

令和 6 年 6 月 17 日現在

機関番号：12602

研究種目：基盤研究(C) (一般)

研究期間：2021～2023

課題番号：21K07997

研究課題名(和文) B型肝炎ウイルスドライバー変異の高感度検出系の開発およびPRSモデルの構築

研究課題名(英文) Highly Sensitive Detection for HBV Driver Mutations and Construction of a PRS Model

研究代表者

西田 奈央 (Nishida, Nao)

東京医科歯科大学・難治疾患研究所・准教授

研究者番号：50456109

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：肝がん組織中に低頻度に存在する変異を検出するためのがんゲノム解析パイプラインを構築した。研究期間中に収集した182症例の肝がん組織試料から抽出したDNAを対象としたターゲットキャプチャーシーケンシングを実施し、現在、低頻度変異の解析を実施中である。全672症例のB型肝炎患者の生体試料を研究協力施設から新たに収集し、肝がんを発症するまでの日数を予測するモデルの構築した。4つの項目(血液検査時年齢、性別、Log10AFP、HLA-A*33:03の有無)を入れたモデルは、肝がん発症累積確率(1年後、3年後)を用いたROC解析で、AUC=0.862、0.863となることが分かった。

研究成果の学術的意義や社会的意義

本研究において、ディープシーケンシングにより肝がんドライバー変異を高感度に検出する測定系を開発した。リキッドバイオプシーや肝生検試料の中に微量に存在する肝がんドライバー変異を高感度かつ高精度に検出することが可能となれば、B型慢性肝炎患者の中で肝がんを発症するリスクの高い患者を囲い込むことが可能となる。治療が必要な患者を選択することで適切な治療や投薬を行うことが可能になると期待される。

研究成果の概要(英文)：The cancer genome analysis pipeline for detecting the mutation which existed in the liver cancer tissue at the low frequency was constructed. Targeted capture sequencing was performed on DNA extracted from 182 liver cancer tissue samples collected during the study period, and low-frequency mutations are currently being analyzed. Liver tissue samples of all 672 hepatitis B patients were newly collected from collaborating institutes, and a model was constructed to predict the number of days until liver cancer developed. ROC analysis using cumulative probability of developing liver cancer (One year, three years from now.) showed that the model including 4 items (Blood test Age, sex, presence/absence of Log10AFP, HLA-A*33:03) had AUC=0.862 and 0.863, respectively.

研究分野：ゲノム解析

キーワード：ターゲットキャプチャーシーケンシング B型肝炎 体細胞変異

1. 研究開始当初の背景

シーケンス技術の著しい発展により全ゲノムシーケンス(WGS)が可能となり、2016年には藤本らが日本人300症例の肝臓がんの腫瘍と正常DNAを対象としたWGSの結果を報告した(Fujimoto A et al. Nat Genet 2016)。その結果、既知のがん関連遺伝子(TERT遺伝子など)のゲノム構造異常だけでなく、新規のがん遺伝子のゲノム構造異常、B型肝炎ウイルス(HBV)やアデノ随伴ウイルス(AAV)の肝臓がんゲノムDNAへの組み込み(virus DNA integration)、遺伝子発現に影響を及ぼす可能性のある非コードRNAの変異などが検出された。さらに中国からZhao LHらが癌部と非癌部の組織を対象としたRNAシーケンシングを実施し、癌組織においてテロメア領域にHBVゲノムの組み込みが高頻度に存在することを報告した(Zhao LH et al. Nat Commun 2016)。がんゲノム異常は多様性に富んでいることが知られており、2013年には30種類のがん組織を対象としてがん組織における体細胞突然変異パターンの解析が実施され、肝臓がん(HCC)は胃がんや子宮がんと共に最大の6種類のパターンが混在することが報告された(Ludmil B. Alexandrov et al. Nature 2013)。

