

一般口演B

[KB4] 一般口演 B  
変異頻度に基づくがんゲノムデータの新規クラスタリング手法とその  
臨床応用

2018年6月22日(金) 10:15 ~ 10:45 第2会場 (2階・中会議室201)

---

[KB4] 変異頻度に基づくがんゲノムデータの新規クラスタリング手法とその  
臨床応用

凌 一葦（新潟大学医歯科総合研究科 バイオインフォマティクス分野）

## 変異頻度に基づくがんゲノムデータの 新規クラスタリング手法とその臨床応用

凌 一葦<sup>\*1</sup>, 渡辺 由<sup>\*1</sup>, 永橋 昌幸<sup>\*2</sup>, 島田 能史<sup>\*2</sup>, 市川 寛<sup>\*2</sup>,  
若井 俊文<sup>\*2</sup>, 奥田 修二郎<sup>\*1</sup>

<sup>\*1</sup>新潟大学医歯科総合研究科 バイオインフォマティクス分野

<sup>\*2</sup>新潟大学医歯科総合研究科 消化器・一般外科学分野

### Mutation Frequency based Novel Clustering Method for Cancer Genome Data

Yiwei Ling<sup>\*1</sup>, Yu Watanabe<sup>\*1</sup>, Masayuki Nagahashi<sup>\*2</sup>, Yoshifumi Shimada<sup>\*2</sup>, Hiroshi Ichikawa<sup>\*2</sup>, Toshifumi Wakai<sup>\*2</sup>, Shujiro Okuda<sup>\*1</sup>

<sup>\*1</sup> Division of Bioinformatics, Niigata University Graduate School of Medical and Dental Sciences

<sup>\*2</sup> Division of Digestive and General Surgery, Niigata University Graduate School of Medical and Dental Sciences

抄録: クリニカルシーケンス技術が臨床医学領域で普及しつつある。その中でも、がんゲノムデータを用いた遺伝子異常の解析はがん発症メカニズムの解明と個別化医療の実現に重要な役割を担っている。しかしながら、現在、遺伝子変異を元に分子標的薬によるがん治療方針の判断を効率的に実施する手段は存在しない。そこで、本研究では、異なる遺伝子変異を持つがん患者コホートを変異頻度の高い特異的な遺伝子変異のパターンでクラスターに分類することで、分子標的薬の効率的な選別を可能にする新規クラスタリング手法を開発した。約 400 のターゲットシーケンス遺伝子に加え、分子標的薬のターゲットになる治療効果が期待される遺伝子でのクラスタリングを実施した。その結果、APC や KRAS など大腸がんの特徴的な変異遺伝子をもつクラスターを別けることに成功した。また、日米での承認薬のターゲット遺伝子に注目したクラスタリング結果の比較から、分子標的薬を中心とした治療方針決定にとって最新のがんゲノム知識ベースの必要性が確認された。

キーワード がんゲノム, 変異頻度, クラスタリング, 変異パターン

### 1. はじめに

次世代シーケンサーを用いた DNA 配列解読技術の飛躍的な進歩により、がんのゲノム変異の網羅的な同定というクリニカルシーケンスが可能になっている。すでに欧米を中心に TCGA (The Cancer Genome Atlas) [1] のような大規模ながんゲノムデータ解析プロジェクトが進められ、多くの遺伝子変異の情報がデータベース化されるに至った。その成果により、現在では、がん細胞の持つ遺伝子変異の解析に対して、数百の特定のがん関連遺伝子のみをクリニカルシーケンスするターゲット遺伝子パネル解析が始まりつつある。遺伝子変異の情報からがん発症メカニズムの解明と分子標的薬治療による個別化医療の実現が待ち望まれている。

がんの種類によって、変異を受けやすい遺伝子や変異のホットスポットが存在する。つまり、どの遺伝子に変異を受けているのかということと、その遺伝子の変異頻度の情報は、そのがん化の分子メカニズムを理解する上で非常に重要な因子である。

