

令和 3 年 6 月 16 日現在

機関番号：82606

研究種目：基盤研究(C)（一般）

研究期間：2018～2020

課題番号：18K11542

研究課題名（和文）機械学習・深層学習に基づいた多層化統合エピジェネティクスデータベースの構築

研究課題名（英文）Construction of a multilayered integrated epigenetics database based on machine learning and deep learning

研究代表者

金子 修三（Kaneko, Syuzo）

国立研究開発法人国立がん研究センター・研究所・ユニット長

研究者番号：10777006

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究課題においては、個体レベルでの“がん”エピジェネティクスの本態解明を目指した。具体的にはシングルセルレベルでのクロマチン構造の解明を目的として、深層学習を用いた細胞画像データのクラス分類を実施した。さらに凍結検体、FFPE検体等、様々な形態のサンプルにおいても、ロボティクス技術を駆使したChIP-seq解析が適用可能であることを示し、得られたNGSデータセットを解析した。その結果として査読付き英語論文を本研究期間内に複数報告した。また、産総研・人工知能研究センターと共同で、転写因子結合部位の同定も可能とするFFPE ChIP-seq技術をヒト型汎用ロボットに実装させた。

研究成果の学術的意義や社会的意義

本研究課題で得られた、FFPE検体を用いたChIP-seq解析手法は、ロボティクス技術を駆使したことによって大規模解析が可能となり、臨床学的に意義のあるデータを得ることが出来た。具体的には、FFPE検体に紐付けられている、予後・転移・薬剤感受性等の臨床情報と組み合わせることで、今後の創薬展開の道を拓いたと判断している。

研究成果の概要（英文）：In this research project, we aimed to elucidate the true nature of cancer epigenetics at the individual level. Specifically, we used deep learning to classify cell image data in order to elucidate chromatin structure at the single cell level. In addition, we demonstrated that ChIP-seq analysis using robotics technology can be applied to various types of samples such as frozen samples and FFPE samples, and analyzed the obtained NGS data sets. As a result, we reported several peer-reviewed papers in English within this research period. In collaboration with the Artificial Intelligence Research Center of AIST, we implemented the FFPE ChIP-seq technology, which can also identify transcription factor binding sites, on a humanoid robot.

研究分野：情報科学、ロボティクス

キーワード：データマイニング ロボティクス エピジェネティクス

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

エピジェネティクスという学問領域が取り扱う分子群は、DNA、タンパク質、RNA と多様であり、個々の機能を詳細に解析する既存の手法だけでは、生体内における本質的な役割を解明するには不十分である。すなわち、統合的に解析を行うための新規アルゴリズムの開発を行うことが、エピジェネティクス学問領域の包括的理解を深めるためには不可欠であると考えられていた。また次世代シーケンス解析を中心とした技術革新に伴って、学術的知見に関しては積み重なっていたものの、対象となるサンプルがモデル生物や細胞株に限定されている場合がほとんどであった。

“がん”という病態の本態解明において、エピジェネティクスの重要性は認識されていたが、その中でも特に、長鎖ノンコーディング RNA やヒストンメチル化が重要な役割を担っていることが明らかになっていった。また、ENCODE や IHEC など高速シーケンス(NGS)解析を用いた大規模データベースの構築が開始され、当該分野の重要性は世界的に認知されるようになっていったものの、依然未解明の部分が多く、現在の解析技術の限界が指摘されていた。さらに腫瘍内不均一性(Intratatumoral Heterogeneity)を伴う遺伝的・環境的要因の多様性から、取得したデータ解析も困難を生ずるケースがあった。

具体的な例として、ヒストン修飾やクロマチン高次構造に関しては ChIP-seq 法や Hi-C 法等の技術革新が進むにつれ、ゲノムワイドに解析することが可能ではあったが、臨床的価値の高い ChIP-seq・Hi-C 解析は世界的に見ても報告例は限られていた。即ち医学研究分野におけるエピジェネティクス機構の統一的な解析手法が確立されていなかったことを背景に、近年急速に進歩しているロボティクス技術や、人工知能の中核的アルゴリズムである機械学習・深層学習を用いたエピジェネティクス解析の可能性が期待されていた。

2. 研究の目的

臨床的応用も可能とするエピジェネティクス解析手法を開発することを目的として、希少検体を用いた新規 ChIP-seq 法の開発を試みた。さらに画像データと親和性の高い深層学習を活かしながら、シングルセルレベルでのクロマチン構造を解明して、NGS 解析と統合しながら個体レベルでの“がん”エピジェネティクスの本態解明を目指した。

