

# 生命情報ビッグデータからの 医学的知見の発見技術

福岡 豊 工学部 電気電子工学科 教授

キーワード: 生命情報, ビッグデータ, データサイエンス, 発見技術, システムバイオロジー, 生物学的知見, 遺伝子, がん, 異種データの統合解析

## 概要

本技術は、ビッグデータを医学・生物学的な知見に基づき**グループ化し、統計的に比較**するものです。次世代シーケンサー等の登場によって生命情報データが大規模化し、異種データの統合解析法も提案されています。しかし、そのようなデータからの有用な知見を抽出する方法は確立されていません。従来は、変化の大きい遺伝子のみに注目していたため、大多数の遺伝子のデータを解析に用いていませんでした。その結果、ごく少数の遺伝子の結果のみからでは、有用な知見を発見することができませんでした。これに対して、本技術では、**従来の方法では活用されていないデータも有効に活用し、測定条件数(患者数など)が少ない場合でも、有用な知見を発見できる**ようにしました。特に、培養細胞を用いた基礎実験(創薬など)では測定条件数を増やすことが難しいので、本技術がきわめて有用です。

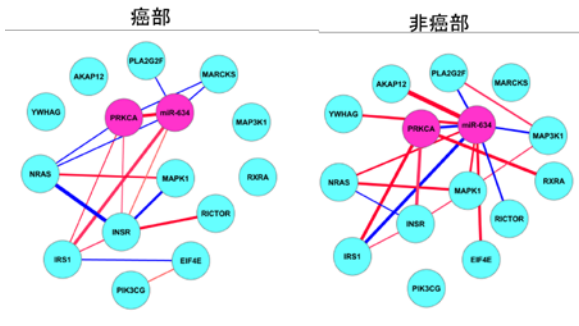
## アピール ポイント

システムバイオロジーの考え方が浸透していますが、異種の生命情報データから**システムレベルでの変化を検出**することを目標にしています。こうした方針に基づき、長年培ってきた解析技術について解説するとともに、がんにおける分子相互作用システムの構造変化についての最新の研究成果も展示します。

## 利用・用途 応用分野

- 小規模な患者データからの疾患バイオマーカーの探索
- 培養細胞を用いた薬剤効果等の検証(創薬など)
- 疾患のシステムレベルでの比較
- 医療・健康分野でのビッグデータに基づく要因解析

| miRNA | p値    |
|-------|-------|
| miR-A | 0.021 |
| miR-B | 0.033 |
| miR-C | 0.034 |



癌部・非癌部の分子相互作用システムの変化

| 遺伝子機能群1(転写に関係する機能)  |                                            |     |          |                                        |
|---------------------|--------------------------------------------|-----|----------|----------------------------------------|
| カテゴリ                | 遺伝子機能                                      | 個数  | p値       | Gene ID                                |
| GO_BP               | transcription, DNA-templated               | 240 | 4.63E-04 | 10346, 147687, 219749, 6774, 7753, ... |
| GO_MF               | DNA binding                                | 208 | 5.65E-04 | 3725, 7494, 147687, 219749, 6774, ...  |
| GO_BP               | regulation of transcription, DNA-templated | 170 | 5.97E-02 | 2146, 4302, 10346, 219749, 147687, ... |
| 遺伝子機能群2(リン酸に関係する機能) |                                            |     |          |                                        |
| カテゴリ                | 遺伝子機能                                      | 個数  | p値       | Gene ID                                |
| GO_MF               | ATP binding                                | 192 | 1.74E-04 | 84930, 79705, 54617, 80216, 29028, ... |
| GO_MF               | protein serine/threonine kinase activity   | 60  | 3.23E-04 | 2580, 84930, 79705, 57538, 80216, ...  |
| GO_BP               | protein phosphorylation                    | 69  | 5.63E-04 | 2580, 79705, 57538, 332, 8573, ...     |
| GO_MF               | protein kinase activity                    | 52  | 6.54E-03 | 2580, 79705, 8573, 6352, 50488, ...    |

