

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 6 日現在

機関番号：84409

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25430180

研究課題名(和文) 家族性及び同胞癌症例ゲノム解析による遺伝性癌原因遺伝子の同定とその応用

研究課題名(英文) Identification of causative genes for hereditary cancer by genome analyses of familial cancer cases and its application

研究代表者

久木田 洋児 (Kukita, Yoji)

地方独立行政法人大阪府立病院機構大阪府立成人病センター(研究所)・研究所・主任研究員

研究者番号：60372744

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：家族性が疑われる肺癌、胃癌、肉腫症例を収集し、エキソーム解析を行った。肺癌症例で検出したCHEK2変異は肺以外の臓器にも癌を発症した多発性肺癌症例の原因異常であると結論付けた。他の症例では原因となる異常は見つかっていない。一方、血液中に微量に存在する癌由来遺伝子変異をdeep sequencingで高感度に検出する方法を開発した。また、分子バーコード技術を用いて、高精度に塩基配列を決定し、DNA分子を絶対定量する方法を開発した。

研究成果の概要(英文)：We collected family members developed lung cancer, gastric cancer or sarcoma, and performed whole-exome sequencing on DNA from their blood. A CHEK2 mutation detected in lung cancer siblings was concluded to be the cause of multiple primary lung cancer with accompanying cancers in other organs. In other cases, mutations that caused cancers were not detected. We developed a high sensitive method to detect rare mutations in blood from cancer patients using deep sequencing. Further, by using molecular barcode technology, we developed a method to determine the nucleotide sequence with high accuracy, and to absolutely quantify DNA molecules.

研究分野：ゲノム科学

キーワード：疾患関連遺伝子 癌

1. 研究開始当初の背景

多くの癌症例ではいわゆる“癌関連遺伝子”に体細胞突然変異が起こっていることが分かっている。例えば、肺癌の中で最も患者数が多い肺腺癌の場合、6割以上で癌関連遺伝子に突然変異が起こっており、日本人患者集団に限ると約5割に上皮成長因子受容体EGFR遺伝子に体細胞突然変異がある。この突然変異を持つ患者では、EGFR阻害剤(ゲフィニチブ、エルロチニブ)による著しい腫瘍縮小効果が観察されるが、他の遺伝子変異を持つ患者には強い副作用が生じることもある。よって、遺伝子診断後にEGFR阻害剤を投与する治療は癌個別化医療の典型になっている。このような遺伝子異常(またはゲノム構造異常)による癌疾患の細分化が個別化医療実現への一方向である。あるカテゴリーに分類される患者数の割合が数%であっても、癌を含めたcommon diseaseにおける該当患者数は数百~数千人規模であり、患者の細分化は絶対的に必要である。

孤発癌症例組織中のゲノム構造異常の探索は、国際癌ゲノムコンソーシウムにより国ごとに解析対象癌種を割り当て、大規模なシーケンスにより進められていた。癌疾患の孤発症例を収集し、それらでの共通の異常を同定することが試みられていたが、解析領域の広さや集団の偏り(人種の違い等)に起因する偽陽性が多く、別コホートで再現されにくいという問題があった。

一方、古典的な方法であるが、ヒト疾患原因遺伝子探索には臨床医の所見や家系情報などにもとづき、解析対象を絞り込んだ研究(家族性疾患や顕著な形質を示す稀な症例の解析)がある。例えば、癌研究での網膜芽細胞腫におけるRB1や家族性大腸腺腫症のAPCの発見などは、連鎖解析により同定された原因遺伝子の異常が一般的な孤発症例のものと同通しており、癌の理解や診断法の開発に大きく貢献してきた。次世代(超並列)シーケンサーの登場で、ヒト全ゲノム配列解析のコスト面での問題が解決されつつあり、小規模家系の稀少疾患の全ゲノム/エキソーム解析を行うことによりそれらの原因遺伝子が特定されるようになっていた。

このように遺伝性疾患症例の全エキソーム解析を行う技術環境が整いつつあったが、症例の収集、ゲノムワイドなデータの処理、そこからヒトゲノム多様性に関する知識に基づいた結果の抽出が必要であり、特に遺伝性の癌疾患についてこれを実行でき且つ推進しつつある国内の研究グループは研究代表者ら以外になかった。

これまでに研究代表者は日本人集団のゲノム構造多様性研究に適したユニークなデータベース(dbQSNP: SNPアレル頻度データベース 2;D-HaploDB:ハプロタイプデータベース)を構築してきた。これらの情報は、全ゲノム/エキソーム解析時に、日本人集団特有のゲノム構造多型を考慮した疾患原因遺伝