B型肝炎患者の中で肝がんを発症するのは10%程度であり、肝がんの発症をできるだけ早期に検出する分析方法の開発が求められている。組織生検でがんの有無を確認することが標準的な方法であるが、リキットバイオプシーの一つであるCTC(血中循環腫瘍細胞)や死滅した細胞から放出されるcirculating tumor DNA(ctDNA)が血液中のバイオマーカーとなるか検討する研究が近年数多く報告されている(Chen VL et al. Clin Gastroenterol Hepatol 2020)。しかしながら、大部分のがん種には確立されたバイオマーカー(特定のDNA変異)が見いだされていないため、ctDNAを介した発がんの追跡ができないという問題がある。

B型肝炎患者における肝がんの発症を高い精度で診断するためには、日本人における肝がん発症の直接の原因となるバイオマーカー(肝がんドライバー変異)の同定が必須となる。藤本らの報告によりTERT遺伝子変異が肝がんドライバー変異の一つであることが明らかとなっているが、肝発がん診断の予測精度を高くするためには、より多くの肝がんドライバー変異を見つけ出す必要がある。また、血液中に微量に含まれるctDNAや組織生検試料に含まれる微量の肝がん細胞に由来する肝がんドライバー変異を正確に検出するためには、高速シーケンスを用いた高感度検出系の開発が最も有効と考える。

2. 研究の目的

B型肝炎を背景とする肝がん症例を対象としたがんゲノム解析を実施し、バイオマーカーとなる肝がんドライバー変異の同定を目指す。本研究では、SIFTやPolyPhenといったソフトを使用してミスセンス変異の注釈付けを最初に行い、各遺伝子が有する変異の質と量を考慮したburden testによってドライバー遺伝子を同定する。さらに、リキットバイオプシー(もしくは組織生検試料)を対象としたディープシーケンスにより肝がんドライバー変異を高感度に検出する測定系の開発を本研究の目的とする。肝がんドライバー変異を高感度に検出することで、B型肝炎患者の中で肝がんを発症するリスクの高い患者を囲い込むことで、適切な治療や投薬を行うことが可能になると期待される。

申請者は、NCCオンコパネル(がんに関連する114遺伝子)に加えて、藤本らが見出した肝がんに関連する30個の遺伝子変異、肝がんとの関連が明らかとなっているTERT遺伝子全領域、HBV integrationを同定するためのHBV配列(Genotype A、B、C)、がん免疫関わるPDL1遺伝子全領域、を対象としたカスタムキャプチャーパネルを構築済みである。さらに、ディープシーケンスによって得られる膨大なデータを解析するためのがんゲノム解析パイプラインの構築もすでに完了している(図1)。国内の協力施設(倫理承認済:九州大学、東京大学、久留米大学)から提供される肝細胞がん組織を使用して本研究を実施する体制が整っており、また本研究を実施するための倫理承認(代表施設:国立国際医療研究センター)も得られていることから即時の研究実施が可能である。

3. 研究の方法

研究開発期間中に、研究協力施設である東京大学および久留米大学において収集された200症例の肝がん患者由来の肝がん組織試料(がん部、非がん部)を対象として、カスタムキャプチャーシーケンスによるがんゲノム解析を実施する。カスタムキャプチャーシーケンスで得られる膨大なゲノムデータを解析するためのがんゲノム解析パイプラインの構築はすでに完了(AMED肝炎等克服緊急対策研究事業2019-2021年度)している。さらに、国立遺伝学研究所のスーパーコンピューターシステムの利用申請も完了しており、遺伝研スパコン上でのがんゲノム解析パイプラインの実装が可能である。本研究で得られる肝がん患者の体細胞変異データから、日本人における肝発がんのドライバー変異(SNV、InDel、SV(構造多型)、Virus integration



