日本版がんゲノムアトラス (JCGA, Japanese version of The Cancer Genome Atlas) の構築 - 日本人がん患者 5143 人から得た新鮮凍結腫瘍組織の解析 -

- 本論文は、「Japanese version of The Cancer Genome Atlas, JCGA, established using fresh frozen tumors obtained from 5143 cancer patients. Cancer Science 111(2):687-699, 2019」 をもとに、一部新たなデータを追加し、補完的な図表を本文中に配置した日本語版です。原著に 加筆、追加した部分はイタリックで区別しました。
- 本論文の内容は、プロジェクト HOPE として実施した日本人がん患者 5,143 症例から得た新鮮凍 結腫瘍組織のマルチオミクス解析の結果です。単一医療機関での研究であり、また、静岡がんセ ンターの臨床データの大部分は電子カルテ化されているため、倫理的な配慮の上で、ゲノムデー タと臨床データとの全面的な突合及び臨床経過の追跡が可能です。
- 3. 主要な解析結果は以下の通りです。
  - すべてのがん種を含む 5,143 症例のがん新鮮凍結組織を対象としたマルチオミクス解析(全 エクソーム、遺伝子コピー数、腫瘍遺伝子変異総数、MSI、がんパネル、融合遺伝子、遺伝子 発現)により 72.2%の症例でがんの原因となる遺伝子変化(ドライバー遺伝子変化)が同定 された。
  - ② がんドライバー遺伝子の候補として、7種のデータベースから 914種の遺伝子を抽出し、詳細に検討した結果、このうち 362種の遺伝子がドライバー遺伝子として検出された。
  - ③ 細胞がん化の原因遺伝子の機能を、11 カテゴリー、27 パスウェイに分類し、検出されたド ライバー遺伝子の分布を明らかにした。
  - ④ 当該がん種に適応のある分子標的薬や免疫チェックポイント阻害剤が応用可能な症例(ドラ ッガブル遺伝子変化を有する症例)は、全症例の11.3%であった。
  - ⑤ ドライバー遺伝子変化、ドラッガブル遺伝子変化はともにがん種による差異が認められた。
  - ⑥ 腫瘍遺伝子変異総数(TMB)により Hypermutator(20 変異以上/MB)と判定された症例 は 5.4%であり、その多くで原因となる遺伝子変化が明らかであった。
  - ⑦ 遺伝性腫瘍は、全症例の 1.1%に認められた。なお、今後精査が必要であるが、その他 9.2% では遺伝性腫瘍の原因とされる生殖細胞系列遺伝子変化が検出された。
  - ⑧ いわゆる二次的所見としての非がん性遺伝疾患については、米国臨床遺伝学会による開示推 奨 10 疾患が 0.4%で認められた。
  - ⑨ クローン性造血は、高齢者を中心に 8.4%で認められた。
- 現在、静岡がんセンターでは、エキスパートパネルにおける「見える化」を目指し、本論文のデー タを活用した「日本版がんゲノムアトラス」を 2020 年 11 月に公開する準備を進めています。
- 5. プロジェクト HOPE では、現在、本コホートで、全ゲノム解析 (Illumina, Novaseq 6000)、リ キッドバイオプシー (Illumina, Nextseg 550Dx) を実施し、ゲノム医療の展開を計っています。
- 6. プロジェクト HOPE で解析された症例は、すべて静岡がんセンターの症例です。静岡がんセンターでは、開設以来、すべての臨床データが電子化されており、このため、個々の症例のゲノム情報とすべての臨床データの突合が可能であり、さらに臨床経過の追跡も可能です。また、大多数の症例について、DNA、RNA、新鮮凍結組織、病理標本、血漿が保管されています。なお、商用利用の同意も得られているため、一定の条件下で創薬や診断薬開発に利用可能です。
- 7. 本研究成果は、静岡がんセンターと(株)エスアールエルの共同研究の成果です。

静岡県立静岡がんセンター プロジェクト HOPE

(問い合わせ窓口)ゲノムメディカルリサーチコーディネーター 水口(m.kojima@scchr.jp)

Japanese version of The Cancer Genome Atlas, JCGA, established using fresh frozen tumors obtained from 5143 cancer patients. Cancer Science 111(2):687-699.