子候補領域を絞り込むのに非常に有用である。そこでこれらを活用して全エキソーム解析により稀少な遺伝性癌症例の原因遺伝子を探索することを計画した。

研究代表者の所属機関は癌専門医療施設であるため、これまでに遺伝性が強く疑われる日本人稀少癌症例(多発性肺癌、肉腫など)を複数収集していた。経験豊富な臨床医師により遺伝様式と明確な臨床所見を基に絞り込まれた症例で、原因遺伝子異常が不明なことから、新規癌関連遺伝子上にgermline突然変異が存在することが強く示唆されていた。これら発癌原因遺伝子及びそれらが関わる反応経路は孤発癌症例とも共通することが予想されるので、原因究明は一般的な体細胞癌の理解、診断法の開発につながると考え、既に収集してあった遺伝性多発肺癌症例(同胞罹患例)については検体DNAを収集し、解析を始めていた。アレイCGH解析と全エキソーム解析を行った後、既知多型変異の除去や同胞間に発見したゲノム構造的特徴を考慮した解析(ホモ接合マッピング)により、DNA修復関連遺伝子CHEK2上に同遺伝子産物の安定性に強い影響を与える突然変異(ホモの非同義置換変異)を同定した。

本研究では、上記症例を引き続き解析するとともに、新しい症例の解析と新しい癌診断方法の開発研究を行った。

2. 研究の目的

様々な疾患に対する個別化医療の実現と普及には病因の特定とそれに続く発症・進行機序の理解や診断・治療法の開発が必須である。そこで本研究では、臨床医の所見により絞り込まれた癌症例検体(非常に稀な症状を呈する家族性及び同胞癌症例)の全エキソーム解析(全遺伝子配列解析)を行い、新規癌原因遺伝子の同定を起点とした、癌の個別化医療に役立つ新規癌バイオマーカーの開発、癌発症・進行機序の解明、創薬への応用を目指す。また、癌患者への身体的負担を抑えるために、非侵襲的に得られる患者検体(血液)から癌の診断ができる新しい方法を開発することも目的とする。

3. 研究の方法

(1) 超並列シーケンサーによる家族性癌症例の全エキソーム解析

所属施設病院及び共同研究施設の医師研究協力者とともに、特徴ある症状を呈する癌患者及びその近縁者から生体試料の研究利用に関する承諾を得て、血液検体の収集を行った。また、癌組織の手術摘出検体を得ることが可能な場合はそれも収集した。収集は期間を通して行った。

対象症例検体収集後、全エキソン抽出試薬(アジレント社 SureSelect)にて全エキソンを回収し、超並列シーケンサーにて配列決定を行った。エキソン回収、シーケンスライブラリー作成、塩基配列決定はシーケンス受託

会社にて行った。シーケンス解析関連ソフトウェア、bwa, samtools, GATK を使って得られた塩基配列データをヒトゲノム参照配列 (hg19)へマッピングし、配列変異検出を行った。変異部位のアノテーションには、RefSeq データベースの配列及び塩基位置情報を利用して独自に考案したアノテーション付加システムと ANNOVAR を使った。また、マッピングからアノテーションまでの一連の操作を半自動的に行う解析システムを構築した。

(2) 癌診断マーカー検出方法の開発

癌患者の血液中に微量に存在する癌細胞由来の遺伝子変異を検出する方法の開発を行った。

Deep sequencing による微量変異の検出
癌患者血液検体は所属施設病院の医師研究協力者ととも、生体試料の研究利用に関する承諾を得て収集された。血液は収集後数時間の内に血漿分離を行い、血漿部分を凍結保存した。解析する DNA は血漿から Qiagen 社 circulating nucleic acid kit を使って抽出した。EGFR 遺伝子のエキソン 19, 20, 21 内の体細胞変異ホットスポットを解析できるように PCR 反応系を設計し、その PCR 産物をデスクトップ型の超並列シーケンサー Ion Torrent 社 PGM を使って塩基配列を決定した。変異部位の同定にはアノマリ検出を使った。

DNA 塩基配列の高精度決定と DNA 分子の絶対定量を行う方法の開発

12 塩基から成る分子バーコードを組み込んだアダプターを制限酵素切断したゲノム DNA に付加し、各 DNA 分子を標識した。その後、標的領域特異的なプライマーとアダプター配列を用いて標的領域を増幅し、Ion Torrent 社 Proton を使って塩基配列を決定した。バーコード配列情報を基にシーケンス読み取りエラーの修正、分子数測定を行った。