凍結組織由来DNAを対象としたがん変異の解析

工程	ゲノム解析	がんゲノム解析	
クオリティチェック/ トリミング/ フィルタリング		FastQC Cutadapt FASTX-Toolkit PRINSEQ	
マッピング/アライメント		BWA mem	
重複リードの検出/ 削除クオリティスコアの補正		SAMtools Picard GATK 3.8	
マッピング/ アライメント後 クオリティチェック		Picard BEDTools FastQC BCFtools SAMtools R	
変異検出・ フィルタリング	SNV	GATK 3.8	MuTect2 (Fisher検定) 頻度法 (Fisher検定)
	Indel	GATK 3.8	MuTect2 (Fisher検定)
	CNV	XHMM	CloneCNA
ウイルス挿入部位検出	Virus integra tion		Virurs-Clip GENE-IS BATVI
構造多型の検出	SV		SvABA
アノテーション			ANNOVAR

がんゲノム解析パイプライン（構築済）

肝がん関連変異の高感度検出系の開発

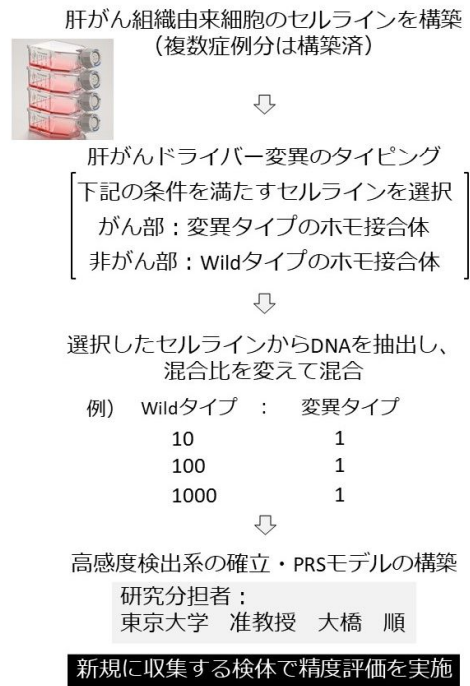


図 1 研究の概要

など)の同定を目指す。さらに、がん組織由来細胞(がん部、非がん部)のセルラインを用いて、見いだされた肝がんドライバー変異の高感度検出系の構築を目指す。複数の肝がん患者から採取したがん組織由来細胞のセルラインはすでに構築済である。本研究で見いだされる肝がんドライバー変異の遺伝子型を各セルラインで決定したうえで、変異型と野生型の変異を有するセルラインの混合比を変えることで、肝がんドライバー変異の高感度検出系を構築する。また、肝がんドライバー変異を有することによる肝発がんリスクを算出するための PRS(Polygenic Risk Score) モデルを構築する。具体的には、ドライバー遺伝子上の変異の質と量によってその遺伝子のスコアを与え、臨床パラメタを調整したうえで肝発がんのリスクとの関連を(多項)ロジスティック回帰分析によって評価する。得られた P 値と偏回帰係数(オッズ比)から、有用な遺伝子と臨床パラメタを用いた PRS モデルを構築する。一般的には、症例を base data と target data に分割して PRS モデルを構築するが、本研究では並べ替えによって多数回(10,000回以上)分割操作を行うことで、より頑健な PRS モデルを開発する(「R」による解析基本プログラムは作成済み)。研究分担施設である九州大学において新規に収集する肝がん患者由来肝細胞がん組織試料および血液試料(50症例を目標)を対象として、カスタムキャプチャーシーケンスによる高感度検出系の精度評価を行う。肝がん患者の血清試料中に微量に含まれる ctDNA 由来のドライバー変異が正確に検出できることを確認するために、同一患者の肝細胞がん組織中のドライバー変異の検出も合わせて行うこととする。

4. 研究成果

(1)肝がん関連変異の高感度検出系の開発

肝がん組織のがん部および非がん部から DNA を抽出した。がん部由来の DNA と非がん部由来の DNA の混合比を変えて、がん部由来 DNA が 100%から 0.14%となるように希釈した(表 1 の Dilution)。合計 8 サンプルを対象として、カスタムキャプチャーシーケンスを実施した。がんゲノム解析パイプラインの MuText2 の実行結果をもとに、がんと正常の reference タイプのリード数と変異リード数の Fisher の正確確率検定、がんと正常組織のアレル頻度などを考慮して変異候補を選択した。

