

# 主 論 文 要 旨

報告番号	① 乙 第	号	氏 名	長 嶺 誠 香
主 論 文 題 目： Genome-wide statistical prediction of interactions between biomolecules (ゲノムワイドな生体分子間相互作用の統計的予測)				
(内容の要旨) 近年、800 種を超える生物種のゲノム解読が既に完了していることが象徴するように、生命を解明するためのデータが多量に生み出されている。ここで、ゲノムという生命の設計図が得られていることに加え、生体機能をより精密に、分子レベルで解明しようとする現在までの分子生物学の研究により、生命を構成する部品の情報が多く得られている。この収集され、また急速に蓄積されていく情報を効率的に活用するための手段のひとつとして、統計的手法の利用が考えられる。 本研究は、この蓄積された生物学的データに対して統計的手法を適用することにより、生体分子間相互作用をコンピュータにより予測・解析する問題を扱う。一般に、生体機能は、タンパク質、代謝物等の生体分子がネットワークを構成することでシステムとして実現されている。その最小の構成単位の一つである生体分子間の相互作用を予測し、明らかにしていくことは、生体機能の解明に貢献するものである。特に疾病原因タンパク質と小分子化合物の間の相互作用を予測することは、創薬におけるリード化合物探索や、薬剤の未知薬効や副作用の推定に貢献するものであり、産業的・経済的な意義と価値を持つ。 本論文第一部では、生命現象解明における網羅的な生体分子間相互作用予測の意義について述べるとともに、本研究で扱う二つの生体分子間相互作用について概説した。 第二部では、生体内で最も重要な生体分子間ネットワークであり情報が蓄積されているタンパク質間相互作用ネットワークを利用した、協調的に作用する転写因子の推定について述べた。酵母 ( <i>S. cerevisiae</i> ) のクロマチン免疫沈降法 (ChIP) データ及びタンパク質間相互作用データを用いたコンピュータ実験から、タンパク質間相互作用ネットワーク上で近い位置に存在するタンパク質群の転写を制御する転写因子は、協調的に作用していることを見出した。これによりタンパク質間の物理的な相互作用ネットワークと、転写制御ネットワークの間に関連性が存在することを明らかにした。 第三部では、生物学と化学、バイオインフォマティクスとケモインフォマティクスの融合を目指しての生体内の標的タンパク質と化合物間の相互作用予測について述べた。入手が比較的容易な化合物の構造情報や質量分析データとタンパク質のアミノ酸配列情報に基づいて、統計的学習手法 Support Vector Machine (SVM) を用いたコンピュータ実験で、既知の認可薬剤とその標的タンパク質間の網羅的な結合予測が高精度で実現可能なことを示した。また、ヒトアンドロゲン受容体を対象とした 1900 万以上の化合物からの網羅的結合リガンド予測を行い、その予測結果に対する共同研究でのウェット実験による予測検証を通じて、コンピュータ予測と実験的検証の有効な統合手法を提案し、インシリコとウェット実験の統合を通じて効率的な相互作用予測が実現可能なことを示した。 第四部では、本研究をまとめるとともに、本研究の今後の方向性について論じた。 結論として、網羅的結合リガンド予測は、創薬におけるリード化合物探索の効率化及びコスト削減に貢献するものであり、本研究は統計的相互作用予測の有用性を提示した。				