理化学研究所　東京大学

PRESS RELEASE

理化学研究所
東京大学

がん全ゲノムにおけるマイクロサテライト不安定性の解明
ーがん免疫療法の適応診断や遺伝性腫瘍診断に有効ー

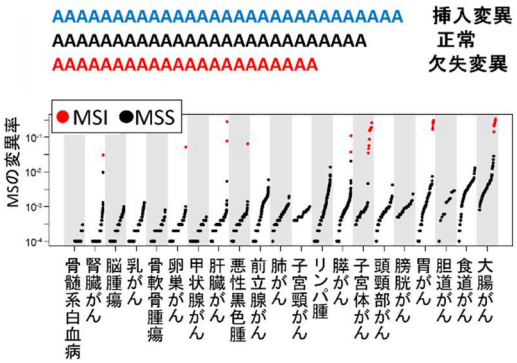
理化学研究所（理研）生命医科学研究センターがんゲノム研究チームの藤本明洋副チームリーダー（研究当時、現 東京大学大学院医学系研究科教授）、中川英刀チームリーダーらの共同研究グループは、2,700 例以上のがん全ゲノムシーケンスデータについて、「マイクロサテライト（MS）[1]」と呼ばれる繰り返し配列の変異を同定し、その特徴を明らかにしました。本研究は、国際連携による全ゲノムがん種横断的解析プロジェクト（PCAWG）[2]の一環として行われ、がん全ゲノムの「マイクロサテライト不安定性（MSI）」[3]に関するこれまでで最も網羅的な研究と診断となりました。

本研究成果は、がん免疫療法の適応診断、遺伝性腫瘍の診断に有効であり、次世代のがんゲノム医療研究の進展に貢献すると期待できます。

高変異率の MS マーカーを用いた MSI 判定は、現在、遺伝性腫瘍[4]の診断やがん免疫療法の適応診断のために行われています。しかし、繰り返し配列の解析は非常に難しく、全ゲノムでどのような MS 変異が起きているかは不明でした。

今回、共同研究グループは、全ゲノムにおける MS 変異同定のための高精度解析法「MIMcall」を開発し、PCAWG で解析された 2,717 例の 21 種類のがんの全ゲノムシーケンスデータについて、1 例あたり平均約 765 万カ所、合計約 200 億カ所の MS 変異解析を行いました。その結果、ほとんどの MS に変異が見つかりませんでしたが、解析できた MS 全体の約 2.6%に当たる約 20 万カ所の MS において複数個のがんゲノムに変異があることが分かりました。31 例のがん（大腸がん、子宮体がん、胃がんなど）に MS 変異が極端に多く見つかり、MSI と判定できました。さらに、高頻度に変異している 20 個の MS マーカーを新たに同定し、既存の MS マーカーと同程度の変異頻度を持ち、MSI 判定の精度も同等であることが分かりました。

本研究は、科学雑誌『*Genome Research*』の掲載に先立ち、オンライン版（3 月 24 日付：日本時間 3 月 25 日）に掲載されました。



21 のがん種ごとの全ゲノムにおけるマイクロサテライト（MS）の変異率

※共同研究グループ
理化学研究所　生命医科学研究センター　がんゲノム研究チーム
　チームリーダー　中川　英刀　（なかがわ　ひでわき）
　上級研究員　藤田　征志　（ふじた　まさし）
東京大学　大学院医学系研究科 国際保健学専攻 人類遺伝学分野
　教授　藤本　明洋　（ふじもと　あきひろ）
東京大学　医科学研究所
　ヒトゲノム解析センター　DNA 情報解析分野
　教授　宮野　悟　（みやの　さとる）
　ヘルスインテリジェンスセンター　健康医療データサイエンス分野
　教授　井元　清哉　（いもと　せいや）
埼玉県立がんセンター　腫瘍診断・予防科
　部長　赤木　究　（あかぎ　きわむ）

