

令和 5 年 4 月 20 日現在

機関番号：82504

研究種目：若手研究

研究期間：2020～2022

課題番号：20K17570

研究課題名（和文）膵癌における人工知能技術を用いた臨床画像からの遺伝子変異・発現予測

研究課題名（英文）Gene mutation/expression prediction from clinical images using artificial intelligence technology in pancreatic cancer

研究代表者

岩立 陽祐（Iwatate, Yosuke）

千葉県がんセンター（研究所）・肝胆膵外科・医長

研究者番号：10815731

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：Radiogenomics解析として、臨床病理学に有用な遺伝子の網羅的検索、12例の膵癌組織でRNA-seqの実行、および107例の免疫染色（IHC）とCT画像を用いた機械学習による遺伝子の発現予測モデルを構築した。ITGB1、ITGAVが同定され、IHCとの間にも有意な相関を認め（ $r=0.552$ $P=0.118$ 、 $r=0.625$ $P=0.039$ ）、IHC高発現群では予後が悪化し、（ $P=0.035$ 、 0.009 ）。再発も同様（ $P=0.028$ 、 0.003 ）であった。ITGAVの予測モデルは一定の検出能を認め（ $AUC=0.697$ ）、高発現と予測された群も予後不良であった（ $P=0.048$ ）。

研究成果の学術的意義や社会的意義

診断時に進行していることが多く、5年生存率が10%程度と予後不良な膵臓癌において、診断のための侵襲的な生検等の検査や外科治療は専門施設やハイボリュームセンターでないと施行できない場合が多い。今回CTによる予後予測や今後治療の一助のなりうる可能性がある遺伝子マーカーの同定およびその発現予測が可能であった。今後症例を重ねることでCT画像等の臨床画像から非侵襲的に予後や治療マーカーの同定が行うことができれば、高額で時間がかかる遺伝子検査でなく簡易で費用対効果が高く、専門機関でなくても比較的短期間で結果が得られる検査となることも考えられる。

研究成果の概要（英文）：As a radiogenomics analysis, RNA-seq was performed on 12 pancreatic cancer tissues for an exhaustive search of genes useful for clinical pathology. A gene expression prediction model was constructed by machine learning using immunostaining (IHC) and CT images of 107 cases. ITGB1 and ITGAV were identified and significantly correlated with IHC ($r=0.552$ $P=0.118$, $r=0.625$ $P=0.039$). The prognosis worsened in the IHC high expression group ($P=0.035$, 0.009). Recurrence was similar ($P=0.028$, 0.003). The predictive model for ITGAV showed a certain detectability ($AUC=0.697$), and the group predicted to have high expression also had a poor prognosis ($P=0.048$).

研究分野：肝胆膵外科

キーワード：膵臓癌 膵管腺癌 Radiogenomics ITGAV ITGB1

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

(1) 膵癌は予後不良な悪性腫瘍であり、本邦で消化器癌死亡の第4位と主要な死因となっている。根治的切除が治癒の可能性を引き出す唯一の手段と考えられるが、根治切除をしえた症例に限っても、その5年生存率は20%程度である(Oettle H, et al. JAMA.2013)。それは現在まで著効しうる有効な治療法が確立されていないからである。そこで、遺伝学的情報に基づき患者の個々のレベルにおいて分析・選択し最適な治療を行うことでその解決に期待が持たれている。しかし、遺伝学的検索に数十万円のコストそして数週間の時間を要しており、全ての癌患者にを提供する為には、コストの他、解析のための時間や設備投資、人材育成等の多くの解決すべき問題がある。さらに、組織を得るための外科治療や生検検査には身体には比較的大きな侵襲が伴ってくるのが通常である。一方、画像診断は診断法として現在、通常の癌診療の必須の検査として比較的安価で侵襲性も低く施設を問わず日常的に汎用されており、体内の悪性腫瘍を含めた診断に大きく貢献している。これは、ひとえにCT等の一般画像が組織病理と強く関連するためである。こうした一般臨床画像から遺伝子情報を予測することができれば費用対効果の高い検査プラットフォームが構築されることになる。