4. 研究成果

(1) 超並列シーケンサーによる家族性癌症例の全エキソーム解析

肺癌症例の解析

本研究は、臨床的に特徴のある肺癌(多発性)の同胞罹患例のエキソーム解析から癌関連遺伝子 CHEK2 上に同遺伝子産物の安定性に強い影響を与える突然変異を同定したことを発端にして開始した。この同胞症例は肺癌以外にも過去に大腸癌と前立腺癌(男性)、子宮筋腫(女性)を発症していたので、CHEK2 以外にも疑わしい遺伝子異常が無いかどうかを調べた。また、解析可能な領域が拡大したエキソン抽出キットが利用できるようになったので、それを使い、上記検体を再度解析した。最初の解析と同様に、CHEK2 の変異が検出されたが、TP53, BRCA1/2, DNA ミスマッチ修復遺伝子群等良く知られた癌遺伝子上には common SNP (健康人集団で良く見られ

る多型)以外の変異は無かった。それで、CHEK2 の変異がこの症例の原因であると結論付け、学術雑誌に投稿していた(revise 中)。

胃癌患者症例の解析

2 家系のエキソーム解析を実施した。1 家系は若年発症患者を有し、患者の両親の両家系に胃癌患者が 1 人または複数人発症していた。この家系からは患者本人とその両親を、もう一方は家系内に複数のスキルス胃癌患者を有していたが、患者本人のみをエキソーム解析した。シーケンスリード数を使ったエキソン領域の染色体数異常解析では両家系に異常は検出されなかった。ヒト多型データベースに無いアミノ酸変化を伴う変異は各人約 300 個ずつ検出された。若年性胃癌の家系には常染色体劣性遺伝形式に従う変異が 6 遺伝子上に検出されたが、癌関連遺伝子として知られているものではなかった。スキルス胃癌患者にも癌と強い関連のある異常は見つからなかった。

肉腫患者症例の解析

家系内で複数人肉腫を有する症例を 2 家系収集した。そのうち 1 家系の母子症例のエキソームを行った。シーケンスリード数を使ったエキソン領域のコピー数解析では疾患の原因と疑われるような異常は検出されなかった。母子で共通かつ、ヒト多型データベースに無いアミノ酸変化を伴う変異は 301 個検出された。そのうちの 18 個は体細胞変異データベース COSMIC に登録されているものであったが、原因と疑われるような異常ではなかった。

胃癌と肉腫症例のエキソーム解析では原因を同定することが出来なかったので、タンパク質コード領域外の探索の為に全ゲノムシーケンス解析を予定していたが、施設内のシーケンサー用の専用試薬の発売が大幅に遅れており、実行できなかった。しかし、今後も上記症例の原因探索を続けていく予定である。

(2) 癌診断マーカー検出方法の開発

癌患者の血液中に微量に含まれる癌細胞由来の遺伝子変異を検出すれば、侵襲性の高い針生検における合併症のリスクを回避して、腫瘍における異常遺伝子の種類を知ることができる。また採血であれば、治療中に複数回行うことが出来るので、癌細胞由来の遺伝子変異の量を測定し、治療効果を詳細にモニタリングすることも可能である。そこで、血中循環癌細胞由来 DNA (ctDNA) を解析する方法の開発を行った。

Deep sequencing による EGFR 遺伝子の微量変異の検出

超並列シーケンサー、Ion Torrent 社 PGM を使って、肺腺癌患者の血漿中に存在する微量の EGFR 体細胞変異を検出する方法を開発した。具体的には、複数の健康人ゲノム DNA を用いて標的領域を 10 万以上の読み取り深度でシーケンスを行い、塩基位置ごとの読み取りエラー率情報を基に統計モデルを構築し、

エラー率を有意に超える頻度で現れる塩基変化を変異とするような検出限界値(アノマリ検出による判定基準)を設定した。この方法により正常アレルに対して0.01%の割合で含まれる微量変異DNAの検出(p 値 2×10^{-5})を可能にした。性能は血漿DNAを使った前向き(22検体)及び後ろ向き(155検体)試験により確認した。本法は統計学的有意差をもって微量変異DNAを検出する信頼性の高い技術である。