1．背景

がんは、ゲノムに変異が蓄積することで発症・進行する「ゲノムの病気」であり、がんの本質は「ゲノムの不安定性」にあるといえます。現在、世界中でがんの網羅的な全ゲノムシーケンス解析[5]や、その結果得られるがんのゲノム変異に基づく薬の開発、個別化医療（がんゲノム医療）が精力的に行われています。2014～2019 年には、がんゲノム分野の国際共同プロジェクト「全ゲノムがん種横断的解析（PCAWG）」により、がんの大規模な全ゲノムシーケンス解析が行われた結果、最も網羅的かつ詳細ながんゲノムカタログ[6]が発表されました[注 1)]。

ヒトゲノムの 30 億塩基配列のうちの約半分を繰り返し配列が占めています。そのうち、1～6 個の塩基が 10～数十回繰り返した配列を「マイクロサテライト（MS）」と呼び、ヒトゲノム中には 1000 万個以上の MS があると考えられています。MS は突然変異率が高く、個体間における繰り返しの長さの遺伝的多型[7]が豊富なことから、遺伝的マーカーとして集団遺伝学や家系解析などの研究に有効です。

がんの診療では、数個の MS マーカーの変異（繰り返しの長さが変化する）の有無によって「マイクロサテライト不安定性（MSI）」の判定が行われており、DNA 修復機構が欠損する遺伝性腫瘍（リンチ症候群[4]）の診断に利用されています。さらに近年、MSI 陽性がんが免疫チェックポイント阻害剤[8]による治療によく反応することが分かってきたため、MSI 判定は免疫チェックポイント阻害剤の適応基準にも使われるようになり、注目を集めています。しかし、これまで MSI 判定は、高変異率を示してきた数個の MS マーカーを用いられていますが、ゲノム全体でどのような MS 変異が起きているのかは不明でした。また、MS の高変異率をもたらす機序についても多くは解明されていませんでした。

近年の次世代シーケンサー（NGS）[9]技術の急速な進展、情報解析技術や IT ハード面の技術革新に伴い、約 30 億塩基の情報からなるヒトの全ゲノムシーケンス解析が、容易かつ安価に行えるようになりました。今後は、全ゲノムシーケンス解析が研究のみならず、がんゲノム医療といった診断や個別化医療においても、重要な解析手法になると予測されています。しかし、現在の NGS の解析技

術では、MS や繰り返し配列部位の解析は極めて難しく、NGS による全ゲノムでの大規模な MS 解析は行われてきませんでした。

注 1）2020 年 2 月 6 日プレスリリース「国際連携によるがん全ゲノムの大規模解析」
https://www.riken.jp/press/2020/20200206_1/

2．研究手法と成果

繰り返し配列である MS には、ポリメラーゼ連鎖反応（PCR）[10]および NGS でのシーケンス解析の段階においてエラー[11]が入りやすくなります。共同研究グループは、この測定でのエラーの頻度やパターンを解明するため、男性の正常 DNA の全ゲノムシーケンスデータの中で X 染色体[12]にある MS に着目しました。男性の X 染色体は 1 本しかなく、他の染色体に位置する MS のように二つのアレル[13]を評価する必要がないからです。そこで、男性 X 染色体のデータから全ゲノムシーケンスでの MS の人為的変異の割合を推定し、この情報を用いて、がんでの MS 変異を同定するための解析法「MIMcall」を開発し、この手法が高い精度を持つことを確認しました。

次に MIMcall を用いて、PCAWG にて解析された 2,717 例に及ぶ 21 種類のがん全ゲノムシーケンスデータ（がんと正常）から、1 例あたり平均約 765 万カ所、合計約 200 億カ所の MS を対象として変異解析を行いました。その結果、ほとんどの MS に変異は見つかりませんでしたが、解析できた MS 全体の約 2.6%に当たる約 20 万カ所の MS において複数個のがんゲノムに変異があり、これら変異 MS のほとんどは、タンパク質をコードしていないゲノム領域に存在していました。

また、変異 MS の個数は、全ゲノムでの SNV（一塩基置換）変異[14]数とは相関していませんでしたが、indel（短い挿入・欠失）変異[14]数と強い相関があり、MS 変異は indel 変異と同じ機構で起こることが分かりました。さらに、変異 MS の特徴を詳しく解析したところ、A（アデニン）や T（チミン）の繰り返しは変異率が高いこと（図 1 上）、MS 変異には DNA 複製のタイミングや DNA の構造が関連することなどが示唆されました。

