142 糖タンパク発現による大腸癌の新規腫瘍マーカーの開発

九州医療センター 消化管外科・がん臨床研究部　　蓮田　博文

【目的】大腸癌、胃癌、食道癌を含む消化管癌は、世界的にも最も罹患率が多く、癌の早期発見が重要である。本研究では、実用性と実現可能性に重点を置き、従来の腫瘍マーカーに加え血清糖ペプチドに着目し、包括的診断モデルを開発した。このモデルは、濃縮糖ペプチド（EGP）の詳細なプロファイリングと機械学習ベースの包括的血清糖ペプチドスペクトル解析（CSGSA）アプローチを統合することによってより高い診断精度を実現し、癌の早期発見を可能にすることを目的としている。

【方法】本研究には大腸癌患者（CRC）296 例、胃癌患者（GC）180 例、食道癌患者（EC）42 例および健常者（HC）590 例を後方視的に登録し、血清サンプルを分析した。従来の 9 つの腫瘍マーカー（CEA、CA19-9、CYFRA、AFP、PSA、CA125、CA15-3、SCC 抗原、NCCST-439）を測定し、1,688 種類の EGP を、液体クロマトグラフィー質量分析法を用いて同定した。CSGSA を用いて、EGP を従来のマーカーと統合し、ニューラルネットワーク（NN）を含む機械学習モデルに組み込み、診断フレームワークの開発と検証を行った。

【結果】約 10,000 個の EGP を包括的に解析した結果、CRC、GC、EC を HC と区別できる 2 つの糖ペプチド、α1-アンチトリプシン（AT271-FSG）とα2-マクログロブリン（MG70-FSG）を同定した。診断精度を向上させるため、従来の腫瘍マーカーと AT271-FSG および MG70-FSG の糖ペプチド、さらに 1,688 個の EGP を統合した包括的な機械学習モデルを開発した。これらのバイオマーカーの寄与を明らかにするために、3 つのモデルを用いて評価した。モデル 1 では 9 つの既存の腫瘍マーカーのみを用い、モデル 2 では AT271-FSG と MG70-FSG の糖ペプチドを加え、モデル 3 では全てのマーカーと 1,688 個の EGP から得られた 100 の主要な特徴を PCA で処理した。モデル 3 を開発する際、XGBoost と NN アーキテクチャの両方で評価した。その結果、NN モデルは、XGBoost よりも、グループの区別において優れた性能を示すことが明らかになった（図 1A）。解析の結果、癌グループを HC から区別するための AUC スコアは、0.854（モデル 1）、0.892（モデル 2）、0.962（XGBoost によって確立されたモデル 3）、0.977（NN によって確立されたモデル 3）であり、既存の腫瘍マーカーを有意に上回ることが示された（図 1B）。さらに、モデル 3 の予測値を、以下の式を用いて CSGSA スコア（0～10）に変換した：CSGSA スコア＝−log10（1-モデル 3 の予測値）図 1C は、CRC、GC、EC の CSGSA スコアヒストグラムである。データは、CSGSA が CRC、GC、EC を HC から効果的に鑑別することを示唆し、AUC 値はそれぞれ 0.966、0.992、0.995 であった。CRC に対する感度は 54.3%、GC に対する感度は 61.6%、EC に対する感度は 67.7%であり、特異度は 99.9%を超えた。CRC、GC、EC の PPV は、それぞれ 54.5%、35.5%、11.1%であった。

【考察】糖ペプチドプロファイリングと NN ベースの機械学習モデルを組み合わせることで、CRC、GC、EC を健常対照から識別することに成功した。このアプローチにより AUC 値は各々0.966、0.992、0.995 および PPV は 54.5%、35.3%、11.1%という優れた値を達成し、既存の腫瘍マーカーを凌駕した。今回の研究結果は、糖鎖ペプチドを活用した新しい診断の枠組みを提示するものであり、癌診断の未来に向けた重要な一歩となる。

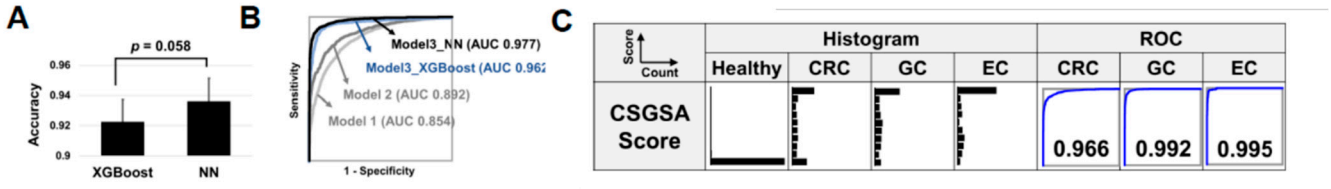


図 1. 機械学習モデルと CSGSA スコアの包括的評価