

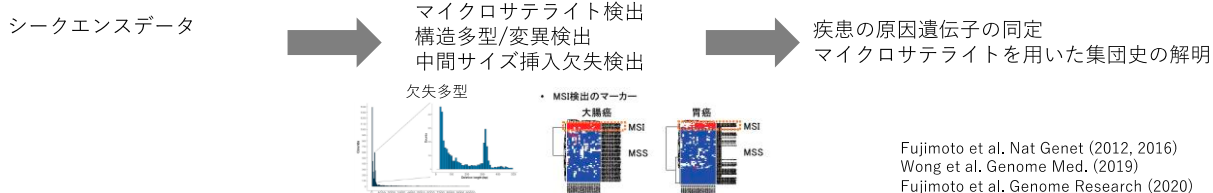
【目的】 人類集団には、様々な遺伝的多様性が存在し、疾患のリスクや表現型の個人差に関わっていることが知られている。また、がんにおいても様々な種類の体細胞変異が発癌の原因となることが知られている。近年のゲノム配列決定技術の発展により、多くの大規模ヒトゲノム、がんゲノムシーケンス研究が行われている。しかしながら、シーケンスデータは、主に読み取り長が 100 bp 程度の短鎖シーケンサー（いわゆる次世代シーケンサー:NGS）を用いて取得されており、情報解析が困難なタイプの変異や多型の検出は行われていない。近年入手可能になった長鎖シーケンス技術は、検出困難な多型や変異の検出に有効であると期待されているが、エラー率の高さなどの欠点があり導入が進んでいない。我々は、1. 情報解析手法を開発し従来の手法では検出が困難な変異/多型の検出、2. 長鎖シーケンサーの情報解析手法の開発による変異/多型の検出、3. 変異/多型の機能的意義の解明、を目的として研究を行った。

【方法】 エラー率が高く検出が難しいマイクロサテライトの変異/多型の検出法の開発、中間サイズ（30~5,000 bp 程度）の挿入・欠失の検出、逆位・転座などの構造異常の検出法の開発を行った。また、がんの体細胞変異や日本人の正常組織の多様性を解析した。長鎖シーケンサーを用いてがんサンプルの全ゲノムシーケンスを行った。また、変異/多型検出する手法を開発し長鎖シーケンスデータを解析した。さらに、ゲノム編集技術を用いることで中間サイズ欠失が遺伝子発現に与える影響を解析した。

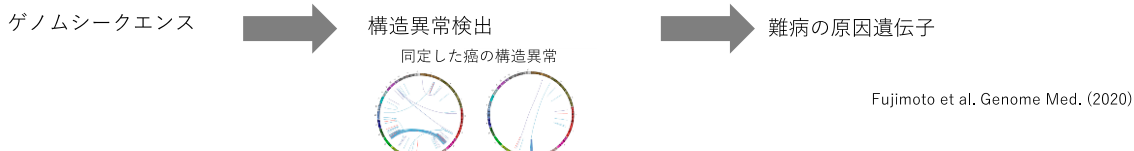
【結果】 3,000 サンプルのがんの全ゲノムシーケンスを解析し、マイクロサテライト領域の体細胞変異を検出した。その結果、変異率が高いがん（マイクロサテライト不安定性）サンプルを検出した。さらに、マイクロサテライトの変異率に影響を与える要因を見出した。中間サイズの挿入や欠失の同定を行い、遺伝子発現の個人差との関係を調査した。その結果、数%の中間サイズの挿入や欠失が遺伝子発現の個人差に関係し、機能的であることが示唆された。CRISPR-Cas9 法を用いて細胞株に欠失を導入し、機能的意義を解析したところ、データ解析の予測と合致する結果が得られた。長鎖シーケンサーを用いてがんサンプルの全ゲノムシーケンスを行った。解析の結果、挿入多型には、トランスポゾン由来のものが多く、欠失の多型は非相同組換えなどのさまざまな要因で生じていることが判明した。

本研究の結果の概要

(1) 解析手法の新規開発によるNGS データからの変異/多型の検出



(2) 長鎖シーケンサーを用いたシーケンスと情報解析



(3) 変異/多型や新規転写産物の機能的意義の解明