(i) HLADP135-1 から、Fisher の正確確率検定の P 値 < 0.01、がんの最小アレレル頻度 0.05、がんのアレレル頻度 > 正常のアレレル頻度のものを変異候補として選出した。8 個の候補が選ばれた。
(ii) (i) で選出した候補を正解として、(i) で検出された候補以外の変異は偽陽性として、作成したフィルタープログラムを実行し評価した。

フィルタープログラムのアルゴリズムは以下の通りとした。

- ・がんと正常の reference タイプのリード数と変異リード数の Fisher の正確確率検定の P 値 < 0.01

- ・がんの最小アレレル頻度 0.005
- ・正常組織のアレレル頻度は 0.03 未満
- ・がんの変異リードの最小数は 5 本
- ・がんのアレレル頻度と正常のアレレル頻度の比が 10 以上
- ・TLOD(mutect が計算) が 10 以上

このアルゴリズムでがんのアレレル頻度 0.005(0.5%)以上の変異の同定を試みた結果、33.33%希釈で 7/8 個検出、11.11%希釈で 8/8 個検出、3.7%希釈でも 5/8 個検出した。また、偽陽性はなかった。かなり希釈したサンプルでも変異が検出されたことから、今回のプログラムは有用であると判断をした。

表 1 低頻度変異の検出結果

	Dilution (%)	Total	Mutation	FN	FP
HLADP135-1	100	129	8	0	0
HLADP135-2	33.33	96	7	1	0
HLADP135-3	11.11	140	8	0	0
HLADP135-4	3.7	106	5	3	0
HLADP135-5	1.23	128	0	8	0
HLADP135-6	0.41	124	0	8	0
HLADP135-7	0.14	105	0	8	0

(2) 肝がんを発症するまでの日数を予測するモデルの構築

全 672 症例の B 型肝炎患者の生体試料を研究協力施設から新たに収集し、肝がんの発症予測モデルに組み込む臨床検査項目の情報を収集した。すべての症例について HLA 遺伝子型を決定し、BBJ 試料において選択された 4 つの検査項目 (年齢、性別、肝硬変の有無、AFP) の情報を整理、Curation を実施した。ここで血液検査日、肝がん診断日、生体試料提供日を収集することにより、血液検査日から肝がんと診断されるまでの日数や中断日までの日数を得た。

4 つの検査項目に欠損のない 162 症例について、HLA-A*33:03、HLA-A*31:01、HLA-DPB1*02:01 を加えてコックス比例ハザードモデルを用いて肝がんの発症を予測するモデルを構築した。その結果、4 つの項目 (血液検査時年齢、性別、Log10AFP、HLA-A*33:03 の有無) が有意となり、下記の予測式が得られた。

$$h(t) = h_0(t) \cdot \exp[0.0567 \cdot \text{age} + 0.601 \cdot \text{sex} (\text{male}=1, \text{female}=0) + 0.642 \cdot \log_{10}\text{AFP} + 0.714 \cdot A^*33:03 (\text{carrier}=1, \text{non-carrier}=0)]$$

