主　論　文　の　要　約

論文題目	バイオインフォマティクスを用いた *Aspergillus oryzae* における転写因子制御ネットワークの網羅的解析に関する研究
氏　　名	岡　大椰

生体内における代謝、発生、免疫など様々な反応は、各反応に関与する遺伝子の発現を介して厳密に制御されている。この遺伝子の発現制御には、転写を制御しているタンパク質群、転写因子が大きな役割を担っている。とりわけ、ゲノム DNA 上に存在する特定のシスエレメント領域を認識･結合することで遺伝子の発現制御を行う DNA 結合型転写因子は、転写制御機構の中枢を担っている。このため、ゲノム DNA に埋め込まれた制御情報を読み取るメカニズムを解き明かすことは、DNA 結合型転写因子を介した転写制御ネットワークの全貌解明に繋がりうる。しかしながら、この転写制御ネットワークは非常に複雑に絡まり合っており、多くの生物種において大部分が未解明であるのが現状である。本論文においては、産業微生物 *Aspergillus oryzae* を標的生物とし、バイオインフォマティクス技術を駆使することで、*A. oryzae* における転写制御機構をより多面的かつ包括的に理解することを目的とした。

第一章では、*A. oryzae* における合理的な有用物質生産プロセスデザイン技術の確立を志向し、当研究グループで開発された転写制御ネットワークの網羅的解析システムを用いてコウジ酸代謝に関連する転写因子 KojR の機能解析を試みた。コウジ酸は、*A. oryzae* に特有の二次代謝産物であり、*A. oryzae* 内でメラニン形成を促すチロシナーゼの抑制に関与している。この特性から、化粧品分野で美白剤としての応用が期待されているだけでなく、医薬品、農薬、殺虫剤、抗原中剤としても利用されている。

転写因子 KojR に対して *in vitro* の DNA-転写因子結合反応により、ゲノム中の転写因子結合部位を網羅的に同定する手法、Genomic SELEX-Seq (gSELEX-Seq)と KojR 欠損株を用いた発現変動解析 RNA-Seq を実施した。KojR のゲノム中の結合部位および発現変動遺伝子の同定を行ったところ、その直接的な標的遺伝子が検出された。しかしながら、この遺伝子リストには、KojR によって発現制御されることが既知の *kojA*、*kojT* が含まれていなかったことから、解析条件のさらなる改善が必要であると考えられた。

第二章では、上記の転写制御ネットワークの網羅的解析システムを *A. oryzae* 由来のキシランやセルロースの代謝などに関与する転写因子 AoXlnR に応用し、その

標的遺伝子を包括的かつ高精度な同定を試みた。

結合領域解析 gSELEX-Seq によって得られた AoXlnR のゲノム中の結合領域に関する情報と DNA マイクロアレイによって得られた AoXlnR 変異体依存的な発現変動値を用いた統合解析により、AoXlnR の直接的な標的遺伝子の網羅的同定を試みた。gSELEX-Seq によって得られた候補制御遺伝子リストとマイクロアレイによって同定された発現変動遺伝子（DEGs）72 種類と照合したところ、その積集合に 51 種類の遺伝子が検出された。この 51 遺伝子には、AoXlnR によって発現制御を受けることが既知の遺伝子 8 種すべてが含まれていた。さらには、gSELEX-Seq によって得られた AoXlnR 結合領域に対して結合モチーフ解析を実施したところ、既知の AoXlnR 結合モチーフが検出された。これらの結果から、本解析法により、転写因子の標的遺伝子を網羅的かつ高精度に同定できることが強く示唆された。

第三章では、AoXlnR の結合に関連する様々なパラメーターと発現量の差の相関を解析することで、さらなる AoXlnR による転写制御機構をより詳細かつ多面的に解析・解明することを目的とした。

