しか示さない相互作用ペアが増加していた。そこで、本論文の後半部においては、共進化の強度が低くなる原因を、2つの異なる観点から調査している。申請者は、まず複数の相互作用パートナーの存在が、共進化の強度に影響を与えるかもしれないと考え、1つ目の観点として、相互作用パートナーの数と共進化の強度との間の関係について注目した。そこで、タンパク質間相互作用ネットワークを構築する相互作用ペアについて共進化の強度を評価した結果、共進化の強度は、相互作用パートナーの共有度と関連することが明らかにされた。したがって、相互作用パートナーをお互いに共有せず、独立してパートナーを多く抱えている相互作用ペアでは、共進化の強度が低くなる可能性がある。このようなペアが予測において擬陰性の原因になる可能性が指摘された。2つ目の観点として、タンパク質は相互作用境界面を介して相互作用することから、複合体の境界面の構造的な特徴と共進化の強度との間の関係について注目した。立体構造が決定されているタンパク質ペアについて共進化の強度を評価した結果、共進化の強度が、相互作用境界面の面積、および平面性などの構造的な特徴と関連することを見出した。すなわち、共進化の強度が低くなる相互作用ペアは、相互作用境界面の面積が小さく、境界面の形状も平面に近い特徴を示すペアである可能性が高かった。このようなペアも擬陰性の原因となることが示唆され、本手法の予測性能向上のために改

善可能な問題点を見出すことに成功している。

論文審査の結果の要旨

タンパク質間相互作用の予測は、その予測精度向上が実験研究者からもバイオインフォマティクス研究者からも求められている重要な研究テーマとなってきた。時代の流れに即した本論文のテーマは、研究の意義と重要性が十分にあるものと考えられる。

本論文において申請者は、まず先行研究であるミラーツリー法の問題点の原因とその解決策を明確に述べている。その過程で、ミラーツリー法の問題点である擬陽性の原因として、共進化の強度を測る相関係数が過度に評価されている点に申請者は注目した。同じ生物種集合から集められたオーソログ配列の中には、それら配列に由来する生物種の進化的情報が共通の情報として含まれている。この共通の進化的情報を除去すれば、過度に評価されていた相関係数を低く抑えることができ、結果として擬陽性を減少させることに繋がる可能性がある。この点に注意し共進化強度を評価する新しい手法を用いることによって、予測精度の向上が期待できる。

次に、この研究仮説を実証するために、2種類の新規手法、(1)射影演算子を用いる方法と(2)偏相関係数を計算する方法を開発し、タンパク質間で観察される共進化の強度の指標を変更することが、予測精度に与える影響を評価した。これらの手法は、それぞれ異なる数学的理論に基づいて開発されているが、オーソログ配列に由来する生物種の進化的関係の情報を除去することが可能であるという点で、上記の研究仮説を検証するにはよいアルゴリズムである。検証には、オーソログ配列が多く存在する大腸菌由来タンパク質の相互作用821ペアの情報が用いられ、その結果、本論文で開発した手法はいずれも擬陽性を大幅に減少させる性能があることが示された。これらの結果は、共通の進化的情報を除去し共進化強度の相関係数を適切に設定すれば擬陽性が減少するという上記研究仮説に、一定の妥当性があることを支持するものである。一方、本論文で開発した2種類の手法間で予測精度に違いが見られたが、この違いについては、本論文では、アルゴリズムの違いに起因する観点から詳細な議論が行われている。

このように、今回開発した手法は予測において擬陽性を減少させることに成功しているが、一方で、擬陰性を多く生じさせるという結果ももたらしている。そこで、申請者は、新規手法がもたらす擬陰性の原因として、相互作用パートナーの数と相互作用境界面の構造的な特徴に注目し、解析を行った。その結果、これらの因子が共進化の強度と統計的に有意な結果で関連していることが示された。このようにさまざまな解析によって特徴付けられた相互作用ペアの情報から、擬陽性の原因となる特徴が、共進化情報を扱う予測法の適応限界の理由となりうることなど、今後、このような予測法の予測精度向上に向けて、改良すべき点を洗い出すことが出来る。したがって、今後のタンパク質間相互作用研究の発展に大いに貢献することが期待される。

以上のように、本論文では、従来手法の問題点の究明からその改善のための研究仮説の設定、実証および検証に至るまでが、論理的かつ明確に述べられている。また、開発した手法の特徴（長所と短所）についても十分な議論が展開されている。

よって、本論文は博士（理学）の学位論文として価値あるものと認める。また、論文内容とそれに関連した事項について試問を行った結果、合格と認めた。