

審査の結果の要旨

氏 名 林 周斗

本論文は、個々人やヒト腫瘍細胞の免疫反応の差異を理解する上できわめて重要なヒト免疫系の構成要素であるヒト白血球抗原 (HLA) 遺伝子について、次世代シーケンサーのデータからきわめて高精度にその遺伝子型決定を行う技法の開発・研究を行ったものである。特に、HLA遺伝子の多様性、多数の類似HLA偽遺伝子の存在、次世代シーケンサーの読み取り精度の制約などから、HLA遺伝子の型を全ゲノムシーケンズデータから解析することはこれまで困難であったが、本論文では、次世代シーケンサーのリードのゲノムへのアラインメントスコアリング手法とベイズモデルを構築することで、その精度の大幅な改善を実現している。また、本論文では、精度の検証に実際の全ゲノムシーケンズデータを用いることで、実応用に近い状況で信頼性の高い精度確認を行っている。

本論文は六章からなり、第一章では本論文の背景と動機として、HLA遺伝子の免疫反応における重要性ならびに、近年の次世代シーケンサーの発展によるがんゲノム研究の発展について論じ、本論文の貢献であるゲノムデータからのHLA遺伝子型予測の重要性について議論を行っている。

第二章では、本論文の議論の基礎となる基本的な知識、技術として、HLA遺伝子の性質、次世代シーケンサーおよびそのデータの性質、マルコフ連鎖モンテカルロ法のいくつかの基本的な手法について解説を行っている。

第三章では、次世代シーケンサーによる全ゲノムシーケンズデータからHLA遺伝子領域のリードを高精度に選択するための、参照ゲノムへのアラインメントおよびそのスコアリング手法を提案している。

第四章では、HLA遺伝子型を予測する既存手法について議論を行うとともに、アラインメントされた個人の全ゲノムシーケンズのリードから実際のHLA型を予測するベイズモデルおよびその学習方法を提案している。さらに、実際の全ゲノムシーケンズデータおよび全エクソームデータに対し提案した手法を適用するとともに既存手法との比較も行い、精度の検証を行っている。

第五章では、第四章の手法をさらに拡張し、体細胞変異した腫瘍細胞のHLA型を予測するための新たなベイズモデルに基づく手法を提案している。その上で実際の全ゲノムデータ、全エクソームデータ上で精度の検証を行っている。

第六章では、本論文の貢献について要約するとともに、今後のHLA遺伝子型研究の発

展の方向性について論じている。

このように本論文は、個人の免疫反応や腫瘍細胞の免疫反応を理解するだけでなくその診療上も重要なHLA遺伝子型の予測を全ゲノムシーケンスデータより行う手法について、従来手法と比べきわめて高精度に予測する新たな手法を構築し、検証を行ったものであり、今後のがんゲノム解析・個人ゲノム医療のさらなる発展にも大きくつながることが期待できる研究である。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。