

論文審査の結果の要旨

氏名 笠原 雅弘

本論文は6章からなり、以下の内容について述べられている。第1章で、遺伝学において最も基礎的な情報であるゲノム配列について総論的な導入を行った後、第2章で、本論文で扱うDNA配列決定の手法について網羅的な解説を行っている。第3章で、DNA配列決定の手法のひとつである全ゲノムショットガン法について解説を行っている。第4章で、論文提出者が開発したRAMENゲノムアセンブラの詳細について解説を行っている。第5章と第6章で、並列計算機上での開発手法に触れながら、上記RAMENゲノムアセンブラの有効性についての検討を行っている。より詳細な概要は以下の通りである。

広範な応用にもかかわらず、ゲノム配列の決定は簡単ではなく、大量の実験と巧みなアルゴリズムを必要とする。最新のDNAシーケンサーは一回の稼働で1キャピラリーあたり、連続する1000塩基対を決定することができるが、例えば脊椎動物のような複雑なゲノムはそのサイズが数十億塩基対に達することもあり、DNAシーケンサーの読み取り長を遙かに超えている。そのため、ゲノムDNAをランダムに裁断しDNAシーケンサーで配列決定する「全ゲノムショットガン法」がゲノム配列を決定するために使われている。

しかし、ランダムなゲノムDNA断片配列はゲノム上での位置が分からない。そこで、ランダムDNA断片配列の集合から元のゲノム配列を再構成するコンピューターアルゴリズムが必要となる。この操作を行うコンピュータープログラムを「ゲノムアセンブラ」と呼び、世界中で多くの研究グループがこの問題に取り組み数百篇にも及ぶ数の論文を公表している。しかし、大きなゲノム配列に対する配列解読プロジェクトにおいて見られるアセンブリ上の問題は、今まで詳細かつ実践的には十分に語られてこなかった。これはおそらく、扱うデータが大きすぎるために、問題を発見し記述すること自体が難しかったからである。

膨大なデータを扱う際には、計算資源の制約から、単純なアルゴリズムやデータ構造しか使うことはできない。また、ゲノムアセンブリの性質は、解読対象の種やシーケンシングセンターに応じて大きく異なることもある。ある状況では有効であった改良が、種やシーケンシングセンターが異なると改悪である場合すらある。このような理由で、大きなゲノムに対するアセンブリアルゴリズムは体系的に論じられたことは無かった。本論文ではまず、大きなゲノム配列を解読するための既存のアセンブリアルゴリズムを体系的に紹介している。ゲノムアセンブリにおける問題点の多くは、DNAシーケンサーの技術的限界に由来しているため、DNAシーケンシング技術についても概説している。

次に、論文提出者が新たに開発したゲノムアセンブラであるRAMENを、ステップ毎の詳細なアルゴリズムとともに報告している。RAMENはメダカゲノム解読を目指して開発された。アルゴリズムの詳細に加えて、大きなゲノム配列解読において遭遇す

る問題点を実際のゲノムプロジェクトを例に報告している。具体的には、ベクターマスク問題や多型、系統的シーケンシングエラー、互い違いのスキュッフォールド、クローンサイズ推定の精度不足などについて論じている。これらの問題点に対処するアルゴリズムと、その背景となる考え方についても同時に述べている。

また、RAMENの性能を測るために、メダカ (*Oryzias latipes*) ゲノムをアセンブルしている。ショットガンリードの配列カバー率は10倍以上であり、総計クローンカバー率53.7倍分のプラスミド・フォスミド・BAC等のメイトペア情報が与えられた。アセンブリの結果、N50スキュッフォールド長は1.41Mbに達した。また、メダカの遺伝学的地図を統合し、ウルトラコンティグ長は5.1Mbに達し、約90%の塩基が染色体と関連付けられた。アセンブリと完成BAC配列との比較から、コンティグの端100bpを除くと塩基精度は99.96%であると推定されている。

これらの結果は、RAMENは脊椎動物のゲノムをアセンブルするのに十分な精度とスケーラビリティを備えていることを示唆するものである。さらに、本論文の最後に、大量のDNAシーケンサーから産出される膨大なデータを処理するためにはアルゴリズムを実装するうえで特別な注意が必要であったという教訓を紹介している。本稿では並列プログラミングのフレームワークについて注目して議論しているが、これは次世代シーケンサー用のゲノムアセンブラを実装する際には特に重要となることを指摘している。

なお、本論文の一部は、武田洋幸博士、森下真一博士、小原雄治博士との共同研究であるが、RAMENゲノムアセンブラの開発と性能評価は、論文提出者が主体となって行ったもので、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。