



TGACCTTGAATGAGTGCAGCGAGGTCATGACGAAC

# 癌の難治性の解析に必要な情報 入手法とその数理統計解析方法

九州大学病院別府病院外科

内 龍太郎・三森 功士

Ryutaro Uchi

Koshi Mimori

(診療科長)

ACTGGAACCTTAGTCCAGTCTCCAGTACTGCTTG

## Summary

癌の難治性の解析には次世代シーケンサーを用いた Exome sequence や RNA sequence, マイクロアレイなどを用いたさまざまな解析が行われる。これらの実験から得られたデータは論文化される際に公共データベースに登録されて広く公開されている。実際に自分で実験を行うことができればそれに越したことはないが、コストや手間を考えれば、これらのデータベースを利用することは非常に有用である。解析も目的に沿ってさまざまな方法が存在するが、あまり大きなデータを扱ったことがなくとも比較的容易に解析が可能な手法も存在し活用されている。

Surgery Frontier 22(3) : 66-70, 2015

## Key Words

公共データベース, 解析手法, TCGA, 次世代シーケンサー, マイクロアレイ

### はじめに

医療の発展にもかかわらず、多くの癌は化学療法のみでは完治が難しく、治療は困難をきわめている。しかし、最近の次世代シーケンサーを代表とする技術進歩により癌のゲノム・エピゲノム・トランスクリプトームの異常が包括的に明らかになるようになり、新たな治療方法の探索に活用されている。特にTCGA (図1)、ICGCなどの国際的なゲノムコンソーシアムは大規模に大量の癌ゲノムデータの集積・解析をすることで、多くの癌腫についてのゲノム異常を明らかとしてきた。論文化された、もしくはされる前のこれらのデータは全て公表され、皆が使えるようにするというのが時代の流れとなっている。データベースを活用することで、自らの研究をサポートするデータを得ることや、逆に新しい研究への足掛かりとすることが可能である。本稿では、公共データベースの紹介とそのデータを用いた解析方法を紹介し

たい。

### 登録される情報の実験手法

公共データベースから得られるデータには多数あるが、まずは実験の種類ごとに得られるデータについて紹介する。

#### 1. 次世代シーケンサーによる解析

次世代シーケンサーによって大量のゲノムを読むことができるようになり、以前はサンガーシーケンスによって長時間かけていたゲノム配列の決定も網羅的もしくは高速に癌ゲノムの異常を同定することができるようになった。癌のゲノム配列を明らかにすることは、癌発生の機序を明らかとするだけでなく、新規の治療ターゲットの同定にも有用であると考えられている。これらの生データには遺伝子多型の情報なども含まれており、取得に申請などの制限があるものもある。