培養細胞の低酸素応答に対する解析結果  
— 従来の技術ではこれほど多くの機能は検出できない

## 関連情報

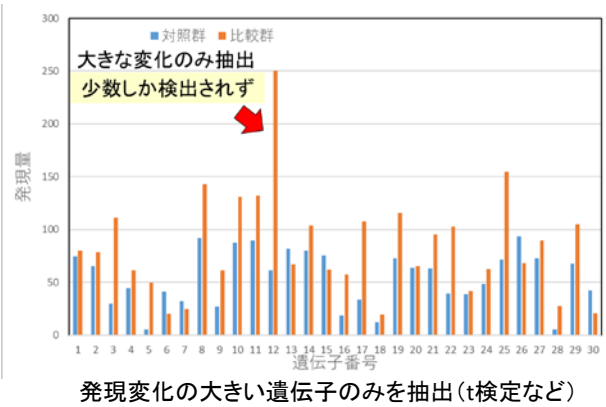
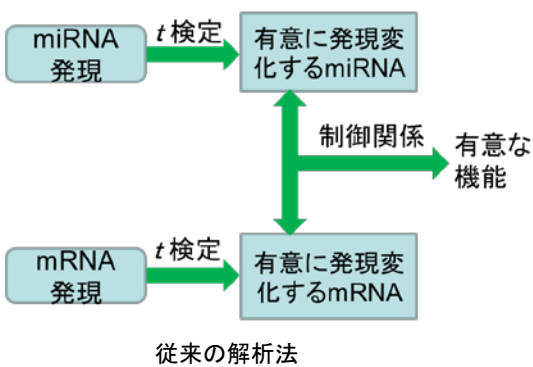
- 関連論文 = 生命情報ビッグデータからの発見技術  
[https://shingi.jst.go.jp/kobetsu/kogakuin/2017\\_kogakuin/tech\\_property.html#pbBlock67655](https://shingi.jst.go.jp/kobetsu/kogakuin/2017_kogakuin/tech_property.html#pbBlock67655)  
バイオインフォマティクスから個別化医療へ  
[https://www.jstage.jst.go.jp/article/jsmbe/57/4-5/57\\_99/\\_article/-char/ja](https://www.jstage.jst.go.jp/article/jsmbe/57/4-5/57_99/_article/-char/ja)
- 関連 URL = <http://www.ns.kogakuin.ac.jp/~wwc1059/>

# 生命情報ビッグデータからの 医学的知見の発見技術

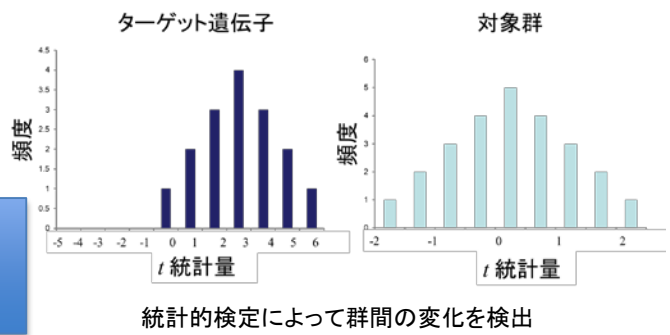
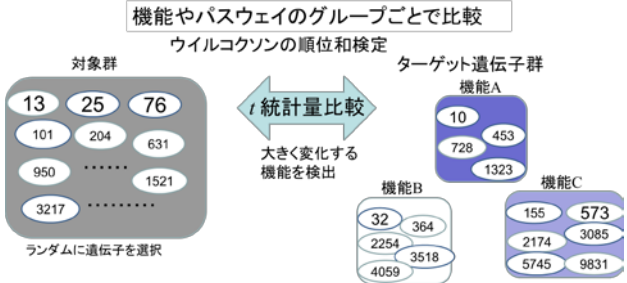
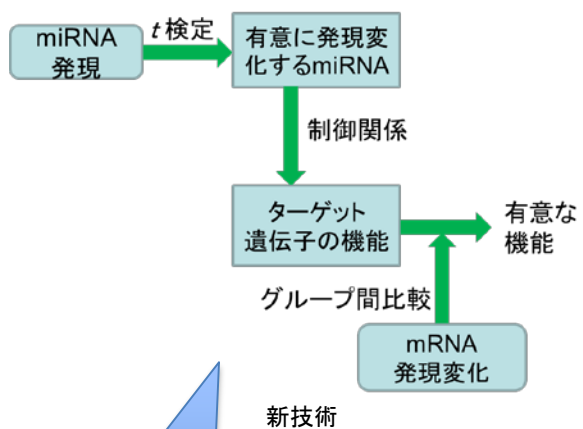
福岡 豊 工学部 電気電子工学科 教授

キーワード: 生命情報, ビッグデータ, データサイエンス, 発見技術, システムバイオロジー, 生物学的知見, 遺伝子, がん, 異種データの統合解析

従来 個別のデータセットに対して検定を行うため、少数の変化しか検出されません。異種データの統合を行うと、複数のデータセットで共通に検出されるデータはほとんどありません。



新技術 本技術では、データ構造を考慮して、これまで捨てていたデータもグループに取り込むことで、精度を高め、適用範囲を広げました。



グループ間比較によって有意な機能の検出率が改善

## 今後の展望

- 外部データベースとの連携で解析性能の向上
- 医療・健康分野でのビッグデータへの適用拡大
- 適用例の増大