

論文審査の結果の要旨

氏名 モハammed バドルル アーサン

DNA から遺伝子が転写される場所を転写開始点と呼ぶ。10 年ほど前までは、1 つの遺伝子の転写開始点の数は少なく、通常は1つと考えられていた。しかし、この認識は誤っており、観測の感度が十分でなかった時代の限界を象徴している。過去10年間 DNA シーケンサー(解読措置)が DNA を解読できる量が5ケタ近く増大した。その結果、転写開始点は1つではなく多数存在する場合があります、広範囲に分布していることも分かってきた。そのため、転写開始点の分布を分類し、どのような機能を持っている遺伝子が、どのような転写開始点の分布をもつか? という研究が盛んになってきている。しかし、転写開始点の分布は多様であり、正規分布等のように数式での記述が容易な分布は少ない。分布を意味あるクラスへと分類するための指標があると便利である、と考えられてきた。そこで本論文は、転写開始点の分布を分類するための指標を提案している。

本論文は8章に分かれている。まず1章で問題を導入したのちに、2章で過去の研究をサーベイし、本論文の結果の位置づけを述べている。3章では転写開始点を収集するための生物学的実験手法について述べ、著者がデータ解析したショウジョウバエの転写開始点の分布を例に、どのようにして情報学的な解析を進めるべきかについて述べている。4章では転写開始点の分布を分類するための手法として、Aggregated Index, Gini Coefficient, Average Neighborhood Distance という指標を導入している。そしてこれらの指標が、従来は分類が難しかった分布を明快に分類するに役立つことを議論している。5章ではショウジョウバエの転写開始点の分布を例に、既に知られているプロモータの転写因子結合部位の配列モチーフが、どのような転写開始点の分布と相関するかについて調べている。6章では代替的プロモータが存在する遺伝子に対する転写開始点の分布が、あたらしく定義した指標により特徴づけやすいことを紹介している。7章は著者が作成したメダカの転写開始点データベースについて触れている。8章では提案した転写開始点分布の指標の有用性と限界について議論している。指標は新しく、従来は分類しにくかった分布を分類しやすくしたという点で評価できる。

なお、本論文は、森下真一をはじめとする多数の研究者との共同研究であるが、論文提出者が主体となって開発、分析及び検証を行ったもので、論文提出者の寄与が十分であると判断する。

したがって、博士(科学)の学位を授与できると認める。

以上1006字