DNA塩基配列の高精度決定とDNA分子の絶対定量を行う方法の開発

Deep sequencingによる微量変異検出法を開発した一方、超並列シーケンサーの解析ではその高い読み取りエラー率による偽陽性が微量変異検出の障害になり、解析遺伝子領域を拡張しづらいという問題点があった。そこで研究代表者らは、超並列シーケンシングと分子バーコード技術を組み合わせ、標的DNA配列を高精度に決定する方法を開発した。この方法ではまず、標的領域近傍を制限酵素で切断したDNA断片分子に12塩基バーコード配列を付加する。12塩基の配列なら理論上 10^7 種類のバーコード配列が存在するので、数千分子のDNA断片を一分子ずつ識別して標識することが出来る。その後、PCR/シーケンシングにより得られた配列データをバーコード配列情報を基にグループ化し、コンセンサス配列を作る。それにより、実験中に挿入されたエラーが除かれるので、標的配列を高精度に決定することが出来る。さらに、検出されたバーコード配列とシーケンス配列数の関係を解析することにより、エラーを含んだバーコード配列を除去する方法を考案した。そして最終的に検出されたバーコード配列の種類を数えることにより、解析に使われたDNA分子の絶対定量ができる。この方法をNon-Overlapping Integrated Reads Sequencing System (NOIR-SeqS)と名付けた。ctDNA検出の実例として、肺癌患者の血中遊離DNAを使用し、KRAS遺伝子変異の検出にNOIR-SeqSを応用した。高精度な配列決定ができるが、読み取りエラーを完全に無くすることはできない。それで変異検出については、解析したDNA分子数と塩基配列情報を使い、エラーと真の変異を統計学的手法により区別する変異判定法を考案して用いた。肺癌患者では、EGFR遺伝子変異とKRAS遺伝子変異は相互排他的にみられることが知られているが、そのことと矛盾なく、以前の研究でEGFR遺伝子変異を持たないことがわかっていいる患者でのみアミノ酸を変えるKRAS変異が検出された(EGFR変異無患者8/19、変異有患者0/11)。また、肺癌患者では通常突然変異が検出されないCTNNB1遺伝子の解析では、アミノ酸を変える変異は1人でのみ検出された(1/30)。これらの結果から、NOIR-SeqSでは、事前に癌組織中の変異の情報を知ることなく、癌患者の血中から癌組織の体細胞変異を検出できることを確認した。今後は、高

精度塩基配列決定及び絶対定量の機能を保持したまま、解析領域を拡大する方向に技術を改良していく予定である。

5. 主な発表論文等

[雑誌論文](計9件)

Imamura F, Uchida J, Kukita Y, Kumagai T, Nishino K, Inoue T, Kimura M, Oba S, Kato K. Monitoring of treatment responses and clonal evolution of tumor cells by circulating tumor DNA of heterogeneous mutant EGFR genes in lung cancer. *Lung Cancer*, 2016, 94:68-73. (査読有)

DOI: 10.1016/j.lungcan.2016.01.023

Uchida J, Imamura F, Kukita Y, Kumagai T, Nishino K, Inoue T, Kimura M, Oba S, Kato K. Dynamics of circulating tumor DNA represented by the activating and resistant mutations in epidermal growth factor receptor-tyrosine kinase inhibitor treatment. *Cancer Science*, 2016, 107:353-358. (査読有)

DOI: 10.1111/cas.12860

Uchida J, Kato K, Kukita Y, Kumagai T, Nishino K, Daga H, Nagatomo I, Inoue T, Kimura M, Oba S, Ito Y, Takeda K, Imamura F. Diagnostic Accuracy of Noninvasive Genotyping of EGFR in Lung Cancer Patients by Deep Sequencing of Plasma Cell-Free DNA. *Clinical Chemistry*, 2015, 61:1191-1196. (査読有)

DOI: 10.1373/clinchem.2015.241414

Kukita Y, Matoba R, Uchida J, Hamakawa T, Doki Y, Imamura F, Kato K. High-fidelity target sequencing of individual molecules identified using barcode sequences: de novo detection and absolute quantitation of mutations in plasma cell-free DNA from cancer patients. *DNA Research*, 2015, 22:269-277. (査読有)

DOI: 10.1093/dnares/dsv010

Hamakawa T, Kukita Y, Kurokawa Y, Miyazaki Y, Takahashi T, Yamasaki M, Miyata H, Nakajima K, Taniguchi K, Takiguchi S, Mori M, Doki Y, Kato K. Monitoring gastric cancer progression with circulating tumour DNA. *British Journal of Cancer*, 2015, 112:352-356. (査読有)