そして、ゲノム全体で解析できた約 765 万カ所の MS のうち 3%以上に変異が見られた腫瘍を MSI 陽性と判定し、31 例の全ゲノムレベルでの MSI 陽性がんを特定しました。その中には、大腸がん（59 例中 7 例）、胃がん（78 例中 6 例）、子宮体がん（50 例中 11 例）が多く含まれ、他にも肝臓がん、膵臓がん、卵巣がん、腎臓がん、悪性黒色腫も 1 ～ 2 例ずつ含まれていました（図 1 下）。

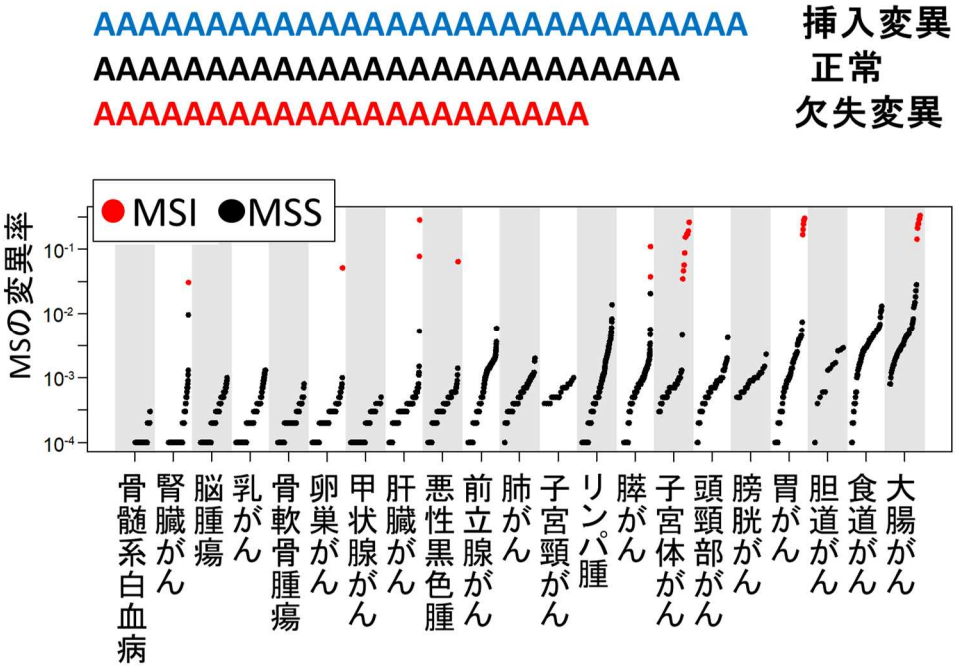


図１　21 のがん腫ごとの全ゲノムレベルでのマイクロサテライト（MS）の変異率

上：A（アデニン）が 26 回繰り返した MS（黒）に対する、挿入変異（青）と欠失変異（赤）を示した。
下：3%以上の MS に変異があるがんをマイクロサテライト不安定性（MSI; Microsatellite Instable）と判定した（赤丸）。大腸がん、子宮体がん、胃がんに MSI 陽性のサンプルが複数検出できた。MS の変異が 3%未満であるがんは、マイクロサテライト不安定性が認められない（MSS; Microsatellite Stable）とした（黒丸）。

この 31 例の MSI 陽性がんのうち、DNA 修復遺伝子の生殖細胞変異が 2 例で見つかりリンチ症候群に相当するとされ、DNA 修復遺伝子の多くの体細胞変異も見つかりました。また、MSI 陽性がんは変異によってネオ抗原[15]が多く発現し、抗原の産生などを誘導する免疫原性が高いと考えられています。本研究により、MSI 陽性がんにはネオ抗原が MSI 陰性がんよりも多く存在すること、MSI 陽性がんの多くはタンパク質をコードする遺伝子内の MS あるいは繰り返し配列の変異からできたことが推測されました。

