

# 主 論 文 要 旨

学籍番号		氏 名	八 谷 剛 史
主 論 文 題 目： Probabilistic Models for Inferring Orthology in Comparative Genome Analysis (比較ゲノム解析におけるオルソロジー推定のための確率モデル)			
(内容の要旨) シーケンシング技術の革新的な進歩により、生物配列のデータ量は指数関数的に増加している。増加し続ける生物配列の膨大なデータの中から生物学的に有意義な知見を得るためには、配列間の進化的な関係を推定する必要がある。 生物配列間の進化的な関係は、オルソロジーとパラロジーという 2 つの関係に分類することができる。オルソロジーの関係にある生物配列は祖先配列と同一の機能を保持しており、パラロジーの関係にある生物配列は系統特有な新たな機能を獲得している、と広く受け入れられている。そのため、配列間のオルソロジー関係を同定することにより、配列の機能を類推することができる。近年の急速なゲノム配列の蓄積を背景に、オルソロジーの推定に関して、2 つの新たな課題が提起されている。第 1 の課題は、染色体間のオルソロジー関係を同定することであり、第 2 の課題は、近隣種ゲノム比較において、遺伝子間のオルソロジー関係を同定することである。 本論文第一章では、オルソロジー推定の重要性について概説した。 第二章では、第 1 の課題に取り組んだ成果として、OSfinder アルゴリズムについて述べた。既存アルゴリズムにおいては、染色体間のオルソロジー関係を同定するために重要な役割を果たすパラメータの値を、ユーザが経験的に決定しなければならない。OSfinder は確率モデルに基づいたスコア体系を定義することにより、パラメータ値を自動的に最適化することを可能にした。この最適化により、OSfinder は既存アルゴリズムよりも高精度に染色体間のオルソロジー関係を同定できることが、哺乳類ゲノム比較及びバクテリアゲノム比較において示された。さらに、OSfinder のパラメータ値を自動的に決定できる機能は、比較ゲノム解析のスループットを大幅に向上させると期待できる。 第三章では、第 2 の課題に取り組んだ成果として、OASYS アルゴリズムについて述べた。近隣種ゲノム比較において遺伝子間のオルソロジー関係を同定するためには、タンパク質配列の相同性情報だけでは不十分であることが知られている。そこで OASYS は、タンパク質配列相同性と遺伝子の並びの保存情報を統合することにより、近隣種ゲノム比較においても遺伝子間のオルソロジー関係を高精度に同定することを可能にした。OASYS は異なるデータに基づく情報を統合するため、確率モデルのアプローチを活用している。原核生物ゲノムを用いた精度評価実験により、遺伝子の並びの保存情報は遺伝子間のオルソロジー関係を同定するために有用であることが示された。また、遺伝子の並びが種間で完全に保存されている“conserved gene cluster”の同定においても OASYS アルゴリズムは有用であることが示された。さらに、101 種の原核生物ゲノム比較及び 15 種の菌類界ゲノム比較に OASYS を適用することにより、ゲノム進化の過程において、遺伝子の並びの進化とタンパク質配列の進化の間に相関があることを明らかにした。 第四章では、OSfinder と OASYS のシステム生物学及び進化生物学への貢献について議論した。			