肝がん発症累積確率 (1 年後、3 年後) を用いた ROC 解析の結果、1 年後では AUC=0.862、3 年後では AUC=0.863 となることが分かった。また AUC、Sensitivity、Specificity を年ごとに計算した結果、この予測モデルは検査日から近い日数での予測精度が高いことが分かった。さらに $S = 0.0567 \cdot \text{age} + 0.601 \cdot \text{sex} (\text{male}=1, \text{female}=0) + 0.642 \cdot \log_{10}\text{AFP} + 0.714 \cdot A^*33:03 (\text{carrier}=1, \text{non-carrier}=0)$ とし、 $S > 4$ と $S \leq 4$ で肝がんの予測精度を評価した結果、 $S > 4$ の症例では 100 日以内の肝がんの発症リスクが有意に高くなることが分かった。構築した肝がん予測モデルによる予測性能を評価するために、162 症例をランダムに 2 : 1 に分けてトレーニングセットとテストセットとし、トレーニングセットで作成したモデル式からテストセットの病態を予測した。1000 回の繰り返し検定の結果、1 年後および 3 年後の C-index はともに 0.75 以上、Sensitivity は 0.7 以上となることが明らかとなった。臨床的には肝がんになるリスクの高い患者を高い精度で予測することが重要であることから、この予測モデルは臨床的にも利用価値の高いモデルであると言える。ヒト遺伝要因を項目として加えた肝がんリスク予測モデルはこれまでに報告がないことから、その点においても特長のある予測モデルと言える (Nishida et al. International Journal of Molecular Sciences 24: 4761, 2023)。

(3) Polygenic な効果を加えた肝がんのリスク予測モデルの構築

肝がん発症予測モデルは以下の手順で構築した。

BBJ の第 2 期症例を除く GWAS データを半数に分割することでモデル構築用データセットとモデル検証用データセットとし、ベストフィットモデルによる肝がん発症リスクの予測モデルを構築した。その結果、今回のモデルで肝がんの発症を説明できる割合は 1.35% となり、使用する SNP 数は 46,581 個となることが分かった。ゲノム解析データから線形混合モデルで推定した肝発がんの遺伝率は 25.2% であり、今回のモデルで説明できる割合が 1.35% であったことから、まだ多くの遺伝要因が未知のままであることが示唆される結果となった。そのため、年齢、性別、PRS 予測値を項目として加えた肝がんの発症リスク予測モデルによる肝がん患者の判別は $AUC=0.562$ と精度の低い結果となることが分かった。

(4) 肝臓組織由来 DNA のターゲットキャプチャーシーケンシング

国内の協力施設（倫理承認済：九州大学、東京大学、久留米大学）から提供された 182 症例の肝細胞がん組織（がん部、非がん部）から DNA を抽出し、ターゲットキャプチャーシーケンシングを実施した。がんゲノム解析パイプラインを用いて、低頻度変異の検出を現在実施中である。

5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 5件/うち国際共著 1件/うちオープンアクセス 5件）

1. 著者名 Ashouri S, Khor SS, Hitomi Y, Sawai H, Nishida N, Sugiyama M, Kawai Y, Posuwan N, Tangkijvanich P, Komolmit P, Tsuiji M, Shotelersuk V, Poovorawan Y, Mizokami M, Tokunaga K.	4. 巻 13
2. 論文標題 Genome-Wide Association Study for Chronic Hepatitis B Infection in the Thai Population	5. 発行年 2022年
3. 雑誌名 Front Genet	6. 最初と最後の頁 887121
掲載論文のDOI（デジタルオブジェクト識別子） 10.3389/fgene.2022.887121	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 Nishida N, Ohashi J, Suda G, Chiyoda T, Tamaki N, Tomiyama T, Ogasawara S, Sugiyama M, Kawai Y, Khor SS, Nagasaki M, Fujimoto A, Tsuchiura T, Ishikawa M, Matsuda K, Yano H, Yoshizumi T, Izumi N, Hasegawa K, Sakamoto N, Mizokami M, Tokunaga K.	4. 巻 24
2. 論文標題 Prediction Model with HLA-A*33:03 Reveals Number of Days to Develop Liver Cancer from Blood Test	5. 発行年 2023年
3. 雑誌名 Int J Mol Sci	6. 最初と最後の頁 4761
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/ijms24054761	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 -
1. 著者名 Watanabe Y, Ohashi J	4. 巻 26
2. 論文標題 Modern Japanese ancestry-derived variants reveal the formation process of the current Japanese regional gradations	5. 発行年 2023年
3. 雑誌名 iScience	6. 最初と最後の頁 106130
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.isci.2023.106130	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 -
1. 著者名 Nishida N, Sugiyama M, Ohashi J, Kawai Y, Khor SS, Nishina S, Yamasaki K, Yazaki H, Okudera K, Tamori A, Eguchi Y, Sakai A, Kakisaka K, Sawai H, Tsuchiura T, Ishikawa M, Hino K, Sumazaki R, Takikawa Y, Kanda T, Yokosuka O, Yatsuhashi H, Tokunaga K, Mizokami M.	4. 巻 11
2. 論文標題 Importance of HBsAg recognition by HLA molecules as revealed by responsiveness to different hepatitis B vaccines	5. 発行年 2021年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 3730, 3741
掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s41598-021-82986-8	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 -