日本版がんゲノムアトラス(JCGA, Japanese version of The Cancer Genome Atlas)の構築 -日本人がん患者 5143 人から得た新鮮凍結腫瘍組織の解析-

長嶋剛史<sup>1,2</sup>、山口建<sup>3</sup>、浦上研一<sup>1</sup>、下田勇治<sup>1,2</sup>、大浪澄子<sup>1</sup>、大島啓一<sup>4</sup>、田邉智絵<sup>1,2</sup>、成岡茜<sup>5</sup>、鎌田福美<sup>1</sup>、 芹澤昌邦<sup>5</sup>、畠山慶一<sup>4</sup>、松村研哉<sup>1</sup>、大浪俊平<sup>1</sup>、丸山宏二<sup>6</sup>、望月徹<sup>4</sup>、楠原正俊<sup>5,7</sup>、塩見明生<sup>8</sup>、大出泰久<sup>9</sup>、 寺島雅典<sup>10</sup>、上坂克彦<sup>11</sup>、鬼塚哲郎<sup>12</sup>、西村誠一郎<sup>13</sup>、平嶋泰之<sup>14</sup>、林央周<sup>15</sup>、清原祥夫<sup>16</sup>、坪佐恭宏<sup>17</sup>、片桐 浩久<sup>18</sup>、庭川要<sup>19</sup>、髙橋かおる<sup>20</sup>、柏木広哉<sup>21</sup>、中川雅裕<sup>22</sup>、石田裕二<sup>23</sup>、杉野隆<sup>24</sup>、髙橋満<sup>25</sup>、秋山靖人<sup>26</sup>

- 1. 静岡がんセンター研究所 診断技術開発研究部
- 2. 株式会社エスアールエル
- 3. 静岡がんセンター
- 4. 静岡がんセンター研究所 遺伝子診療研究部
- <sup>5.</sup> 静岡がんセンター研究所 新規薬剤開発・評価研究部
- <sup>6.</sup> 静岡がんセンター研究所 実験動物管理室
- 7. 静岡がんセンター研究所 地域資源研究部
- 8. 静岡がんセンター病院 大腸外科
- 9. 静岡がんセンター病院 呼吸器外科
- 10. 静岡がんセンター病院 胃外科
- 11. 静岡がんセンター病院 肝・胆・膵外科
- 12. 静岡がんセンター病院 頭頸部外科
- 13. 静岡がんセンター病院 乳腺外科
- 14. 静岡がんセンター病院 婦人科
- 15. 静岡がんセンター病院 脳神経外科
- 16. 静岡がんセンター病院 皮膚科
- 17. 静岡がんセンター病院 食道外科
- 18. 静岡がんセンター病院 整形外科
- <sup>19.</sup> 静岡がんセンター病院 泌尿器科
- 20. 静岡がんセンター病院 乳腺センター
- 21. 静岡がんセンター病院 眼科
- 22. 静岡がんセンター病院 再建・形成外科
- 23. 静岡がんセンター病院 小児科
- 24. 静岡がんセンター病院 病理診断科
- 25. 静岡がんセンター病院
- 26. 静岡がんセンター研究所 免疫治療研究部

## 責任著者

山口建

静岡がんセンター

静岡県駿東郡長泉町下長窪 1007 番地

Tel: 055-989-5230, Fax: 055-989-5793, E-mail: k.yamaguchi@scchr.jp

### 要約

本研究の目的は、各種固形がん 5143 症例の新鮮凍結腫瘍組織を対象として、そのゲノム異常を次世代シ ークエンサー等により解析し、日本版がんゲノムアトラスを構築することである。これは静岡がんセンタ ーで実施されている「プロジェクト HOPE」の研究成果の一部である。

本研究では、各種固形がん患者の DNA および RNA 検体を対象に、全エクソンシークエンス、がん遺伝 子パネル、融合遺伝子パネル、全遺伝子発現(マイクロアレイ)を実施し、遺伝子変異、構造変化、発現量 異常を網羅的に解析し、細胞がん化の要因を探るとともに、有効な薬剤の有無を評価した。データ解析に は、独自に開発した「静岡マルチオミックス解析プロトコル」を用いた。