現在、遺伝子変異を元に分子標的薬によるがん治療方針の判断を効率的に実施する手段は存在しない。そこで、本研究では、遺伝子変異の頻度を考慮できる形でがんゲノムデータをクラスタリングする新しい手法を開発した。本手法を使うことで特定の遺伝子変異によるがん化症例の層別化を可能にする。これを分子標的薬のターゲット遺伝子で実施することで、これまで実現できなかった遺伝子変異の情報から分子標的薬によるがん治療方針の判断を効率的に実施で

きる可能性を示す。

## 2. 方法

### 1) がんゲノムデータ

新潟大学医歯学総合病院消化器・一般外科と関連の病院で採取されたステージ I-IV にわたる大腸がんサンプル 201 例に対して、平均 500 倍のカバレッジでがんに関連する 400 個以上の遺伝子を対象としたターゲットシーケンシング解析が実施された[2]。本研究ではこの日本人大腸がんの遺伝子変異の情報を利用した。

### 2) クラスタリング

同じメカニズムでがんになったサンプル同士は、同じ遺伝子変異を持つ可能性が高い。サンプル間で共通の変異遺伝子の数に基づいた行列を作成し、ノーマライズを実施した。この行列データに最適なクラスタリング結果を得るため、複数の距離、クラスタリング法、クラスター数の組合せから、クラスタリングの安定性解析を実施した。

R ソフトウェアで利用される `clValid` パッケージ[3]を用い、距離、クラスタリング法、クラスター数の組み合わせに対して、APN (average proportion of non-overlap), AD (average distance), ADM (average distance between means), FOM (figure of merit) と、クラスターサイズについての指標を合わせた 5 つのクラスター安定性の指標を計算し、クラスタリング安定性の評価指標とした。

## 3. 結果

### 1) 日本人大腸がん症例でのクラスター

日本人大腸がん 201 症例のがんゲノム解析の結果から、Hypermutation と判定された 17 症例を除き、クラスタリング安定性指標の計算を実施した。その結果、日本人大腸がんの症例では距離の定義として euclidean 距離、クラスタリング手法として ward.D2 を用いてクラスタリングした場合に、15 のクラスターに分割する条件が最も安定性が良いという結果となった。他のクラスタリング手法では期待したように得られなかった APC, TP53, KRAS, SMAD4, PTEN 等の大腸がんのがん化との関連が示唆されている遺伝子が顕著に変異しているグループがクラスター化されていることが示された。

### 2) 分子標的薬のターゲット遺伝子での日米

## 比較

次に、日本および米国 FDA によって承認された分子標的薬がターゲットとする遺伝子を対象にクラスタリングを実施した。その結果、PTEN や KRAS のような日本ではまだ薬事承認されていない医薬品のターゲットでのクラスター化に差が認められ、日米での治療選択に一定の人数の患者で差が生じることが予想された。

## 4. 考察

本来、クラスター毎にその対象となる分子標的薬の効果を知らなければ、より効果的な治療戦略立案に繋がる。現状では、実際に最新の分子標的薬を投薬された患者やそのゲノム情報の不足から、そのレベルにまでは至っていない。遺伝子変異と臨床情報を統合化したがんゲノム知識ベースに加え、迅速にデータ解析するシステムが必須である。変異遺伝子パターンから分子標的薬の効果を推測できれば、治療効果だけでなく医療費抑制の効果も期待できる。

## 5. 結語

本研究では、がんゲノムデータを変異の頻度に注目した形でクラスタリングする方法を開発した。日本人大腸がんのデータから得られたクラスタリングでは、特徴的な変異遺伝子によるクラスターが得られること、さらに、分子標的薬に対する効果が期待できる遺伝子群でのクラスタリングから、現状のがんゲノム医療の可能性および限界について議論できるようになった。

## 参考文献

- [1] Cancer Genome Atlas Research Network: Comprehensive genomic characterization defines human glioblastoma genes and core pathways, *Nature*, 455(7216):1061-1068, 2008.
- [2] Nagahashi, M., Wakai, T., Shimada, Y., et al.: Clinical implications of comprehensive genomic sequencing for precision medicine, *Genome Medicine*, 8(1):136, 2016
- [3] Brock, G., Pihur, V., Datta, S., Datta, S.: `clValid`: An R Package for Cluster Validation, *Journal of Statistical Software*, 25(4):1-22, 2008