さらに、高頻度に変異している 20 個の MS マーカーを新たに同定し、埼玉県立がんセンターにある大腸がんと子宮体がんの DNA を用いて、その MSI の診断精度を検証しました。その結果、新しい MS マーカーは、既存の MS マーカーと同程度の変異頻度を持ち、MSI 判定の精度も同等であることが分かりました（図 2）。

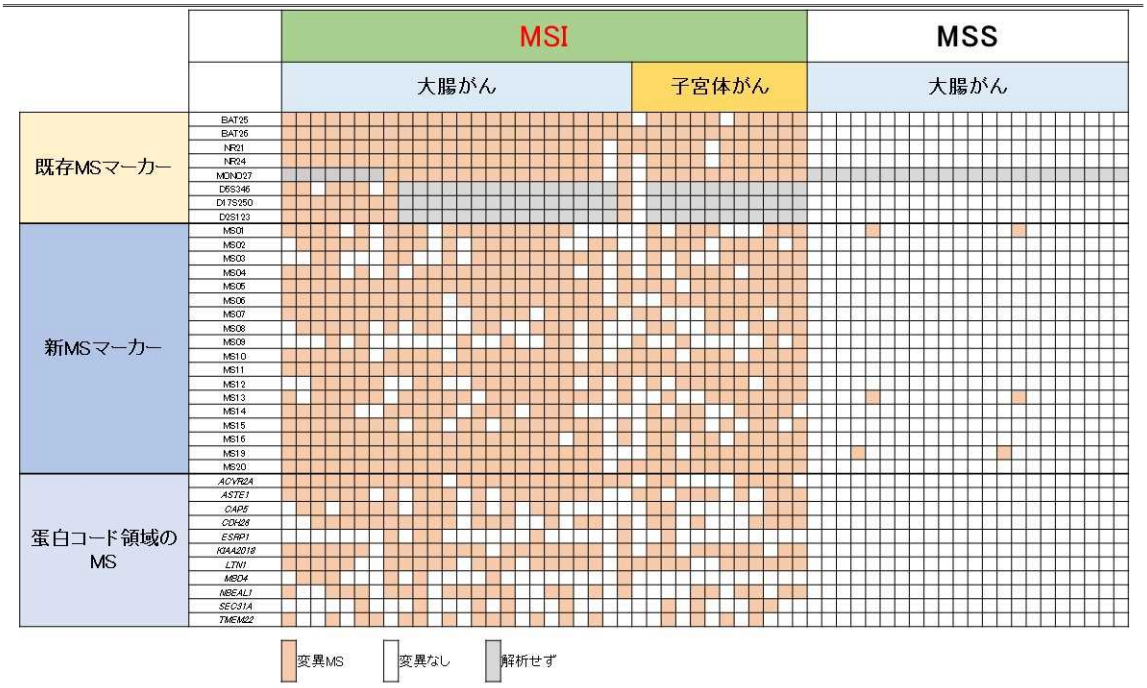


図 2　新規 MS マーカーによる大腸がんと子宮体がんの MSI 判定

既存の MSI 診断マーカー（8 個）、今回の高変異率の新規 MS マーカー（20 個）、およびタンパク質コード領域（11 個）の MS マーカーでの変異をオレンジで記している。MSI 陽性の大腸がん（24 例）、子宮体がん（12 例）では、同程度の変異頻度を持つことが分かった。MSI 陰性の大腸がん（22 例）における変異頻度も同程度であることから、MSI 判定の精度も同等であることが分かった。

3．今後の期待

今後、DNA シーケンス解析技術の革新に伴うコストの低下により、全ゲノムシーケンス解析が研究分野のみならず、がんの診断や個別化医療の分野においても標準的なゲノム解析手法になると予測されています。

MSI は、がんの本質である「ゲノムの不安定」に直接つながる現象であり、今回の大規模な全ゲノムシーケンスデータの解析による MSI についての知見は、がんの本体の解明につながるものであり、今後、この現象を標的とした治療法が開発されると考えられます。

また、今回開発した全ゲノムシークエンスデータからの MS 変異解析法（MIMcall）や MSI 判定法は、全ゲノム時代でのがん免疫療法の適応および遺伝性腫瘍の診断に使われるようになると期待できます。