3. 研究の方法

シングルセルレベルでのクロマチン構造の解明を目的として、深層学習(Deep learning)を用いた細胞画像データのクラス分類を行った。解析対象としては、ホルマリン固定パラフィン包埋の肝がん検体を利用し、DAPI で染色した細胞核を正確にセグメンテーションできるかどうかを検討した。本研究課題では医用画像や組織染色画像を対象とした場合に用いられる Unet を参考にした end to end アルゴリズムを構築し、深層学習に基づいた細胞核セグメンテーションを実施した。我々の典型的な Unet では、3つのマックスプーリング層と3つのアップサンプリング層を持つ。パラメータ数の制限を目的として、コンボリューションフィルタでは 3×3 のサイズのフィルタを用いた。活性化関数として ReLU を採用し、更にドロップアウト層を追加した。また事前学習した重みを初期値としてエンコーダの箇所に追加した。本手法はパラメータ収束を著しく向上させることが既に報告されており、我々はオプションとして、VGG19 もしくは resnet152 モデルを重みの初期値として加えた。このモデルを用いてセマンティックセグメンテーションを実施して、3つの異なったクラス(背景、核、境界領域)に分割することを試みた。



図1| バイオメディカル
研究支援 LabDroid

次に臨床的応用も可能とするエピジェネティクス解析手法の開発を目的として、FFPE 組織検体や凍結組織検体等、様々なサンプル形態に対応した ChIP-seq 解析手法の開発を試みた。具体的には、国立研究開発法人・産業技術総合研究所(産総研)・人工知能研究センターに設置されているヒト型汎用ロボット LabDroid「まほろ」(図1)に、我々が独自で開発した FFPE ChIP-seq 解析手法を実装させて、非常に煩雑な実験手法である ChIP-seq 解析の Wet 実験の自動化を試みた。さらに LabDroid「まほろ」の開発元であるロボティック・バ

イオロジー・インスティテュート株式会社 (RBI 社) と共同で、凍結組織検体を用いた ChIP-seq 自動化の技術支援を行い、大規模 ChIP-seq 解析の検証実験を試みた。

4. 研究成果

深層学習(Deep learning)を用いた細胞画像データのクラス分類については、2018 年の Data Science Bowl(<https://www.kaggle.com/c/data-science-bowl-2018/discussion/56316>)にて使用された 500 以上の細胞画像 (source domain) および我々が取得した細胞画像 (target domain) を用いた結果、エポック数 200 程度で、核検出においては良好な結果を得た。しかしながら Instance Segmentation においては、target domain において、いくつかの誤りを観測した。以上の結果をまとめ、第 1 回日本メディカル AI 学会学術集会にて報告した結果、学会奨励賞の優秀賞を頂いた。

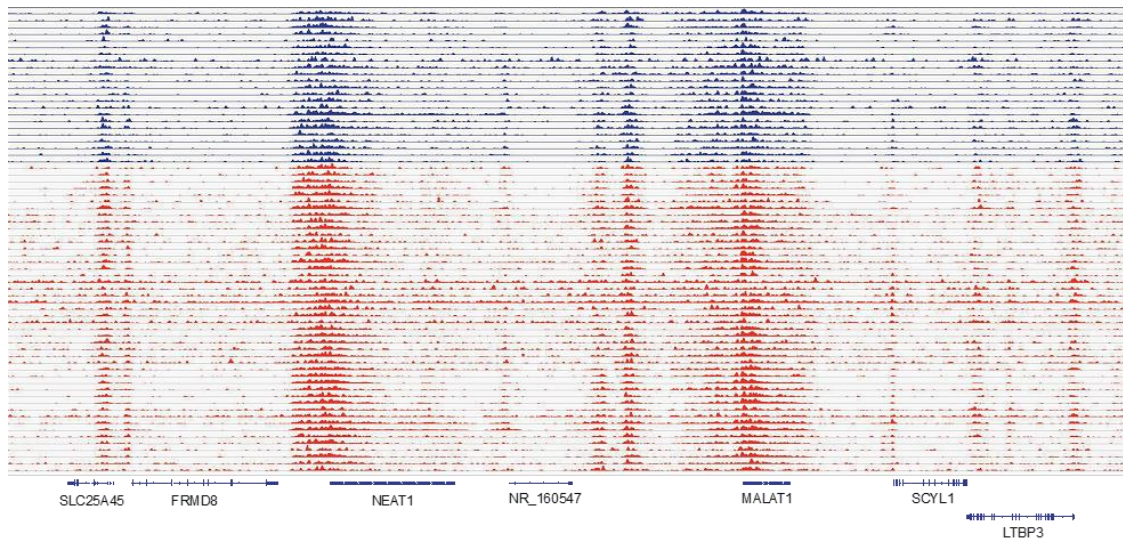


図 2 | LabDroid「まほろ」を用いた FFPE ChIP-seq

肺癌 FFPE 組織検体 69 検体から得られた H3K27ac ChIP-seq のデータを IGV で可視化した例を示す。常に活性化された遺伝子である NEAT1, MALAT1 遺伝子領域に H3K27ac ヒストン修飾の蓄積が見られており、FFPE ChIP-seq 実験結果が良好であったことを示している。