まず、各標的遺伝子の AoXlnR の発現依存的な発現変動レベルの比（AoXlnR の高発現株と AoXlnR 破壊株の遺伝子発現の変動差）と AoXlnR の結合に影響すると考えられるいくつかのパラメーターとの相関を解析した。その結果、プロモーター上に存在する AoXlnR 結合モチーフの個数と発現変動の間には優位な相関があることが明らかとなった。また、これら解析過程で、開始コドン上流に AoXlnR 結合モチーフを保有するにも関わらず、標的転写因子の高発現に応答しない遺伝子が多数検出されることが明らかとなった。このことから、転写因子-DNA 間相互作用には転写因子の結合モチーフ情報に加え、未知の要因が関与していることが示唆された。

そこで、AoXlnR の転写制御情報を基に、結合モチーフ付近の配列や、それら配列に基づく DNA 構造情報を抽出することで、配列環境情報から機械学習による転写因子-DNA 結合メカニズム予測モデルの構築を試みた。配列情報は結合モチーフの前後 25 bp からワンホットエンコーディング法によって、DNA 構造情報は結合モチーフの前後 25 bp から DNA 構造に関連するパラメーターHelT、ProT、MGW、Roll を DNAshapeR によって算出した。これらのパラメーター値と発現変動を教師値として機械学習モデルの構築を行い、結合部位周辺の配列と DNA 構造情報を用いて各結合モチーフの機能予測を実施した。その結果、AoXlnR 結合モチーフの下流に位置する遺伝子の発現変動の有無は、結合モチーフの周辺配列とその DNA 構造から高精度に予測できることが示された。このことは、転写因子による転写制御が、結合モチーフだけではなく、結合モチーフ周辺の特定の配列や DNA 構造情報にも大きく影響を受けていることが強く示唆された。

さらに、プロモーター内に存在する結合 DNA モチーフの個数に応じて遺伝子を分類し、それぞれ構築して用いた予測モデルを比較したところ、分類しない場合に比べ、発現変動をより高精度に予測できた。また、非常に興味深いことに、これら

の予測モデルにおいて、発現量の差に寄与するパラメーターがそれぞれ異なっていた。これは、プロモーター中の結合 DNA モチーフの個数によって AoXlnR を介した転写制御機構が異なることを示唆していた。

最後に、今回構築した AoXlnR による発現変動の予測モデルを用い、野生株の持つプロモーターと同等の発現が期待される DNA 構造を有する新たなプロモーター設計し、それらを相同組み換えによって麹菌ゲノムに導入することで、細胞内における変異型プロモーターの発現量を解析した。その結果、予測モデルによって発現を期待されたプロモーターでは、麹菌内で野生株と同等以上のキシラン誘導への発現応答が観察された。一方、予測モデルにおいて下流に位置する遺伝子がキシラン誘導に応答しないと判定されたプロモーターでは、麹菌内で野生型プロモーターより少ない発現量を示した。この結果は、本予測モデルの信頼性と妥当性を示すと同時に、転写因子による転写制御において、結合モチーフだけではなくその周辺の DNA 構造を考慮する必要性を改めて強調する結果となった。

本研究では、結合領域解析である gSELEX-Seq と各発現変動解析を組み合わせた新規転写因子解析システムを用いて、*A. oryzae* 由来の転写因子の結合メカニズムを解析した。*A. oryzae* は食品や醸造分野だけではなく、その高い安全性とタンパク質生産能力により、様々な分野への応用が期待されている微生物である。現在既に、*A, oryzae* 内で CRISPR/Cas9 を利用したゲノム改変が既に行われており、同技術と今回得られる AoXlnR 結合機構情報を応用することで大腸菌などでは現在発現が困難とされているタンパク質に対し *A. oryzae* を用いた異種タンパク質発現を検討している。本転写因子解析システムと CRISPR-Cas9 を用いた効率的な物質生産システムを組み合わせることが可能となれば、*A. oryzae* の産業応用に向けて大きな足がかりになると考えられる。また、本研究は、高速 DNA シーケンシングから機械学習に及ぶ、高度かつ最先端のバイオインフォマティクス解析と微生物を用いた物質生産に関する実践的な解析を同時に実施している数少ない研究であり、2 つの離れた領域の架け橋として、重要な意義を持つと考えられる。