

論文内容の要旨

申請者氏名 山下 理宇

転写による遺伝子発現制御機構は、基本的には様々な転写因子とその DNA 結合部位であるシスエレメントとの相互作用によって行われている。この解析には、シスエレメントを含むプロモータ配列が必要であり、それには正確な転写開始点 (transcription start site : TSS) の同定が不可欠である。申請者は、TSS のデータベース Data Base of Transcription Start Sites (DBTSS) を構築した。DBTSS には、mRNA の 5' 端配列を選択的に抽出できる oligo-capping 法(ヒト)または Cap-trapper 法(マウス)によって決定された 5' 端配列を、ゲノムにマッピングした情報が記載されている。これには、11,234 ヒト遺伝子と 7,524 マウス遺伝子の転写開始点が、それぞれ 190,964、195,446 クローンの配列として登録されている。

このデータベースを使って TSS を網羅的に見た結果、TSS には揺らぎがあり、そのパターンは二つに大別されることがわかった。一つは、揺らぎの標準偏差が 100 塩基以下であり、こちらは転写開始の揺らぎを反映していると考えられた。もう一つは、揺らぎの標準偏差が 100 塩基を越えるグループであり、こちらは alternative promoter の可能性が考えられた。また、従来 TSS 付近にあるとされている、4 種のシスエレメントを探索した結果、TATA-box や initiator は TSS から一定の距離に顕著に検出できたが、downstream promoter element (DPE) や TFIIB recognition element (BRE) は、明確には検出できなかった。さらに、TSS 付近の配列に着目すると、1) initiator を持つもの、2) terminal oligo-pyrimidine tract (TOP) を持つもの、3) それ以外の 3 グループに分類できることがわかった。

続いて、翻訳レベルで調節を受けるとされている TOP 配列を持つ遺伝子群 (TOP 遺伝子) を、ヒトゲノムから網羅的に予測した。検出された TOP 遺伝子群は、リボソームタンパク質、翻訳伸長因子等、従来知られていたものに加え、いくつかの翻訳開始因子、構造タンパク質、サイクリン等、タンパク質合成および細胞分裂に関わる主要な遺伝子群を含んでいた。さらに、TOP 遺伝子は、従来考えられていた翻訳関係の遺伝子にあるばかりではなく、一部の膜タンパク質、構造タンパク質等、より広範囲に存在する可能性が示唆された。

プロモータ領域の CpG island と遺伝子発現の組織特異性については、矛盾した報告があるので、これらの関係について検討した。まずプロモータ領域に CpG island を持つ遺伝子群 (ヒト:6600、マウス 2948) と、持たない遺伝子群 (ヒト:2948、マウス 1830) に大別した。UniGene には、ある遺伝子に属する cDNA 配列がどの組織・ライブラリー由来であるかというソースの情報があり、この数はおよその組織特異性の指標となり得る。これを用いて解析した結果、CpG を持たない群では UniGene のソース数が少ないことがわかり、CpG を持つ群に比較してより組織特異的に発現していることが示唆された。

論文審査結果の要旨

申請者氏名 山下 理宇

mRNA の 5' 端配列を選択的にクローニングできる oligo-capping 法(ヒト)あるいは cap-trapper 法(マウス)によって決定された、大量の 5' 末端配列をゲノム配列上にマップした転写開始点データベース (Data Base of Transcription Start Sites; DBTSS)の構築に申請者は中心的役割を果たした。そのデータベースには、11,234 ヒト遺伝子と 7,524 マウス遺伝子の転写開始点が、それぞれ 190,964、195,446 クローンの配列として登録されており、その規模は世界的にも類をみないものである。

申請者は、構築したデータベースを解析した結果、転写開始点が厳密に決まっている遺伝子に加えて、転写開始点には揺らぎがある遺伝子も大きな割合を占め、そのパターンは二つに大別されることを明らかにした。一つは、揺らぎの標準偏差が 100 塩基以下であり、こちらは一つのプロモーター配列の中での転写開始部位の揺らぎを反映しているものであり、他方は、揺らぎの標準偏差が 100 塩基を越えるグループであり、alternative promoter であると考えられた。また、従来転写開始点付近にあるとされている、4 種のシスエレメントの内、TATA-box や initiator は転写開始点から一定の距離に検出できるが、downstream promoter element (DPE)や TFIIB recognition element (BRE)は、明確には検出できないことを示した。TSS 付近の配列に着目すると、1) initiator を持つもの、2) terminal oligo-pyrimidine tract (TOP)を持つもの、3)それ以外の 3 グループに分類できることも明らかにした。こうした、ヒト・マウスのプロモーター領域の構造的な特徴の全体像は、申請者が構築した大規模なデータベースを用いて、始めて明らかにできたものである。

さらに申請者は、翻訳レベルで調節を受けるとされている TOP 配列を持つ遺伝子群 (TOP 遺伝子) を、ヒトゲノムから網羅的に予測した結果、TOP 遺伝子は、従来考えられていた翻訳関係の遺伝子にあるばかりではなく、一部の膜タンパク質、構造タンパク質等、より広範囲に存在する可能性を示した。また、プロモーター領域の CpG island と遺伝子発現の組織特異性についても、大規模な解析を行い、CpG を持たない群は、CpG を持つ群に比較してより組織特異的に発現していることを示した。

以上のように、本論文では、ヒト・マウスについての大規模な転写開始点データベースを構築し、それを用いて、転写開始点領域の構造的な特徴の解析を行っており、学術上、応用上貢献するところが少なくない。よって審査委員一同は、本論文が博士 (バイオサイエンス) の学位論文として価値あるものと認めた。