

氏名	伊藤真純
学位(専攻分野)	博士(理学)
学位記番号	理博第3212号
学位授与の日付	平成19年11月26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	ゲノム比較による真核生物のタンパク質ドメイン組み合わせの特異性と進化解析

論文調査委員 (主査) 教授 金久 實 教授 七田芳則 教授 岡 穆宏

### 論文内容の要旨

タンパク質は複数のドメインからなり、真核生物のタンパク質が持つ多様なドメイン構造は、真核生物の複雑な高次機能の実現に大きく寄与していると考えられている。本論文において、申請者は真核生物の各系統グループに特異的にみられるドメイン構造に着目した比較ゲノム解析を行っている。

本論文では、データセットとして47の真核生物のゲノム配列にみられる全タンパク質セットを用いている。申請者は、ドメイン構造について、同一のタンパク質上で隣接するドメインの組み合わせに着目して解析しており、各生物種のタンパク質セットにみられるドメイン構造から各生物種の持つドメインとその組み合わせを定義している。このうち、原核生物と真核生物で共有するものを除き、真核生物にのみ存在するドメインと組み合わせについて真核生物内での系統グループ特異性について調べた。系統グループは47種の真核生物の系統関係に基づき階層的に定義されており、系統グループ特異性はドメインおよび組み合わせの出現時期から推定している。この際、水平伝播や異なる系統で独立して出現する同一の組み合わせについて考慮する必要があるため、Mirkinらによって開発されたオーソログ遺伝子の系統樹上での出現・欠失の最節約的なシナリオの推定法を用いている。申請者は、二分木を前提に開発されたMirkinらの手法を多分木に拡張し、系統関係を多分木として表現することで、系統関係の曖昧な種や、ドラフト段階のゲノムデータの影響を緩和している。

このようにして同定された系統グループ特異的なドメインおよび組み合わせについて、その特徴の比較を行っている。動物および後口動物に多数の特異的な組み合わせが観察されており、その組み合わせに類出するドメインは動物と後口動物では異なるものであった。動物特異的な組み合わせには、チロシンリン酸化に関わるドメインが類出していたのに対し、後口動物特異的な組み合わせには補体・凝血系に関わるドメインが類出していた。これらの機能はそれぞれの共通祖先で進化したと考えられるものであり、この結果は、真核生物における高次機能の進化とドメイン構造の複雑化の関連性を示唆するものである。

次に、申請者はドメインのグループ特異性と出現時期を関連づけ、出現時期と組み合わせにみられる頻度について調べている。その結果、動物以外の系統では出現時期の古いドメインほど多くの組み合わせを形成する傾向があることが観察された。この傾向は、原核生物との比較が主であった先行研究においては観察されなかったものである。一方で、後口動物を含む動物の系統では、動物特異的なドメインが非常に多くの組み合わせを形成していた。後口動物の分岐後に出現した多数のドメインが形成する組み合わせが少数であることから、後口動物にみられる組み合わせの増加には、動物特異的なドメインの出現が大きく寄与していることが示唆される。

さらに、ドメインと組み合わせの出現時期から、ドメインの組み合わせネットワークの変化を推測し、ヒトおよび出芽酵母の系統についてネットワークの進化過程の比較を行っている。その結果、ドメインの組み合わせネットワークの変化は、それぞれの系統に特異的なドメインと組み合わせの出現とその欠失を反映して変化しており、一様でないことを観察している。このことは、ネットワークの進化モデルとして提唱されている単純な優先選択モデルではドメインの組み合わせネット

ワークの進化は説明できないことを示している。

## 論文審査の結果の要旨

シーケンシング技術の高速化を背景に、様々な種のゲノム配列が決定されており、これらのゲノムデータを比較し、種間で共通するあるいは種特異的な特徴と生命機能やその進化との関連性を調べる比較ゲノム解析が盛んに行われている。特に、真核生物は多様な系統特異的な高次機能を持つため、系統関係を考慮した比較ゲノム解析は真核生物の高次機能を理解する際に非常に有効であると考えられる。高度な多細胞体制を持つ真核生物は、より複雑なドメイン構造を持つタンパク質を持つことが知られており、かねてより機能の高度化とドメイン構造の複雑化の関連性についての解析がなされてきた。バイオインフォマティクス分野におけるドメイン構造の解析ではネットワーク解析の手法が用いられており、真核生物など superkingdom に特異的な組み合わせが存在することや、その進化モデルについての知見が得られている。しかしながら、真核生物のドメイン構造の網羅的な解析は、原核生物との比較や、少数の真核生物間の比較によるものであり、真核生物内の系統グループ間の比較解析は行われてこなかった。

申請者は、近年のゲノムプロジェクトの進展に伴い利用可能となった47種の真核生物のタンパク質セットを網羅的に用い、種の系統関係を考慮することで、タンパク質のドメイン構造の特異性と機能進化との関連性についての解析を行っている。この結果、動物および後口動物に特異的な組み合わせは、それぞれの共通祖先で出現した機能に関わるものが多く、高次機能の進化にドメイン構造の複雑化が大きく寄与していることを指摘している。また、原核生物と真核生物の比較では否定されていた、ドメインの出現時期と組み合わせのパートナー数との相関を観察しており、先行研究の知見をさらに深めることに成功した。これらの結果は、真核生物を一様に扱わず、なおかつ、個別の種の比較ではなく系統グループごとに比較することによって初めて得られたものであり、真核生物の系統グループ特異性に着目して解析した意義は大きい。

また、手法の面でも Mirkin らが開発したオーソログ遺伝子の出現時期の推定アルゴリズムを多分木に拡張して用いることで、系統関係の曖昧な種やドラフトゲノムデータしか存在しない種についても解析対象としている点も評価に値する。さらに、申請者は同手法によって得られる結果を祖先生物種のドメイン構成の推定に応用しており、既存のネットワーク解析における進化モデルについての検討を行っている。この結果は、先に述べた個別のドメインや組み合わせの特徴と、ネットワークの大域的な解析を結びつけるものとして非常に興味深い。また、本論文において、申請者が拡張・適用した特異性の同定手法は、より一般の比較ゲノム解析へ応用することの出来るものであり、今後増加する真核生物のデータに対して、同手法を用いることでさらなる知見を得ることが期待される。

以上、申請者が行った、真核生物のドメインの組み合わせの特異性と進化に関する比較ゲノム解析は基礎的な研究成果として高く評価できる。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。なお、論文内容とそれに関連した試問の結果、合格と認めた。