4．論文情報

＜タイトル＞

Comprehensive Analysis of Indels in Whole-genome Microsatellite Regions and Microsatellite Instability across 21 Cancer Types

＜著者名＞
Akihiro Fujimoto, Masashi Fujita, Takanori Hasegawa, Jing Hao Wong, Kazuhiro Maejima, Aya Oku-Sasaki, Kaoru Nakano, Yuichi Shiraishi, Satoru Miyano, Go Yamamoto, Kiwamu Akagi, Seiya Imoto, and Hidewaki Nakagawa
＜雑誌＞
Genome Research
＜DOI＞
10.1101/gr.255206.119

5．補足説明

[1] マイクロサテライト（MS）
1～6 個の塩基が 10～数十回繰り返した配列のこと。ヒトゲノム中には 1000 万個以上の MS が存在すると考えられている。MS は Microsatellite の略。

[2] 全ゲノム横断的がん解析プロジェクト（PCAWG）
国際連携共同体である ICGC/TCGA 内のがんの全ゲノムシーケンス解析のデータを集積し、ICGC/TCGA の共同作業にてがんの横断的（PanCancer）解析を行うプロジェクトとして、2014 年に始動した。約 2,800 例のさまざまながんの全ゲノムシーケンスのビックデータを東京大学医科学研究所ヒトゲノム解析センターのスーパーコンピュータ「SHIROKANE」を含む世界 10 カ所のデータセンターで仮想空間を作り、分担して解析を行った。生シーケンスデータだけで約 1 ペタバイト（1000 兆バイト）の情報量になる。PCAWG は、PanCancer Analysis of Whole Genomes の略。

[3] マイクロサテライト不安定性（MSI）
細胞が分裂する際や環境曝露によって起こる、DNA の配列ミスを修復する機能が低下している状態のこと。がんでは、数個の MS マーカーの変異の有無（繰り返しの長さが変化する）によって MSI の判定が行われ、DNA 修復機構が欠損する遺伝性大腸がんの一つであるリンチ症候群の診断に使われている。MSI は Microsatellite Instability の略。

[4] 遺伝性腫瘍、リンチ症候群
がん全体の約 5%は、遺伝子の配列の変化により生まれつきがんを発症しやすい体質が原因で発症すると考えられている。これを遺伝性腫瘍という。遺伝性腫瘍にはいくつかの種類があるが、遺伝性乳がん卵巣がんや「リンチ症候群」が主である。リンチ症候群は、大腸がんや子宮体がんなどの発症リスクを高める遺伝性疾患で、全大腸がんのうち 2～5%の原因になっていると推定されている。

[5] 全ゲノムシーケンス解析
次世代シーケンサーを使って、個人（約 30 億塩基）やがんの全ゲノム情報を解読し、塩基配列の違いや変化を同定すること。データが大量になるため、大型計算機を使って情報解析を行うのが一般的である。タンパク質をコードする 1～2%の範囲のエクソンだけでなく、遺伝子の発現を制御するゲノム領域の変異やさまざまな構造異常（大きなゲノム配列異常）も検出可能で、究極のゲノム解析手法といえる。がんの場

合は、がんの DNA と同一患者由来の正常 DNA の全ゲノムシーケンス解析を行い、その差分を調べる。

[6] がんゲノムカタログ

30 億塩基からなるヒトゲノムマップで、どの位置にどういった型の変異が蓄積しているのかを、網羅的にカタログ化したもの。例えば、最も多く見つかる *KRAS* 遺伝子の変異は、12 番染色体の 25245350 の位置に C から T への変異が見られる。

[7] 遺伝的多型

同じヒトであっても個人個人のゲノムの塩基配列は多種多様であり、人口の 1 %以上の頻度で存在する配列の違いを遺伝子多型という。一つの塩基が他の塩基に置き換わっている一塩基多型（SNPs）や MS の長さの違いなどが、ゲノム解析の多型マーカーとして使われている。

