

論文審査の要旨および学識確認結果

報告番号	㊶／乙第 号	氏 名	Vasanthan Jayakumar
論文審査担当者：			
主査	慶應義塾大学教授	博士（理学）	榊原 康文
副査	慶應義塾大学教授	工学博士	岡 浩太郎
	慶應義塾大学教授	博士（地球環境科学）	土居 信英
	慶應義塾大学准教授	博士（工学）	舟橋 啓
	慶應義塾大学教授	博士（工学）	櫻井 彰人
(論文審査の要旨)			
<p>Dr. G. R. Damodharan college of science (India) 修士課程(Bioinformatics) 修了の Vasanthan Jayakumar 君提出の学位請求論文は「Computational pipelines for assembly, analysis, and evaluation of genome sequences (ゲノム配列のアセンブリ, 解析および評価のための計算パイプラインの開発)」と題し, 4章から構成されている。</p> <p>ゲノムとは, 生物の全遺伝情報が書かれた生物の設計図である。DNA シークエンサー (配列読取り装置) は, さまざまな生物のゲノム配列の解読を可能にした。DNA シークエンサーで読まれた短い配列はリードと呼ばれ, 大量に読まれたリードから元のゲノム配列を組み立てる情報処理をアセンブリと言う。サンガー法に代表される第一世代シークエンサーに続く第二世代シークエンサーは多くの生物のゲノム配列を解読するのに役立ったが, その欠点は短いリード長であり, 断片化されたゲノムアセンブリにつながった。最近開発された第三世代シークエンサーが提供する最大の利点は, リード長が指数関数的に増加したことである。その長いリードはロングリードと呼ばれ, リード長は反復 (リピート) 配列よりも長くなる。</p> <p>本研究は, 第三世代シークエンサーのロングリードを用いたゲノムアセンブリ, 遺伝子コード領域やリピート配列などのアノテーション作業, そしてアセンブリされたゲノムの品質評価のための計算パイプラインを開発した。最初の研究では, 開発した計算パイプラインを応用して, 日本の有名な開花植物であるアサガオ (学術名 <i>Ipomoea nil</i>) のゲノムを組み立て, 多様な突然変異パターンを作り出す因子を同定した。そのアセンブリの過程で高品質なアセンブリを得るためにさまざまなパラメータを調整する必要があった。そこで第二の研究では, パイプラインのパラメータをチューニングし, それを様々な生物のゲノムアセンブリに適用すること, およびロングリード用のアセンブリプログラムを評価する方法に焦点を当てた。</p> <p>本論文の第1章では, ゲノムシークエンスとアセンブリ手法の重要性について述べるとともに, 本研究で扱う二つの課題について概説した。</p> <p>第2章では, 疑似染色体レベルまで復元したアサガオゲノムのアセンブリについて報告した。完成したゲノム配列は, アセンブリ精度の尺度の一つであるコンティグ N50 の値が 1.8Mb にまで到達し, 最新のゲノム解読プロジェクトの中でもトップクラスの高精度である。ゲノム解読により, <i>I.nil</i> の主要な突然変異原として知られている <i>Tpn1</i> トランスポソンの網羅的同定とカタログ化を完成した。トマトやジャガイモなどの近縁種との比較ゲノム解析により, 全ゲノムが二倍に増える全ゲノム重複のイベントが起こった時期を正確に同定した。</p> <p>第3章では, バクテリアから植物までゲノムサイズの異なる代表的な4つの生物種のロングリードデータセットを用いて, ロングリード用アセンブリプログラムの評価をさまざまな基準に基づいて報告した。この評価結果は, ゲノムアセンブリを行うためのパラメータを効率的に調整するガイドとして役立つ。</p> <p>第4章では, 本研究を総括するとともに, 開発したゲノムアセンブリパイプラインについて今後の応用可能性を議論した。</p> <p>以上結論として, アサガオの高精度なゲノム配列を決定することにより植物学や遺伝学分野の発展に資するとともに, ロングリードを用いた効果的なアセンブリのパラメータ調整について示した。よって, 本論文の著者は博士(理学)の学位を受ける資格があるものと認める。</p>			
学識確認結果	<p>学位請求論文を中心にして関連学術について上記審査会委員で試問を行い、当該学術に関し広く深い学識を有することを確認した。</p> <p>また、語学（英語）についても十分な学力を有することを確認した。</p>		