

審査の結果の要旨

氏 名 森山 卓也

本論文は、個々のがんを理解し、治療する上で必須の情報である、後天的に生じたゲノム上の変異(体細胞変異)を、次世代シーケンサーのデータから検出する手法の開発・研究を行ったものである。体細胞変異はがんの原因となる事から、がんの治療・研究の双方において、検出精度の改善は喫緊の課題である。通常、体細胞変異の検出は、単一検体のがん組織由来のシーケンスデータに基づく、もしくは一人の患者の複数検体に基づいて行われるが、本論文は、これらの精度改善に関していくつかの事前知識に着目した新たな階層ベイズモデルの設計を考案することで、これらのいずれの状況においても精度の大幅な改善に成功している。

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、シーケンサーの技術発展と体細胞変異検出のがん研究・治療との関係性を論じ、体細胞変異検出の重要性について議論を行なっている。

第二章では、本論文の議論の基礎となる知識・技術について解説を行なっている。まず、体細胞変異検出に用いる次世代シーケンサーおよびそのデータの性質について解説している。次に、階層ベイズモデルを利用する上で必要な、変分ベイズやマルコフ連鎖モンテカルロ法などの計算手法を解説している。最後に、本論文で議論する事前知識の一つである、進化系統樹の定義と性質を解説している。

第三章では、単一検体に基づく変異検出手法の精度改善を行っている。この章の冒頭では、既存手法やそこで用いられる事前知識について解説する。次に、この章での階層ベイズモデルの設計指針となる確率モデルの統合方法を議論する。ここでは、分割に基づく確率モデルの統合方法を提案し、ベイズモデル平均化との違いを議論している。さらに、単一検体に基づく変異検出手法のための階層ベイズモデルの設計を提案している。最後に、シミュレーションデータや実際のシーケンスデータに基づいて、提案手法の性能評価を行い、提案手法による精度の改善を確認している。

第四章、複数検体に基づく変異検出手法の精度改善を行っている。まず、既存手法やこの章で用いる事前知識である変異共有の想定について解説し、この想定に基づく検出方法に関して二つの問題点を指摘している。さらに、これら問題に対する解決策を考案し、これを元に変異検出のための階層ベイズモデルを設計している。最後に、実データに基づいたシーケンスデータのシミュレーションにより提案手法の性能評価を行い、

提案手法による精度の改善を確認している。

第五章では、複数検体に基づく変異検出手法の精度改善に関して、その事前知識としての進化系統樹の性質について議論する。この章では、既存手法が用いている進化系統樹の性質を解説し、この性質が体細胞変異の検出にどれだけの影響を与えるかを、いくつかの仮定を置いた元で考察している。この考察の結果、進化系統樹の性質を用いると、高い検出特異度が保証する場合が存在することを明らかにしている。

第六章では、本論文の貢献について要約するとともに、今後の体細胞変異検出手法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、がん治療・研究において必須の情報と考えられている、体細胞変異の検出手法について、階層ベイズモデルの設計を通じ、考慮されていない事前知識を利用することで検出精度の大幅な改善に成功している。体細胞変異の検出は、がんゲノムの分野における最も重要な解析手法の一つあり、本論文の研究はがんゲノムの分野の今後の発展に大きくなることが期待できる研究である。さらに、考案した階層ベイズモデルの設計は体細胞変異の検出以外の解析手法に対しても有益であることが期待される。

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。