[8] 免疫チェックポイン阻害剤

免疫のチェックポイントである PD-1 分子を抗体で抑制することで、腫瘍細胞への免疫抑制を解除し、免疫細胞が腫瘍細胞を排除できるように導く薬剤。現在、日本ではニボルマブとペムブロリズマブが、肺がん、胃がん、悪性黒色腫、腎臓がん、そして MSI 陽性の大腸がんなどについて承認され、治療に使われている。

[9] 次世代シーケンサー（NGS）

ヒトゲノムの全配列を 1,000 米ドル以下のコストで解読すべく、欧米の政府や企業が技術開発を行った結果、より高速高精度の性能を持つシーケンサーが開発された。この次世代シーケンサーは従来のものに比べ、超大量の DNA シーケンス反応を並列して行うことができる。NGS は Next Generation Sequencer の略。近い将来、100 米ドル以下のコストになるものと予測されている。

[10] ポリメラーゼ連鎖反応（PCR）

熱耐性を持つ DNA 依存性 DNA 合成酵素（DNA ポリメラーゼ）を用いて、鋳型 DNA と PCR プライマーの結合、相補 DNA の合成、二本鎖 DNA の解離の 3 ステップを繰り返すことで、DNA を連続的かつ指数関数的に増幅させることができる反応。PCR は Polymerase Chain Reaction の略。

[11] シーケンスの段階におけるエラー

細胞内における DNA の複製や修復と同様、PCR や次世代シーケンサーでの化学反応は主に DNA ポリメラーゼ（DNA 合成酵素）が担っており、一定の割合で間違い（エラー）が生じる。ただし、細胞内におけるエラーの多くは DNA 修復機能によって訂正される。

[12] X 染色体

ヒトの全遺伝情報（ゲノム）は、22 本の常染色体が 2 セットと、2 本の性染色体からなる。女性の性染色体は 2 本とも X 染色体である一方、男性では X 染色体と Y 染色体が 1 本ずつ存在する。このことから、MS のように極めて多型に富む（長さが異なる）マーカーを解析する際には、男性の X 染色体を用いれば 2 本の染色体における長さの変化を考える必要がない。そのため、男性の X 染色体は精度解析に適してい

る。

[13] アレル
哺乳類は母親と父親から同じ遺伝子セットを持つ染色体を 1 組ずつ受け継ぐ。この両親から受け継いだ 1 対の遺伝子セットを対立遺伝子またはアレルと呼ぶ。

[14] SNV（一塩基置換）変異、indel（短い挿入・欠失）変異
SNV 変異は、ゲノム上の一塩基（A、T、C、G）が変換される変異。Indel 変異は、1～数塩基配列の挿入または欠失による変異。SNV は Single Nucleotide Variant の略。

[15] ネオ抗原
タンパク質をコードする遺伝子に変異が入ると、アミノ酸が異なるタンパク質が発現するようになる。その一部のペプチド配列が腫瘍細胞の表面に提示されると、免疫細胞が非自己と認識し、腫瘍細胞を攻撃できるようになる。MSI 陽性がんでは MSI 陰性がんに比べてこのような変異が圧倒的に多く、ネオ抗原を多く発現する。そのため免疫原性が高く、免疫チェックポイント阻害剤の効果が期待できる。

6．発表者・機関窓口

＜発表者＞　※研究内容については発表者にお問い合わせください。

理化学研究所　生命医科学研究センター　がんゲノム研究チーム
　チームリーダー　　中川　英刀（なかがわ　ひでわき）
TEL：045-503-9286　FAX：045-503-9294
E-mail：hidewaki[at]riken.jp

東京大学　大学院　医学系研究科 国際保健学専攻 人類遺伝学分野
　教授　　藤本　明洋　（ふじもと　あきひろ）
　（研究当時：理研生命医科学研究センターがんゲノム研究チーム
　副チームリーダー）
TEL：03-5841-3692
E-mail：afujimoto[at]m.u-tokyo.ac.jp

＜機関窓口＞
理化学研究所　広報室　報道担当
TEL：048-467-9272　FAX：048-462-4715
E-mail：ex-press[at]riken.jp

※上記の[at]は@に置き換えてください。