

健康な日本人の腸管免疫と腸内細菌データベースの構築

研究分担者 水口 賢司
国立研究開発法人医薬基盤・健康・栄養研究所
バイオインフォマティクスプロジェクト プロジェクトリーダー

<目的> 食事・栄養摂取状況や身体活動・運動など詳細な生活習慣情報の得られた被験者を対象に腸内細菌叢を解析し、腸内細菌叢、腸管免疫、生活習慣データおよび公共のデータベースからの情報を統合した基盤データベースを設計する。それらのデータについてバイオインフォマティクスを用いて横断的に分析することにより、生活習慣、腸内細菌叢、腸管免疫、疾患発症との相互関係を明らかにすることを目的とする。

<方法> 食事・栄養摂取状況、身体活動・運動などの生活習慣のデータおよび腸内細菌叢のデータをデータベース化し、別途構築した遺伝子、タンパク質、疾患、化合物、パスウェイ情報等を統合したデータベースとともに、多変量解析や機械学習等を用いることによって、分子メカニズムや各種測定量の関係を解析する。

<結果> 遺伝子、パスウェイ情報等を鍵とした独自のデータウェアハウス技術を拡張し、化合物や疾患などの、より多様なデータを統一的に解析できる枠組みを構築した。また、腸内細菌叢のデータを米国国立生物工学情報センター(NCBI)の提供する生物種IDとともにデータベース化する予定である。さらに、食事・栄養摂取状況、身体活動・運動などの生活習慣のデータを入手し、予備的な多変量解析と相関解析を行った。その結果、赤血球数と最大酸素摂取量や腹囲と動物性脂肪摂取量などに既知の相関がみられることを確認した。

<まとめ> 本研究では、詳細な生活習慣情報と腸内細菌叢や腸管免疫データについてバイオインフォマティクスを用いて横断的に分析することによって、分子機序の推定や各種測定量の間の相関予測モデルを構築することを目指している。今年度は、少数のデータを予備的に解析し、それぞれのデータの性質を理解した上で、データをどのように格納するかを設計を行った。この設計に基づいてデータベースを構築し、来年度以降データの格納を行い、異なる種類のデータの間の横断的な解析ができる環境を整える予定である。

A. 研究目的

最近の研究によって、腸内細菌叢が食事や肥満、代謝性疾患などと関連することが明らかとなってきている。本研究では、食事・栄養摂取状況や身体活動・運動など詳細な生活習慣情報の得られた被験者を対象に腸内細菌叢を解析し、腸内細菌叢、腸管免疫、生活習慣データおよび公共のデータベースからの情報を統合した基盤データベースを設計する。それらのデータについてバイオインフォマティクスを用いて横断的に分析することにより、生活習慣、腸内細菌叢、腸管免疫、疾患発症との相互関係を明らかにすることを目的とする。

B. 研究方法

食事・栄養摂取状況、身体活動・運動などの生活習慣のデータおよび腸内細菌叢のデータをデータベース化し、別途構築した遺伝子、タンパ

ク質、疾患、化合物、パスウェイ情報等を統合したデータベースとともに、多変量解析や機械学習法等を用いることによって、分子メカニズムや各種測定量の関係を解析する。

(倫理面への配慮)

すべてのデータは匿名化されており、個人を特定する情報は含まれていない。

C. 研究結果

(1) データ統合技術の開発

遺伝子、パスウェイ情報等を鍵とした独自のデータウェアハウス技術を拡張し、化合物や疾患などの、より多様なデータを統一的に解析できる枠組みを構築した(Chen *et al.*, 下記発表論文)。この技術を活用して、生活習慣や腸内細菌叢のデータ統合を実行していくことを計画している(図1)。

(2)腸内細菌叢のデータ解析

腸内細菌叢のデータは16SリボソームRNA遺伝子の配列を次世代シーケンサーで解析し、細菌分類における門、綱、目、科、属、種のレベルで分類したリード数のデータである。これらのデータをサンプル間で比較、分類する方法について検討した。今後得られるデータについては、米国国立生物工学情報センター (NCBI) の提供する生物種IDとともにデータベース化する予定である。

(3)生活習慣・生理指標データの多変量解析

食事・栄養摂取状況、身体活動・運動などの生活習慣のデータの中から数値化可能なデータを抽出したところ321項目あった。これらのデータについて、予備的な多変量解析と相関解析を行った(図2)。その結果、赤血球数と最大酸素摂取量や腹囲と動物性脂肪摂取量などに既知の相関がみられることを確認した。