(2) 近年になり癌の遺伝学的背景は次世代シーケンサーの登場により解明が飛躍的進み、こういった膨大な遺伝情報を扱う Genomics 研究が進んできている。またCT、MRI、PET-CTといった一般臨床画像から量的な特徴を多数抽出し解析する技術である Radiomics 研究が行われるようになり、これらを融合させた Radiogenomics 解析が行われるようになってきている。

(3) 膵臓癌にておいてはDNA,RNA シークエンスにおける多症例の解析、研究結果が大規模に行われ、遺伝学的背景が解明されてきているが、一般臨床画像を用いた Radiomics 解析研究の報告は少なく、Radiogenomics 解析はほとんど行われていない。

(4) 比較的新しい技術であるシーケンスデータと Radiomics で抽出されるデータ、これら膨大な量的データを扱う Radiogenomics 解析において人工知能を用いることで、さらなる独自性、創造性を有し、一般臨床画像から遺伝子情報を予測するモデル構築の可能性と正確性の向上が期待された。

2. 研究の目的

本研究の目的は、特に難治性と考えられる膵癌において、時間、コスト、侵襲を伴う遺伝子情報を一般画像から導くという Radiogenomics 解析を臨床応用可能なレベルまで引き上げることである。

3. 研究の方法

(1) ゲノム異常、及び網羅的遺伝子発現プロファイルの解析 (30例程度)

次世代シーケンサー(NGS)を用いて、網羅的なゲノム異常、mRNA、ncRNAの解析を行う。

(2) 画像解析 (Radiomics) (30例程度)

本研究での画像解析は、各患者の造影CT、MRI、代謝画像としてFDG-PETを使用する。これら画像情報を、造影される腫瘍と推定される部位をマニュアルで選択し、3D segmentationを行う。画像的特徴量を抽出し数値化する。

(3) ゲノム情報と抽出された画像的特徴量を千葉大学 大学院工学研究科・工学部に解析依頼をし、有意なAI技術の選択、特徴量選択と遺伝子発現予測モデルの構築を行う。

(4) 統合解析と検証 (100例程度)

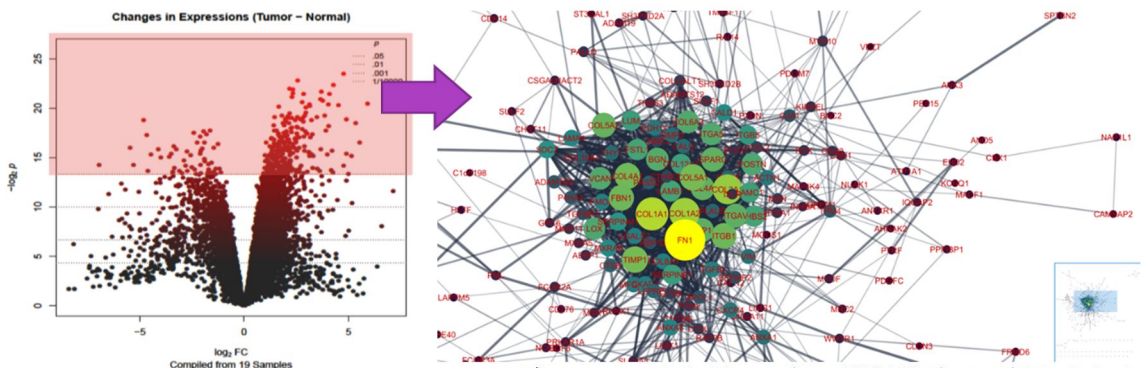
これらの解析により絞り込んだ因子に対して薬剤感受性や予後について多検体を用いたリアルタイム半定量RT-PCRや免疫染色等を用いて検証を行う。さらにこれら因子の発現が、構築された予測モデルでも予測可能かの検証を行う。