次に臨床的応用も可能とするエピジェネティクス解析手法の開発を本研究課題の期間内に実施した結果、培養細胞、凍結検体、FFPE 検体等、様々な形態のサンプルにおいても、ロボティクス技術を駆使した ChIP-seq 解析が適用可能であることを示し (図 2)、得られた NGS データセットを解析した。その結果として査読付き英語論文を本研究期間内に複数報告した。また知財および 産学連携の成果としては、産総研・人工知能研究センターと共同で、転写因子結合部位

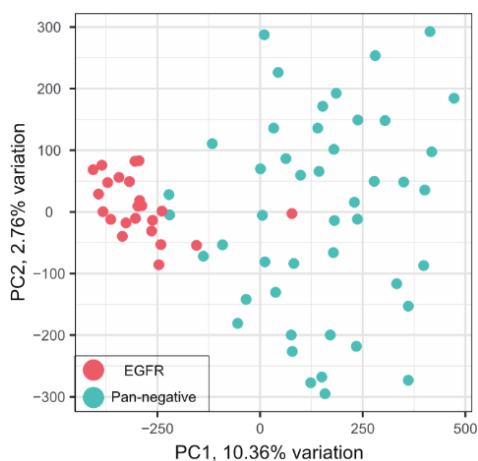


図 3 | ヒストン修飾 (H3K27ac) 状況から導き出されるドライバー変異

EGFR ドライバー変異を伴う肺腺がん 23 症例と Pan-negative 46 症例を用いた H3K27ac ChIP-seq 解析を実施して得られたデータセットを主成分分析(Principal Component Analysis; PCA)を用いた次元削減処理後、特徴抽出を行い、2次元上にプロットした。PCA は Singular Value Decomposition に基づいて実施した。

の同定も可能とする FFPE ChIP-seq 技術をヒト型汎用ロボットに実装させた結果、著作権法の保護対象となる職務発明として認定された (整理番号: 2020-26, 発明の名称: 「LabDroid MAHOLO 用 Clinical ChIP-seq プロトコル」)。以上の実験から得られたデータセットを解析したところ、アジア人に高頻度で見られる EGFR ドライバー変異を伴う肺腺がん症例で、エピジェネティクスの観点から均一性を保っていることが判明しており、さらにその特徴が Pan-negative 症例とは明確に分類出来ることを見出した (図 3)。この結果は、EGFR ドライバー変異を伴う肺腺がん患者を対象とした場合の、分子標的薬(EGFR-TKI)の高い奏効率(～

