

審査の結果の要旨

氏名 市川 和樹

クロマチン情報とは、DNA 修飾、ヒストン修飾、ヌクレオソーム位置、DNA の核内折畳み構造等を示す概念である。21世紀に入り、ヒトや脊椎動物を対象として全ゲノム概要配列の解読が進んでいる。それと並行して全ゲノムからクロマチン情報を観測する方法も開発され普及しつつある。その情報量は膨大であり、ゲノム配列の情報量を凌ぐデータが収集されている。ゲノム全域のクロマチン情報を分析するには、従来とは異なる高速なアルゴリズムが必要となることも知られている。

クロマチン情報を分析する問題は多岐にわたるが、論文提出者は3つの問題に取り組んだ。1つ目はヌクレオソーム位置を分類してクラスタリングする問題である。ヌクレオソーム位置は遺伝子転写を制御していると考えられており、いくつかのパターンが存在し、組織に依存して変化することも過去の研究で示唆されている。そのパターンを分類（クラスタリング）することは遺伝子発現解析では重要である。ところが数万個の遺伝子転写開始点前後の1000塩基の領域で観測したヌクレオソーム位置を従来の方法でクラスタリングするには、計算時間がかかるという問題があった。論文提出者は同一のクラスタリング結果を出力することを保証しながら、従来法と比べて5~26倍高速化するアルゴリズムを研究開発することに成功した (Ichikawa, Morishita. 2014)。

2つ目は、特定のDNA修飾とヒストン修飾が組み合わさることで遺伝子発現が制御される領域を枚挙する問題である。全ゲノムで観測する技術がこの10年近く普及し、様々な組織において適用されつつある。ただし、全ゲノムから数万に及ぶ領域を精度良くかつ高速に枚挙することは自明でない。たとえば、多能性をもつ初期胚から細胞運命決定する過程では、DNA中のCpGが低メチル化する一方で、ヒストンH3の転写抑制マーク(H3K27me3)が上昇する領域(K27HMD)が多数存在する。特に発生関連遺伝子が顕著にみられることが2013~4年に発見された。このK27HMD領域を推定する経験的手法はあるが、精度を保証しつつ高速に検出するためのアルゴリズム的考察は無かった。論文提出者は、既存のアルゴリズムのパラメータを調節することにより、経験的手法を補強できることを示した (Ichikawa, Morishita. 2015)。

3つ目は、DNAが核内で折り畳まれている情報をゲノムアセンブリの高精度化、たとえばセントロメア領域の解読に応用する問題である。DNAは核内では折り畳まれており、物理的に近接する領域を架橋してDNA解読装置で読み取る技術(Hi-C法)が普及しつつある。この方法を論文提出者は工夫して、繰り返し配列が稠密に存在するセントロメア領域を解読し、メダカゲノムを対象にセントロメア領域の約10%を染色体地図上に対応させることに成功した (Ichikawa *et al.* 2017)。約10%の領域ではあるものの、脊椎動物の中では最大級のセントロメア領域の解析である。今後、セントロメア領域を解析する研究に寄与する可能性が高い。

以上のように、論文提出者はクロマチン情報解析の3つの問題について、新しい情報

学的な考察をし、新手法を提案した。なお、本論文は、数名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解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1337字