細胞がん化に関わる体細胞ドライバー変化は 72.2%の検体で検出され、原因となる遺伝子の種類は 362 種に及び、1 検体あたりの平均数は 2.3 個であった。検出された遺伝子変化に対し、有効とされる薬剤の存 在状況は、現在、日本あるいは世界各国でそのがん種への適用が承認されている薬剤が全体の 11.3%で存 在し、他のがん種で承認されている薬剤や研究段階の医薬品を含めると 55.0%で候補薬剤が存在していた。 また、免疫チェックポイント阻害剤の効果判定のバイオマーカーとされる高頻度変異症例(腫瘍遺伝子変 異総量(TMB)が 20 変異/Mb 以上)は 5.4%を占めた。

生殖細胞系列変異解析においては、全体の 9.2%で遺伝性腫瘍の原因とされる病原性変異が検出され、そ の内 12.2%は臨床的に、あるいは確認検査でも病的であることが確認された。非がん性遺伝性疾患の病原 性変異は、米国臨床遺伝学会が患者への情報開示を推奨する 34 疾患、12 遺伝子を対象とした場合 0.4%の 症例で検出された。血液細胞で認められるクローン性造血は 8.4%で同定された。

本研究の成果である日本版がんゲノムアトラスは、日本人がん患者に対するゲノム診療の実践において 有用な情報であり、臨床現場での活用が期待される。また、この研究によって体細胞ドライバー遺伝子変 化が検出されなかった 27.8%については、現在、全ゲノム解析(Illumina, Novaseq 6000)を実施し、ゲノ ム非コード領域変異、ゲノム構造異常などの有無を検索中である。

# 略語

CCP: Comprehensive Cancer Panel: がん遺伝子パネル

CNV: Copy Number Variation: コピー数多型

FP: Fusion Panel: 融合遺伝子パネル

GEP:Gene Expression Profiling:全遺伝子発現解析

HOPE : High-tech Omics-based Patient Evaluation : オミックス解析に基づく患者評価

JCGA : Japanese version of The Cancer Genome Atlas : 日本版がんゲノムアトラス

SMAP : Shizuoka Multi-omics Analysis Protocol : 静岡マルチオミックス解析プロトコル

SNP : Single Nucleotide Polymorphism : 一塩基多型

SNV: Single Nucleotide Variant: 一塩基置換

TMB: Tumor Mutation Burden: 腫瘍遺伝子変異総量

WES: Whole Exome Sequencing: 全エクソンシークエンス

### 1. はじめに

がんゲノム解析によって、がん発症の分子機構、分子標的薬の対象となるアクショナブルなゲノム変化、 免疫チェックポイント阻害剤のバイオマーカーとしての腫瘍遺伝子変異総量(TMB)、遺伝性腫瘍および クローン性造血に関わる遺伝子異常などが明らかにされてきた<sup>1-7</sup>。しかし、これらの知見は、主として欧 米人がん患者を対象とした大規模なゲノム解析の結果に基づいて得られたものであり、日本人がん患者を 対象とし、がん種横断的にゲノム異常を網羅したデータセットは未だ報告されていない。がんゲノム解析 結果の活用に際しては、ドライバー変異における人種間差<sup>8-10</sup>、人種特異的な一塩基多型(SNP)の存在(例 えば日本人 SNP の 15%は日本人固有のものである)<sup>11</sup>を考慮した上で、これまでに得られた知見と照合す る必要がある。このように、日本人のがんゲノム異常の正確な評価には、データ解析を日本人向けに最適 化した方法で実施することが必須である。

本研究では、日本人がん患者 5143 症例から得られた新鮮凍結腫瘍組織 5521 検体を対象として、全エク ソンシークエンス(WES)、がん遺伝子パネルシークエンス(CCP)、融合遺伝子パネルシークエンス(FP)、 マイクロアレイに基づく全遺伝子発現解析(GEP)を実施し、そのゲノム異常を同定した。得られた結果 に基づき、細胞がん化の要因および有効な薬剤の有無を評価した。データ解析は、日本人 SNP ならびに日 本人がん患者における遺伝子異常の情報を統合した独自開発の解析パイプライン「静岡マルチオミックス 解析プロトコル(SMAP)」を用いて実施した。