80%) をエピジェネティクスの観点からも説明可能であることを示していた。

5. 主な発表論文等

〔雑誌論文〕 計14件（うち査読付論文 14件／うち国際共著 3件／うちオープンアクセス 13件）

1. 著者名 Kaneko S, Mitsuyama T, Shiraishi K, Ikawa N, Shozu K, Dozen A, Machino H, Asada K, Komatsu M, Kukita A, Sone K, Yoshida H, Motoi N, Hayami S, Yoneoka Y, Kato T, Kohno T, Natsume T, Keudell GV, Saloura V, Yamaue H, Hamamoto R	4. 巻 13
2. 論文標題 Genome-Wide Chromatin Analysis of FFPE Tissues Using a Dual-Arm Robot with Clinical Potential	5. 発行年 2021年
3. 雑誌名 Cancers	6. 最初と最後の頁 2126 ~ 2126
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/cancers13092126	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 Ozawa T, Kaneko S, Szulzewsky F, Qiao Z, Takadera M, Narita Y, Kondo T, Holland EC, Hamamoto R, Ichimura K	4. 巻 9
2. 論文標題 C11orf95-RELA fusion drives aberrant gene expression through the unique epigenetic regulation for ependymoma formation	5. 発行年 2021年
3. 雑誌名 Acta Neuropathologica Communications	6. 最初と最後の頁 36
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s40478-021-01135-4	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 Wada M, Kukita A, Sone K, Hamamoto R, Kaneko S, Komatsu M, Takahashi Y, Inoue F, Kojima M, Honjoh H, Taguchi A, Kashiyama T, Miyamoto Y, Tanikawa M, Tsuruga T, Mori-Uchino M, Wada-Hiraike O, Osuga Y, Fujii T	4. 巻 10
2. 論文標題 Epigenetic Modifier SETD8 as a Therapeutic Target for High-Grade Serous Ovarian Cancer	5. 発行年 2020年
3. 雑誌名 Biomolecules	6. 最初と最後の頁 1686 ~ 1686
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/biom10121686	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 Hamamoto R, Suvarna K, Yamada M, Kobayashi K, Shinkai N, Miyake M, Takahashi M, Jinnai S, Shimoyama R, Sakai A, Takasawa K, Bolatkan A, Shozu K, Dozen A, Machino H, Takahashi S, Asada K, Komatsu M, Sese J, Kaneko S	4. 巻 12
2. 論文標題 Application of Artificial Intelligence Technology in Oncology: Towards the Establishment of Precision Medicine	5. 発行年 2020年
3. 雑誌名 Cancers	6. 最初と最後の頁 3532 ~ 3532
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/cancers12123532	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 Papadaki V, Asada K, Watson JK., Tamura T, Leung A, Hopkins J, Dellett M, Sasai N, Davaapil H, Nik-Zainal S, Longbottom R, Nakakido M, Torii R, Veerakumarasivam A, Kaneko S, Sagoo MS, Murphy G, Mitani A, Tsumoto K, Kelly JD, Hamamoto R, Ohnuma S	4. 巻 12
2. 論文標題 Two Secreted Proteoglycans, Activators of Urothelial Cell?Cell Adhesion, Negatively Contribute to Bladder Cancer Initiation and Progression	5. 発行年 2020年
3. 雑誌名 Cancers	6. 最初と最後の頁 3362 ~ 3362
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/cancers12113362	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 Takahashi S, Asada K, Takasawa K, Shimoyama R, Sakai A, Bolatkan A, Shinkai N, Kobayashi K, Komatsu M, Kaneko S, Sese J, Hamamoto R	4. 巻 10
2. 論文標題 Predicting Deep Learning Based Multi-Omics Parallel Integration Survival Subtypes in Lung Cancer Using Reverse Phase Protein Array Data	5. 発行年 2020年
3. 雑誌名 Biomolecules	6. 最初と最後の頁 1460 ~ 1460
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/biom10101460	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 Kageyama S, Du J, Kaneko S, Hamamoto R, Yamaguchi S, Yamashita R, Okumura M, Motegi A, Hojo H, Nakamura M, Tsuchihara K, Akimoto T	4. 巻 155
2. 論文標題 Identification of the mutation signature of the cancer genome caused by irradiation	5. 発行年 2021年
3. 雑誌名 Radiotherapy and Oncology	6. 最初と最後の頁 10 ~ 16
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.radonc.2020.10.020	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 Kojima M, Sone K, Oda K, Hamamoto R, Kaneko S, Oki S, Kukita A, Kawata A, Honjoh H, Kawata Y, Kashiya T, Sato M, Taguchi A, Miyamoto Y, Tanikawa M, Tsuruga T, Nagasaka K, Wada-Hiraike O, Osuga Y, Fujii T	4. 巻 20
2. 論文標題 The histone methyltransferase SMYD2 is a novel therapeutic target for the induction of apoptosis in ovarian clear cell carcinoma cells	5. 発行年 2020年
3. 雑誌名 Oncology Letters	6. 最初と最後の頁 1 ~ 1
掲載論文のDOI (デジタルオブジェクト識別子) 10.3892/ol.2020.12014	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 Matsuno Yusuke, Atsumi Yuko, Shimizu Atsuhiko, Katayama Kotoe, Fujimori Haruka, Hyodo Mai, Minakawa Yusuke, Nakatsu Yoshimichi, Kaneko Syuzo, Hamamoto Ryuji, Shimamura Teppei, Miyano Satoru, Tsuzuki Teruhisa, Hanaoka Fumio, Yoshioka Ken-ichi	4. 巻 10
2. 論文標題 Replication stress triggers microsatellite destabilization and hypermutation leading to clonal expansion in vitro	5. 発行年 2019年
3. 雑誌名 Nature Communications	6. 最初と最後の頁 3925 ~ 3925
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41467-019-11760-2	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 Kojima Machiko, Sone Kenbun, Oda Katsutoshi, Hamamoto Ryuji, Kaneko Syuzo, et. al.,	4. 巻 19
2. 論文標題 The histone methyltransferase WHSC1 is regulated by EZH2 and is important for ovarian clear cell carcinoma cell proliferation	5. 発行年 2019年
3. 雑誌名 BMC Cancer	6. 最初と最後の頁 455 ~ 455
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s12885-019-5638-9	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 Hamamoto Ryuji, Komatsu Masaaki, Takasawa Ken, Asada Ken, Kaneko Syuzo	4. 巻 10
2. 論文標題 Epigenetics Analysis and Integrated Analysis of Multiomics Data, Including Epigenetic Data, Using Artificial Intelligence in the Era of Precision Medicine	5. 発行年 2019年
3. 雑誌名 Biomolecules	6. 最初と最後の頁 62 ~ 62
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/biom10010062	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 Kim Sangchul, Bolatkan Amina, Kaneko Syuzo, Ikawa Noriko, Asada Ken, Komatsu Masaaki, Hayami Shinya, Ojima Hidenori, Abe Nobutsugu, Yamaue Hiroki, Hamamoto Ryuji	4. 巻 9
2. 論文標題 Deregulation of the Histone Lysine-Specific Demethylase 1 Is Involved in Human Hepatocellular Carcinoma	5. 発行年 2019年
3. 雑誌名 Biomolecules	6. 最初と最後の頁 810 ~ 810
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/biom9120810	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 Kukita Asako, Sone Kenbun, Oda Katsutoshi, Hamamoto Ryuji, Kaneko Syuzo, et. al.,	4. 巻 513
2. 論文標題 Histone methyltransferase SMYD2 selective inhibitor LLY-507 in combination with poly ADP ribose polymerase inhibitor has therapeutic potential against high-grade serous ovarian carcinomas	5. 発行年 2019年
3. 雑誌名 Biochemical and Biophysical Research Communications	6. 最初と最後の頁 340 ~ 346
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.bbrc.2019.03.155	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Lee Chul-Hwan, Yu Jia-Ray, Kumar Sunil, Jin Ying, LeRoy Gary, Bhanu Natarajan, Kaneko Syuzo, Garcia Benjamin A., Hamilton Andrew D., Reinberg Danny	4. 巻 70
2. 論文標題 Allosteric Activation Dictates PRC2 Activity Independent of Its Recruitment to Chromatin	5. 発行年 2018年
3. 雑誌名 Molecular Cell	6. 最初と最後の頁 422 ~ 434. e6
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.molcel.2018.03.020	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5件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 金子修三
2. 発表標題 ロボティクス技術を駆使した大規模クリニカルChIP-seq解析の確立
3. 学会等名 第79回日本癌学会学術総会
4. 発表年 2020年