4. 研究成果

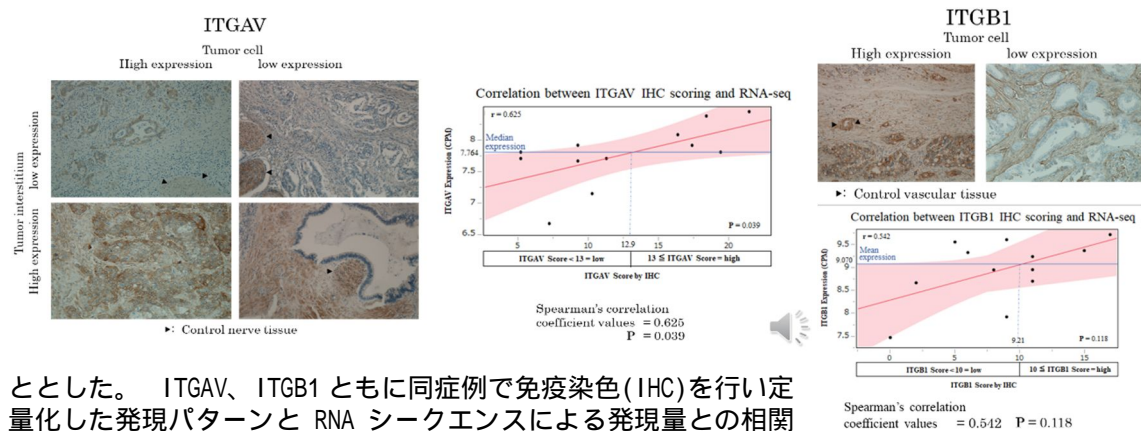
(1) 研究中に術前化学療法の有用性が認められたため、化学療法による遺伝子修飾がない手術症例、検体が著減したため、臨床病理学に有用な遺伝子の網羅的検索には12例を対象としての膵臓癌組織と周辺正常膵組織とでRNAシーケンスが実行された。12症例での少数例では画像解析による発現予測モデル構築に際し十分でないため、検証症例を含む107症例すべてでモデル構築および検証を行う、Cross validationを採用することとした。

Total RNA シークエンスを行い、まずはmRNAの発現解析を行った。正常膵と比べ発現量が有意

に多い、もしくは少ない遺伝子 ($P < 0.0001$) を発現変動遺伝子として 312 個同定し、タンパク相互間ネットワーク解析を行った。これはタンパクもしくは RNA 間の既知の相互関係が登録さ



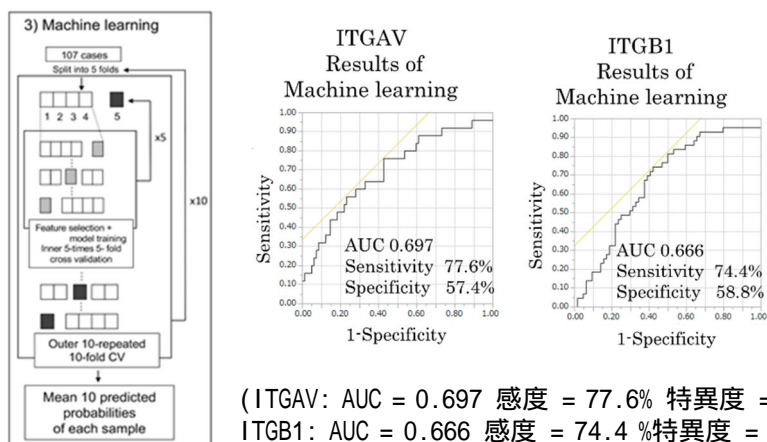
れたデータバンクを用いて相互関係構造の可視化を行うもので Cytoscape (<https://cytoscape.org/>) を用いて行った。ネットワークの中核にある RNA 群をハブ遺伝子とした。これらが臨床病理学的に有用かどうかの判定に、RNA の発現量と予後のデータが登録してあるオープンソースのデータバンクである The Cancer Genome Atlas (TCGA) から引用し予後解析を行える R2: Genomics Analysis and Visualization Platform (<https://hgserver1.amc.nl/cgi-bin/r2/main.cgi>) を用いて事前解析を行った。結果、ITGAV, ITGB1, TIMP1 ($P = 0.036, 0.033, 0.041$) で予後の悪化と関連がみられ上位 2 個の ITGAV, ITGB1 についての検証解析を行い、および有意であった場合は予測モデルの構築を行うこ