DOI: 10.1038/bjc.2014.609

Kimura M, Endo H, Inoue T, Nishino K, Uchida J, Kumagai T, Kukita Y, Kato K, Imamura F, Inoue M. Analysis of ERBB ligand-induced resistance mechanism to crizotinib by primary culture of

lung adenocarcinoma with EML4-ALK fusion gene. Journal of Thoracic Oncology, 2015, 10:527-530. (査読有)
DOI: 10.1097/JTO.0000000000000381
Kanemoto M, Shirahata M, Nakauma A, Nakanishi K, Taniguchi K, Kukita Y, Arakawa Y, Miyamoto S, Kato K. Prognostic prediction of glioblastoma by quantitative assessment of the methylation status of the entire MGMT promoter region. BMC Cancer, 2014, 14:641. (査読有)
DOI: 10.1186/1471-2407-14-641
Tahira T, Yahara K, Kukita Y, Higasa K, Kato K, Wake N, Hayashi K. A definitive haplotype map of structural variations determined by microarray analysis of duplicated haploid genomes. Genomics Data, 2014, 2:55-59. (査読有)
DOI: 10.1016/j.gdata.2014.04.006
Kukita Y, Uchida J, Oba S, Nishino K, Kumagai T, Taniguchi K, Okuyama T, Imamura F, Kato K. Quantitative identification of mutant alleles derived from lung cancer in plasma cell-free DNA via anomaly detection using deep sequencing data. PLoS One, 2013, 8:e81468. (査読有)
DOI: 10.1371/journal.pone.0081468

〔学会発表〕(計8件)

Kukita Y, Matoba R, Uchida J, Takada R, Hamakawa T, Doki Y, Imamura F, Katayama K, Kato K. Non-overlapping integrated reads sequencing system (NOIR SeqS) for de novo detection and absolute quantitation of mutations in circulating tumor DNA. Advances in Genome Biology and Technology 2016. 2016, Orlando, FL
久木田洋児, 的場亮, 内田純二, 浜川卓也, 土岐祐一郎, 今村文生, 加藤菊也. Detection of mutations in plasma DNA from cancer patients by high-fidelity target sequencing of individual molecules (DNA分子の高精度配列決定による癌患者由来血漿DNA中の突然変異検出). 第74回日本癌学会学術総会, 2015, 名古屋
久木田洋児, 内田純二, 的場亮, 今村文生, 加藤菊也. 次世代シーケンサーを使った癌患者血漿DNA中突然変異検出システムの開発. NGS現場の会 第四回研究会, 2015, 筑波
久木田洋児. DNAバーコード技術を使った高精度DNAシーケンシング. 第20回JBICバイオ関連基盤技術研究会, 2015, 東京
久木田洋児. 次世代シーケンサーによ

る肺癌患者血中DNA変異の定量的検出. NGS現場の会 血液核酸マーカー分科会 第1回研究会, 2014, 東京
久木田洋児, 的場亮, 加藤菊也. High-fidelity DNA sequencing of single molecules monitoring errors introduced into barcode sequences (バーコード配列内読み取りエラー検出による1DNA分子高精度シーケンシング). 第37回日本分子生物学会年会, 2014, 横浜
久木田洋児, 的場亮, 加藤菊也. A method to trim erroneous barcode-tags for molecular counting of target DNA with massively parallel sequencing(超並列シーケンシングによる標的DNA配列の分子数測定). 第73回日本癌学会学術総会, 2014, 横浜
久木田洋児, 内田純二, 大羽成征, 西野和美, 熊谷融, 谷口一也, 奥山貴子, 二木俊江, 今村文生, 加藤菊也. Quantitative detection of mutations in plasma DNA of cancer patients by use of anomaly detection with deep sequencing(超並列シーケンシングによる癌患者血漿DNA中遺伝子突然変異の検出), 第72回日本癌学会学術総会, 2013, 横浜

〔図書〕(計1件)

久木田洋児, 加藤菊也. 血中腫瘍DNA遺伝子医学MOOK オミックスで加速するがんバイオマーカー研究の最新動向(監修:今井浩三, 編集:山田哲司, 金井弥栄). 2015, メディカルドゥ

〔産業財産権〕

出願状況(計1件)

名称: 核酸分子数計測法
発明者: 加藤菊也、久木田洋児、的場亮
権利者: 同上
種類: 特許
番号: PCT/JP2015/069114
出願年月日: 平成27年7月2日
国内外の別: 国際

6. 研究組織

(1) 研究代表者

久木田 洋児 (KUKITA, Yoji)
大阪府立成人病センター・研究所・主任研究員
研究者番号: 60372744

(2) 研究分担者

加藤 菊也 (KATO, Kikuya)
大阪府立成人病センター・研究所・部門長
研究者番号: 60194809