1. 著者名 Kikuchi M, Kobayashi K, Nishida N, Sawai H, Sugiyama M, Mizokami M, Tokunaga K, Nakaya A.	4. 巻 8
2. 論文標題 Genome-wide copy number variation analysis of hepatitis B infection in a Japanese population	5. 発行年 2021年
3. 雑誌名 Human Genome Variation	6. 最初と最後の頁 22, 29
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41439-021-00154-w	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7件(うち招待講演 1件/うち国際学会 0件)

1. 発表者名 西田奈央, 杉山真也, 溝上雅史
2. 発表標題 ゲノム解析で明らかとなったHLA遺伝子と肝発がんの関連
3. 学会等名 第58回日本肝臓学会総会
4. 発表年 2022年

1. 発表者名 西田奈央, 土浦貴代, 石川美由紀, 徳永勝士, 溝上雅史
2. 発表標題 2種類のHBワクチンに異なる応答性を示すHLA遺伝子型
3. 学会等名 第30回日本組織適合性学会大会
4. 発表年 2022年

1. 発表者名 西田奈央
2. 発表標題 ゲノム解析が明らかにするB型肝炎由来肝がんのメカニズム
3. 学会等名 2022年度 大学共同利用機関法人 情報・システム研究機構 国立遺伝学研究所 研究会(招待講演)
4. 発表年 2023年

1. 発表者名 杉山真也, 西田奈央, 溝上雅史
2. 発表標題 HLA型ごとのHBV多様性解析によるB型肝炎の発症リスクの予測精度の向上
3. 学会等名 第58回日本肝臓学会総会
4. 発表年 2022年

1. 発表者名 大橋順
2. 発表標題 ポリジェニックスコアの基礎
3. 学会等名 大阪大学 ポリジェニックスコアを勉強する会
4. 発表年 2023年

1. 発表者名 西田奈央, 杉山真也, 土浦貴代, 石川美由紀, 徳永勝士, 溝上雅史
2. 発表標題 ヘパタバックス IIとピームゲンに対する応答性とHLA遺伝子の関連
3. 学会等名 日本人類遺伝学会第66回大会
4. 発表年 2021年

1. 発表者名 西田奈央, 杉山真也, 溝上雅史
2. 発表標題 ヘパタバックス-IIとピームゲンに対するワクチン応答性に関わる遺伝要因
3. 学会等名 第57回日本肝臓学会総会
4. 発表年 2021年

〔図書〕 計0件

〔出願〕 計1件

産業財産権の名称 B型肝炎ワクチン応答性に関わる遺伝因子	発明者 西田奈央、溝上雅史、徳永勝士	権利者 同左
産業財産権の種類、番号 特許、特願2021-28721	出願年 2021年	国内・外国の別 国内

〔取得〕 計0件

〔その他〕

-

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
研究分担者	大橋 順 (Ohashi Jun) (80301141)	東京大学・大学院理学系研究科(理学部)・教授 (12601)	
研究分担者	吉住 朋晴 (Yoshizumi Tomoharu) (80363373)	九州大学・医学研究院・教授 (17102)	

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
---------	---------