

未来医療を拓く

2003年、ヒトのゲノム解析が終了した。以来、疾患と関連のある特定の遺伝子情報が次々と明らかになり、医学研究がますます盛んになってきた。

ヒトの23本ある染色体のうち6番染色体にある「HLA」と呼ばれる領域は、ヒトゲノムの中でも、多数の病気の発症を決定づけることで知られている。この領域にあるHLA (Human Leukocyte Antigen) 遺伝子とは、膨大な種類がある白血球の血液型を決める遺伝子群で、自己免疫疾患やアレルギー性疾患、悪性腫瘍、精神疾患、感染症などの発症に関わる。中でも特に免疫に関わる情報を多く含んでいることが分かっている。

ただし、HLA 遺伝子には数十から数百種類もの遺伝子配列があり、あまりにも複雑なために解析

だ、異なる人種のデータが日本人を対象とした疾患発症リスク解析として有効かどうかは不明です。このままでは国際競争に取り残されてしまう恐れもあり、どうしても日本人集団の参照データを作りたいがたかたかです。

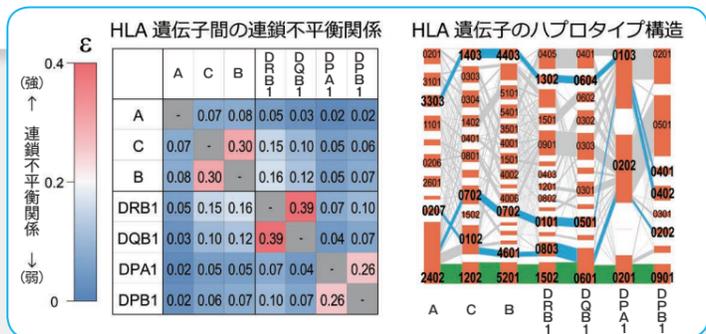
岡田講師は、HLA インピュテーション法を開発したハーバード大学の研究グループの隣の研究室に留学していた経験を持つ。その際に足繁く隣室に通って解析ソフトの開発を学んだのだという。帰国後、東京大学や理化学研究所などの共同研究により、約900人の日本人集団によるHLA 遺伝子配列の参照データを作成。各方面に頼み込んでより多くのデータを集めるのに3年もの歳月を費やした。

バセドウ病発症のバイオマーカーを同定

「日本人の参照データを集めてようやく、人種によるHLA 遺伝子型の違いの有無を知ることができました」

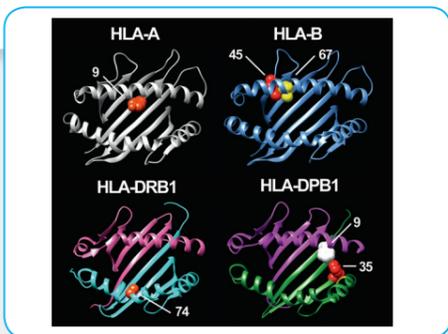
岡田講師はさらにデータを解析するため、高次元ビッグデータを二次元化して可視化する手法を開

(図1) ビッグデータ解析によるHLA遺伝子配列構造の可視化



高次元ビッグデータ可視化手法の開発で、HLA遺伝子構造における人種特異性の可視化に成功。日本人集団に特異なHLAハプロタイプが明らかになった。

(図2) バセドウ病の発症リスクを有するHLA遺伝子部位



日本人集団のバセドウ病のGWASデータに対してHLA imputation法を適用し、原因となるHLA遺伝子部を特定した。

Research Worker Number.16

白血球にある免疫情報の複合体 HLA遺伝子の日本人データを作成

大学院医歯学総合研究科 疾患多様性遺伝学分野 岡田随象 テニュアトラック講師



がなかなか進まなかった。次世代シーケンサーを使ってHLA 遺伝子の塩基配列を調べることも可能だが、1人あたり数万円もかかるなど、コスト面での問題も大きい。

2012年、米国ハーバード大学の研究者らが「HLA インピュテーション」法という解析方法を開発した。「推測する(imputation)」という意味を持つこの手法では、あらかじめ入力しておいた学習データを参照に、コンピュータ上でHLA 遺伝子型の個人差を統計学的に推測。低コストでのHLA 遺伝子解析を可能にした。

約9000人の日本人の参照データを作成

HLA インピュテーション法の参照データとなるのは、HLA 遺伝子の配列やSNP (スニップ) 塩

発。日本人集団に特異的なHLA 遺伝子型の組み合わせや組成を明らかにしたところ、やはり人種別の参照データが必要だと分かったという(図1)。

「アジア人はHLA 遺伝子の個人差が特に顕著だといわれています。同じアジア人でも、シンガポール人や中国人と比べると日本人はかなりの相違点があります。人類の歴史から見ると、HLA 遺伝子は感染症への応答を決めてきたもので、地域特異的な感染症や寄生虫などに応じて地域ごとの遺伝子多型が生じてきたのでしょう」

岡田講師はこの手法を用いて、日本人集団9000人分のバセドウ病の大規模GWAS (ゲノムワイド関連解析) の再解析を行った。その結果、HLA-DPB1 遺伝子の35番目のアミノ酸に変異があった場合にバセドウ病の発症リスクが高くなることが分かった。この遺伝子配列はバセドウ病の疾患バイオマーカーとなり得る(図2)。

「甲状腺機能の異常をもたらす自己免疫疾患の1つであるバセドウ病の発症については、以前からHLA 遺伝子が関連していることが分かっていました。しかし、HL

基多型)などのデータベース。ヒトの30億塩基のゲノム上のうち約1000万カ所に存在しているSNPは、1塩基だけが他の塩基に置き換わっているもので、現在は全ゲノムにおける膨大なSNP データベースがある。このデータをHLA 領域の予測に利用する。

参照データに含まれる人数が多ければ多いほど予測精度が上がります。ハーバード大学の研究グループは約5000人分の参照データを集めた。その結果、関節リウマチのリスク因子である遺伝子多型を見出すなど、この数年間に数々の重要な発見があった。

しかし、日本人の参照データはなかなか作られることがなかった。岡田随象講師は次のように語る。「欧米人に続いて、東アジア人の参照データも作られました。た

A 遺伝子のどの部分が関与しているかは不明でした。それが今回の解析により明らかになったのです」

遺伝統計学の人材育成にも注力

HLA 遺伝子は、疾患発症リスクだけでなく、薬の副作用や効きやすさ、発症後の予後などにも関わる。この領域の遺伝子多型を調べることでヒトゲノムに基づいた「個別化医療」の可能性も見えてくる。

その実現に向けては、膨大なゲノムデータの中から、いかに有益なデータを引き出すことができるかが鍵となる。そのため、適切な解析と解釈がさらに重要になると岡田講師は話す。

「私が専門とする遺伝統計学は、海外では専門家が多い分野ですが、日本ではほとんど存在していません。この分野が発展しなければ今後のゲノム解析やゲノム医学のポータルネットワークになってしまいう可能性ががあります。医師は臨床経験を通じて論理的に考える能力を身につけており、本来数学が得意な人も多いので、この分野に興味を持ってもらいたいですね」

おかだ・ゆきり
2005年東京大学医学部医学科卒業。2011年同大学院医学系研究科内科学専攻博士課程修了(医学博士)。2009年東京大学医学部附属病院にて初期研修プログラム。その後、東京大学医科学研究所ヒトゲノム解析センターGCOリサーチアソシエイト、Broad Institute、Harvard Medical School (Brigham and Women's Hospital)、日本学術振興会海外特別研究員などを経て、2013年より現職。研究分野は遺伝統計学。