ととした。ITGAV、ITGB1 とともに同症例で免疫染色 (IHC) を行い定量化した発現パターンと RNA シークエンスによる発現量との相関解析を行うと相関がみられた。($r = 0.625, P = 0.039, r = 0.552, P = 0.118$) 高発現群となる IHC 染色パターンを基に多症例での検証と行うこととした。ITGAV, ITGB1 高発現群 (IHC) による 107 例での検証ではともに独立した予後規定、および再発因子となった。($P = OS: 0.009, 0.035, DFS: 0.003, 0.028$) したがって、ITGAV, ITGB1 に対する一般臨床画像からの発現予測モデルの構築を試みることにした。

(2) 107 例の膵臓癌患者の術前造影 CT を使用、MRI、PET-CT を参考とし 3D Slicer image computing platform (<https://www.slicer.org/>) を用いて腫瘍をマニュアルにて囲い VOI (volume of interest) を抽出しその範囲および周囲+8 mmまでの画像解析を行い、Python (<https://www.python.org/>) および、その追加ライブラリである PyRadiomics v2.2.0 (<http://www.radiomics.io/pyradiomics.html>) を用いて得られた画像的特徴量から 3748 個の特徴量を抽出した。

(3) 上記、画像的特徴量をランダムフォレストやニューラルネットワーク、サポートベクターマシン、xGBoosting (xGBoosting) などの機械学習法を Five-fold cross validation にて検討し、ハイパーパラメータの設定を行ったところ xGBoosting による判別能が最もよい結果となった。



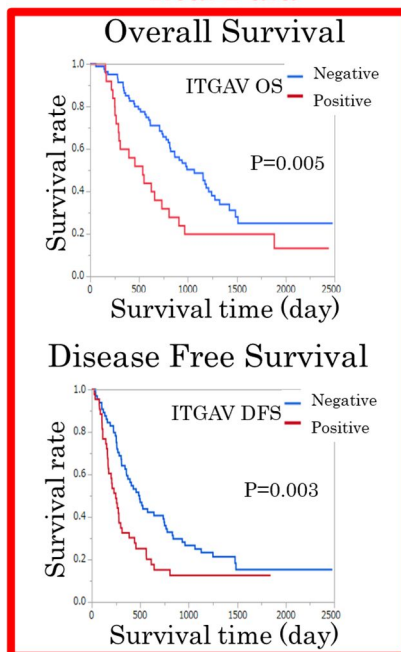
(ITGAV: AUC = 0.697 感度 = 77.6% 特異度 = 57.4%, ITGB1: AUC = 0.666 感度 = 74.4% 特異度 = 58.8%)

R version 3.5.1 (R
Foundation for
Statistical Computing)

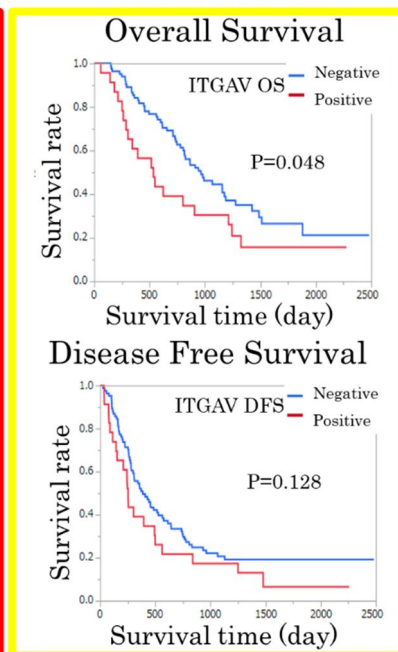
(4) これら予測モデルを用いた全生存期間(OS)、無再発生存期間(DFS)では、ITGAV 高発現群と予測された群で有意に全生存期間の短縮がみられた。
($P = 0.048$)

ITGAV の発現において CT 画像より一定の検出能を持った予測モデルの構築に成功したが、ITGB1 においては十分な結果を得ることはできなかった。多症例での検討や MRI, PET-CT 等の他モダリティを合わせた特徴量の抽出等が検討された。