本研究は、日本版がんゲノムアトラス(JCGA)の構築を目的として、2014 年から静岡がんセンターで 実施している「プロジェクト HOPE」<sup>12</sup>の研究成果の一部である。

#### 2. 材料と方法

2-1. 倫理規定

本研究は、静岡がんセンター内の倫理審査委員会により承認を受けた研究プロトコルに則って実施した (認証番号:25-33)。本研究への参加に際して、全ての患者から書面によるインフォームドコンセントを 得た。臨床検体を用いた全ての研究は、関連するガイドライン<sup>13</sup>に従って実施した。

#### 2-2. 臨床検体

静岡がんセンター病院で外科的手術を受けた患者から得られた腫瘍組織、隣接する正常組織、血液検体 を全エクソンシークエンス、がん遺伝子パネルシークエンス、融合遺伝子パネルシークエンス、全遺伝子 発現解析に用いた。新鮮切除標本は、病理医により研究用検体として分取され、その中で100 mg を超える 腫瘍組織を分析に用いた。本研究は、単一医療機関で実施し、一定期間内に手術組織が入手できたがん患 者について、病態生理学的特徴やがんの種類を特定せず、希少がんを含む全ての検体を分析した。単一医 療機関実施の研究であるため、すべての解析データは、臨床データとの突合が可能である。この点は、TCGA をはじめとするこれまで海外で実施された、特定のがん種を対象とした大規模ゲノム解析プロジェクトと 大きく異なる本研究の特徴である。

腫瘍特異的なゲノム変化を同定するために、腫瘍組織は常に対応する正常コントロールとあわせて解析 した。全エクソンシークエンスとがん遺伝子パネルシークエンスにおいては血液検体を、全遺伝子発現解 析においては腫瘍隣接正常組織を対照とした。また、融合遺伝子パネルシークエンスについては腫瘍組織 のみを用いた。全エクソンシークエンス、がん遺伝子パネルシークエンス、融合遺伝子パネルシークエン スは次世代シークエンサー(イオンプロトン、サーモフィッシャーサイエンティフィック社)を、全遺伝 子発現解析はマイクロアレイ(アジレント・テクノロジーズ社)を用いて行った。実験手法の詳細につい ては以前の報告<sup>14-17</sup>を参照されたい。全エクソンシークエンスとがん遺伝子パネルシークエンスにおける 平均リード深度(解析対象領域1塩基あたりの配列読取回数)は、それぞれ130と1169であった。 2-3. 変異検出

本研究では全エクソンシークエンスとがん遺伝子パネルシークエンスの2通りの方法によって変異解析 を行った。両手法の解析対象遺伝子数はそれぞれ18835個と409個である。本研究では一塩基置換(SNV) と挿入欠失(indel)のうち、タンパク質をコードするエクソン領域およびスプライス部位で検出された変 異を解析対象とした。

シークエンス結果は、装置付随のソフトウェアTorrent Suiteによって配列データに変換し、読み取った 塩基配列に合致する参照用ヒトゲノム配列(UCSC hg19)の位置を特定した。その結果は、血液と腫瘍を それぞれ別のBAMファイルに保存した。

体細胞変異解析においては「血液において参照用ゲノム配列と同一であり、かつ、腫瘍において参照用 ゲノム配列と異なる箇所」を同定する。この処理はIon Reporter AmpliSeqエクソームペア解析ワークフロ ーを用いて行った。生殖細胞系列変異解析においては「血液において参照用ゲノム配列と異なる箇所」を 同定する。これは変異解析ソフトウェアTorrent Variant Callerを用いて行った。

得られた変異リストから解析精度が低いと考えられる変異を除去し、その結果を以降の解析で用いた。 低クオリティ変異の判定には、シーケンスクオリティ値、変異のクオリティ値、リード数、変異リード割 合の情報を用いた。