D. 考察

腸内細菌叢に関する研究は、ここ数年の間に非常に多く行われるようになった。これは次世代シーケンサーの普及により、細菌を培養することなく大規模に解析できるようになったことが大きい。これまでに腸内細菌叢に関する多くの研究がなされ、食事や疾患との関連が明らかにされてきている。しかしながら、腸内細菌叢を身体活動や運動との関わりで解析された例は少ない。本研究では、300項目以上にわたる詳細な身体データや生活習慣のデータと腸内細菌叢や腸管免疫データについてバイオインフォマティクスを用いて統合的に解析する。この解析から、これまでに知られていなかった生活習慣と腸内細菌との関わりが見いだせる可能性があると考えている。そこからどのような腸管免疫や腸内細菌叢を形成することが健康を維持するうえで重要かが明らかとなり、将来的には生活習慣の改善による従来の予防法に留まらず、プロ・プレバイオティクス、新しい治療薬、予防薬や疾患発症予測のバイオマーカーの開発につながる可能性がある。

E. 結論

本研究では、詳細な生活習慣情報と腸内細菌叢や腸管免疫データについてバイオインフォマティクスを用いて横断的に分析することによって、分子機序の推定や各種測定量間の相関予測モデルを構築することを目指している。今年度は、少数のデータを予備的に解析し、それぞれのデ

ータの性質を理解した上で、データをどのように格納するか設計を行った。この設計に基づいてデータベースを構築し、来年度以降データの格納を行い、異なる種類のデータの間の横断的な解析ができる環境を整える予定である。

F. 研究発表

1. 論文発表

Yi-An Chen, Lokesh P. Tripathi, Kenji Mizuguchi: An integrative data analysis platform for gene set knowledge discovery in a data warehouse framework. *Database* (in press).

2. 学会発表

【国内学会: 招待講演】

水口賢司, データウェアハウスによる創薬関連データ統合と解析の実際, 第361回 CBI学会研究講演会, 大阪, 2015.4.24 (招待講演)

水口賢司, 健康な日本人の生活習慣と腸管免疫・腸内細菌データベースの構築: バイオインフォマティクスの視点から, シリーズ「薬づくりの新しいR&Dモデルを探る」第11回「薬づくりと健康食品開発を結ぶ」, 東京, 2016.1.15 (招待講演)

【国内学会: 一般講演】

陳怡安, ロケシュ テリパチ, 水口賢司, 創薬の初期研究における統合データウェアハウスTargetMine, トーゴーの日シンポジウム2015, 東京, 2015.10.5 (ポスター)

長尾知生子, 五十嵐芳暢, 森田瑞樹, 陳怡安, 深川明子, 坂手龍一, 水口賢司, 創薬・疾患研究のためのデータベース検索システム Sagace & Toxygates, トーゴーの日シンポジウム2015, 東京, 2015.10.6 (ポスター)

Chen Y.A., Tripathi L.P., Mizuguchi K., The integration of biological data and its application to drug discovery, CBI学会2015年大会, 東京, 2015.10.27 (ポスター)

水口賢司, 坂手龍一, 深川明子, 五十嵐芳暢, 陳怡安, 長尾知生子, 医薬基盤・健康・栄養研究所の創薬支援データベースとツール, 第38回日本分子生物学会年会第88回日本分子生化学会大会 合同大会, 神戸, 2015.12.1 (ポスター)