Real Data



Predicted Data



5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 1件/うちオープンアクセス 2件）

| | |
|--|-------------------------|
| 1. 著者名 YOSUKE IWATATE, HAJIME YOKOTA, ISAMU HOSHINO, FUMITAKA ISHIGE1, NAOKI KUWAYAMA, MAKIKO ITAMI, YASUKUNI MORI, SATOSHI CHIBA, HIDEHITO ARIMITSU, HIROO YANAGIBASHI, WATARU TAKAYAMA, TAKASHI UNO, JASON LIN, YUKI NAKAMURA, YASUTOSHI TATSUMI, OSAMU SHIMOZATO and HIROKI NAGASE | 4. 巻 60 |
| 2. 論文標題 Machine learning with imaging features to predict the expression of ITGAV, which is a poor prognostic factor derived from transcriptome analysis in pancreatic cancer | 5. 発行年 2022年 |
| 3. 雑誌名 International Journal of Oncology | 6. 最初と最後の頁 1791-2423 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.3892/ijo.2022.5350 | 査読の有無 有 |
| オープンアクセス オープンアクセスとしている(また、その予定である) | 国際共著 該当する |

| | |
|---|---------------------------|
| 1. 著者名 Iwatate Yosuke, Hoshino Isamu, Yokota Hajime, Ishige Fumitaka, Itami Makiko, Mori Yasukuni, Chiba Satoshi, Arimitsu Hidehito, Yanagibashi Hiroo, Nagase Hiroki, Takayama Wataru | 4. 巻 123 |
| 2. 論文標題 Radiogenomics for predicting p53 status, PD-L1 expression, and prognosis with machine learning in pancreatic cancer | 5. 発行年 2020年 |
| 3. 雑誌名 British Journal of Cancer | 6. 最初と最後の頁 1253 ~ 1261 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41416-020-0997-1 | 査読の有無 有 |
| オープンアクセス オープンアクセスとしている(また、その予定である) | 国際共著 - |

〔学会発表〕 計3件（うち招待講演 0件/うち国際学会 0件）

| |
|---|
| 1. 発表者名 岩立陽祐, 星野敢, 横田元, 石毛文隆, 伊丹真紀子, 千葉聡, 有光秀仁, 柳橋浩男, 高山亘 |
| 2. 発表標題 Radiogenomics: Machine learning with imaging feature predict ITGB1, ITGAV expression which are poor prognostic factors derived from transcriptome analysis in pancreatic cancer. |
| 3. 学会等名 第80回 日本癌学会学術総会 |
| 4. 発表年 2021年 |

| |
|--|
| 1. 発表者名 岩立陽祐, 星野敢, 横田元, 石毛文隆, 伊丹真紀子, 千葉聡, 有光秀仁, 柳橋浩男, 高山亘 |
| 2. 発表標題 Radiogenomics解析を用いた 膵癌におけるトランスクリプトームおよび予後の予測 |
| 3. 学会等名 第120回日本外科学会定期学術集会 |
| 4. 発表年 2020年 |

| |
|--|
| 1. 発表者名 岩立陽祐, 星野敢, 横田元, 石毛文隆, 伊丹真紀子, 千葉聡, 有光秀仁, 柳橋浩男, 高山巨 |
| 2. 発表標題 Radiogenomics analysis: prediction of transcriptome and prognosis in pancreatic cancer. |
| 3. 学会等名 第75回日本消化器外科学会総会 |
| 4. 発表年 2020年 |

〔図書〕 計0件

〔産業財産権〕

〔その他〕

| |
|--|
| <p>難治性消化器癌を対象とした治療効果予測、予後予測のためのRadiogenomics理論の創出 https://www.pref.chiba.lg.jp/gan/center/gaiyo/rinsyoukenkyuuosirase.html</p> |
|--|

| | | |
|---------------------------|-----------------------|----|
| 6. 研究組織 | | |
| 氏名 (ローマ字氏名) (研究者番号) | 所属研究機関・部局・職 (機関番号) | 備考 |

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

| | |
|---------|---------|
| 共同研究相手国 | 相手方研究機関 |
|---------|---------|