1. 発表者名 金子 修三
2. 発表標題 ChIP-seq解析に基づいた“がん”創薬ターゲット探索を加速させるバイオ実験自動化システムの構築
3. 学会等名 第78回日本癌学会学術総会
4. 発表年 2019年

1. 発表者名 金子 修三
2. 発表標題 Toward the application of precision medicine: Multi-omics analysis including clinical ChIP-seq datasets may reveal pivotal regulation of gene expression in cancer
3. 学会等名 第2回日本メディカルAI学会学術集会
4. 発表年 2019年

1. 発表者名 Syuzo Kaneko and Ryuji Hamamoto
2. 発表標題 臨床応用を志向したChIP-seq技術に基づいたがんエピゲノム解析の開発
3. 学会等名 The 1st Annual Meeting of the Japanese Association for Medical Artificial Intelligence
4. 発表年 2018年

1. 発表者名 Syuzo Kaneko and Ryuji Hamamoto
2. 発表標題 Toward precision medicine: developing new toolbox to study the epigenetics applied for cancer treatment
3. 学会等名 The 77th Annual Meeting of the Japanese Cancer Association
4. 発表年 2018年

〔図書〕 計1件

1. 著者名 金子修三	4. 発行年 2020年
2. 出版社 医歯薬出版株式会社	5. 総ページ数 5
3. 書名 AIが切り拓く未来の医療	

〔産業財産権〕

〔その他〕

2019年研究論文刊行成果
https://www.ncc.go.jp/jp/ri/division/molecular_modification_and_cancer_biology/paper/010/20190409184120.html
 がん分子修飾制御学分野 研究室の紹介
https://www.ncc.go.jp/jp/ri/division/molecular_modification_and_cancer_biology/index.html
 2018年研究論文刊行成果
https://www.ncc.go.jp/jp/ri/division/molecular_modification_and_cancer_biology/paper/020/20190409175902.html
 がん分子修飾制御学分野 研究室の紹介
https://www.ncc.go.jp/jp/ri/division/molecular_modification_and_cancer_biology/index.html

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	浜本 隆二 (Hamamoto Ryuji) (80321800)	国立研究開発法人国立がん研究センター・研究所・分野長 (82606)	

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