

# 肺腺癌の発生・進展に関わる遺伝子の同定法の検討

修了年月 2012年3月

先端生命科学専攻 がん先端生命科学分野

指導教官 江角浩安教授・土原一哉准教授

学籍番号:106345 氏名:松島洗達

キーワード:大規模シーケンス、エキソーム解析、肺腺癌

## 序論

肺腺癌は全肺がんで最も多い組織型で、アジア系、女性、非喫煙者の頻度が高く、また EGFR 変異陽性肺腺癌には EGFR 阻害剤が奏効するなどの特徴がある。成因に不明な点が多い肺腺癌について 100 例規模で全エクソンにおける遺伝子変異を探索するプロジェクトに参加した。しかし、大規模シーケンスデータから変異箇所を同定する現行のプログラム (GATK v1) では、一塩基変異 (SNV) の判定にしばしば誤りが生じることが明らかになった。本研究では SNV 誤判定の原因を明らかにし、より正確に判定を行う方法の開発を行うこと、さらにこれを利用して肺腺癌 58 症例の全エクソンシーケンス結果から体細胞変異を同定し、変異遺伝子およびシグナルパスウェイの分布を検討して新たな肺腺癌関連遺伝子の探索のための情報を得ることを目的とした。

## 結果・考察

### SNV 誤判定の原因と特徴

全エクソンシーケンスデータで検出された SNV 候補周辺のリードのアライメントを観察したところ、SNV の誤判定の主要な原因は、サンプルゲノム上の indel を無視し

たアライメント(ミスアライメント)によるものとわかった。またミスアライメントの際にはリードと参照配列との相違が増え、SNV と判定した塩基以外が参照配列と一致する perfect match (PM) リードが減少することが予想された (Fig. 1)。

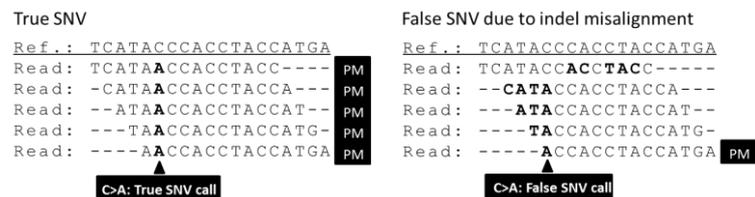


Fig.1 True SNV と False SNV

### PM リードを用いた誤判定除去フィルターの検討

正しいアライメントに基づいて判定された SNV (true SNV) とミスアライメントに起因する誤判定 (false SNV) を、判定に用いられた PM リードの本数によって層別化した結果、false SNV では PM リードが明らかに少なかった。この結果から PM リードの本数に基づく誤判定除去のフィルターを設定し、PM リード数が 2 本以上、未満で分類すると、最も効率よく false SNV が除去できることがわかった (Fig. 2)。

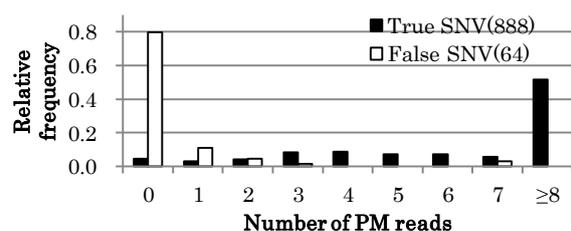


Fig. 2 PM リードによる SNV の層別化

## フィルターの精度の検証

このフィルターを利用することで、True SNV と False SNV をふるい分けられるか、アライメントの観察とサンガーシーケンスで検証したところ、偽陰性 (false SNV が残る) 率が低いことが確認できた。また、indel realignment を併用して SNV を判定する GATK の改良型 (GATK v3) による判定の結果と比較したところ、ほぼ同等の成績を示した。今回開発したフィルターは、計算機負荷が小さく、他のアルゴリズムへの拡張性も高い点が長所であると考えられた (Fig. 3)。

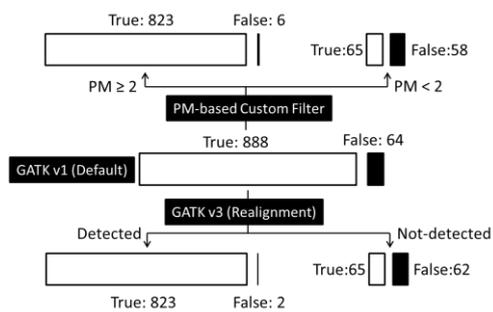


Fig.3 PM リード数に基づく SNV 誤判定  
フィルターと GATK v3 との性能比較

## 肺腺癌エクソンシーケンス 58 症例の体細胞変異の集計

肺腺癌 58 症例の癌部ゲノムの DNA のエクソンシーケンスデータをもとに体細胞変異の同定を行い、症例間での重複した変異を検討した。塩基レベル・コドンレベルでは *EGFR*・*KRAS* の変異が多数例で確認され、遺伝子レベルでは新たに *TP53* で変異の重複が認められた。ただし変異が見つかった 5,304 遺伝子のうち過半数は 1 症例にのみ変異が認められるものであり、個々の肺腺癌症例における遺伝子変異パターンの多様性がうかがわれる結果であった (Table 1)。

Table 1. 58 症例で体細胞変異が  
確認された遺伝子

Number of cases	Number of genes
30	1
26	1
16	1
10	1
9	1
8	7
7	13
6	29
5	50
4	119
3	383
2	1,063
1	3,635
Total	5,304

## EGFR 変異との排他性の検討

5 症例以上で変異が起きていた 103 遺伝子について、*EGFR* 変異と排他的に起きている変異なのか調べたところ 15 遺伝子では有意差をもって *EGFR* 変異陰性症例に変異が集積していた。これらの中には *EGFR* が関与するシグナルパスウェイに属するものもあり、*EGFR* 変異陰性例で変異型 *EGFR* の機能を代替している可能性も想定された。

## パスウェイ解析

KEGG PATHWAY データベースを用いて肺腺癌症例で変異が集積するパスウェイの探索を行ったところ、多くの臓器の癌で共通した遺伝子変異の異常が生じていることが示唆された。また *EGFR* が含まれるパスウェイの中でも ErbB signaling pathway などは *EGFR* 変異陰性症例群でも有意に変異症例が重複しており、ここに含まれる遺伝子変異は肺腺癌の生物学的特性への寄与が高い可能性が示唆された。

## 今後の展望

全エクソンシーケンスから得られた変異情報に臨床病理学的因子などを加味した解析を行い、肺腺癌やその亜型に特徴的な変異遺伝子・機能グループの探索を進める。また解析結果をもとにできるだけ多くの変異遺伝子をクローニングし、培養細胞実験系において今回同定された変異とがんの生物学的特性との関連を検証する。