G. 知的財産権の出願・登録状況

無し

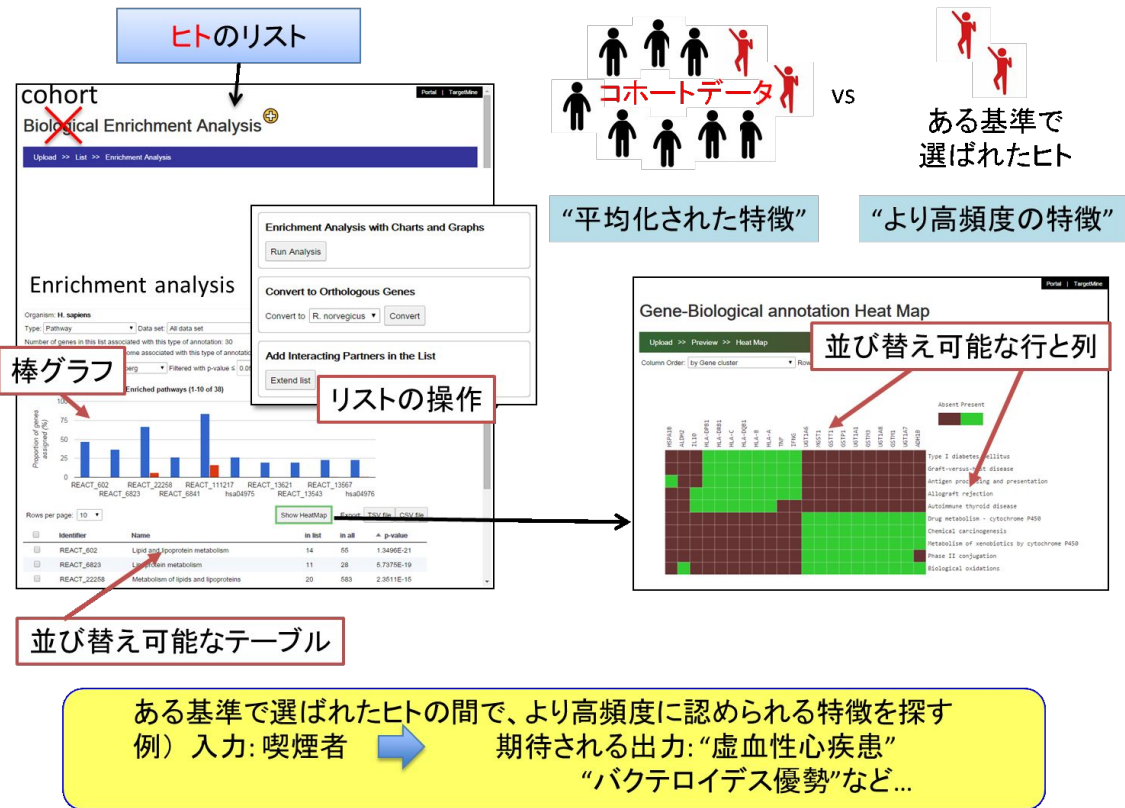


図1. 独自データウェアハウス技術を用いたデータ統合とエンリッチメント解析

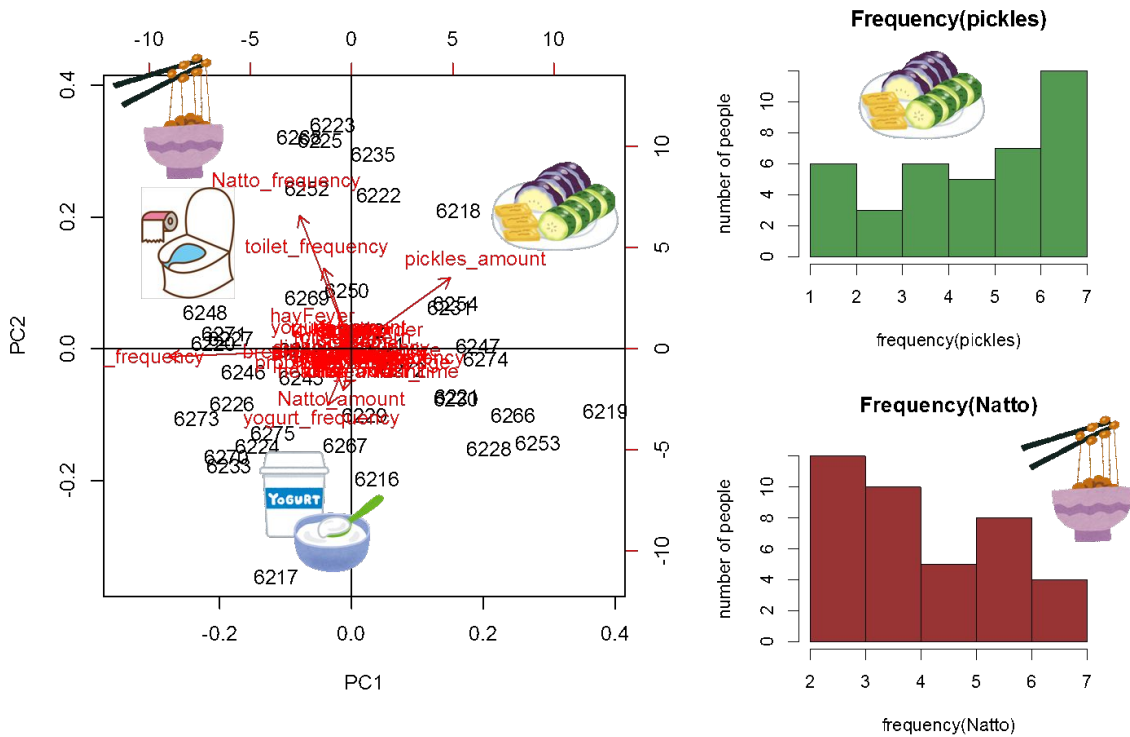


図2. 多項目のデータ間での多変量解析の例(主成分分析)