2-4. アノテーション

遺伝子および遺伝子変異の各種データベース情報に基づく評価・意味づけ(アノテーション)は以下の 順に実施した。

- 1. 変異記述方法の標準化(解析ソフトウェアvt normalize<sup>18</sup>を使用)
  - a. 遺伝子変異を記載する方法として一般的に使用されているVCFファイルにおいて、一部の挿 入欠失や複数塩基置換について、全く同一の変異が複数の表記法によって記載されることが 知られている。標準化処理によりこのような記述の曖昧性を統一する。
- 2. 遺伝子変異に対する下記3種類の情報の付与(解析ソフトウェアSnpEff<sup>9</sup>を使用)
  - a. 変異がアミノ酸配列に与える影響の予測
    - アミノ酸置換、終始コドンの獲得、フレームシフトによる構造変化など
    - b. 影響度合いの分類
      - 終始コドン獲得、フレームシフト変異、スプライスサイト変異の影響度は大、その他の アミノ酸置換の影響度は中、同義置換変異の影響度は低
  - c. 遺伝子情報(遺伝子シンボル、コード配列変化、アミノ酸配列変化、配列登録番号)
- 3. 遺伝子および遺伝子変異に対する各種データベース情報(遺伝子機能、薬剤、疾患、変異頻度、デ ータベース中の登録番号)の付与
- 4. 細胞がん化の要因となるドライバー変化の評価(後述)
- 5. 分子標的薬の対象となるアクショナブル変化の評価(後述)
- 6. 当該がん種に使用可能な分子標的薬、免疫チェックポイント阻害剤などのドラッガブル変化の評価 (後述)

この際、転写産物すなわちmRNAの参照データベースとしてRefSeq<sup>20</sup>を使用した。各種データベースは解 析実施時点における最新版を使用した。

2-5. 代表転写産物の選択

選択的スプライシングによって1つの遺伝子から複数の転写産物が生成される場合、以下の手順に従って 代表転写産物1種を選択し以降の解析に用いた。

- 1. タンパク質配列データベースUniProt<sup>21</sup>において、各遺伝子に対する代表配列として登録されてい るアミノ酸配列と完全に一致するもの
- 2. UniProtが提供しているデータベース間の配列登録番号照合表の中で、UniProtに代表配列として 登録されている登録番号と一致するもの

上記手順で複数の結果が得られる場合、単一の転写産物が特定されるまで以下の順序で絞込みを行った。

- RefSeq mRNA配列の種別、(1) タンパク質をコードするmRNA、(2) タンパク質をコードしな いmRNA、(3) mRNAの配列が実験的に確認されたもの、(4) ゲノム解析から予測されたmRNA 配列
- 4. アミノ酸配列の長さが最長のもの
- 5. 転写産物配列の長さが最長のもの
- 6. RefSeq mRNA配列登録番号の文字列で並べ替えたとき先頭にくるもの

#### 2-6. 腫瘍遺伝子変異総量

1Mb(100万塩基)あたりの腫瘍特異的体細胞変異の数を腫瘍遺伝子変異総量(TMB)と定義し、以下 の式で算出した。

• TMB = N/L

「N」は腫瘍特異的体細胞変異の数、「L」は有効解析長である。有効解析長は、血液と腫瘍の両者におい てリード深度が20以上であった領域のサイズである。5395検体における平均は33.16Mbであった。従って、 腫瘍遺伝子変異総量 20個/Mbは、全エクソンシークエンスの結果663.2個の腫瘍特異的体細胞変異が検出 されたことに相当する。腫瘍遺伝子変異総量は全エクソンシークエンスの結果に基づき、一塩基置換と挿 入欠失、同義置換変異と非同義置換変異のすべてを合わせて計算した。

#### 2-7. 変異シグネチャ

変異シグネチャは、腫瘍特異的体細胞遺伝子変異で見られる塩基置換パターンに加えて前後1塩基を含め た合計96種類のパターンから定義されるプロファイルである。ある1つの検体について、その検体において 同定されたすべての体細胞変異から算出された96パターンの出現類度すなわち96個の数値の組合せが一つ のプロファイルとなる。喫煙、紫外線暴露、DNAミスマッチ修復系の損傷が特定の塩基置換を引き起こす ことが知られており、それぞれ異なる変異シグネチャを持つ。ここでは、全エクソンシークエンスで得ら れた体細胞変異リストに基づき、解析ソフトウェアdeconstrutSigs<sup>22</sup>を用いて変異シグネチャの寄与度を 算出した。COSMICに登録されている30種類の変異シグネチャ(version 2)<sup>23</sup>を参照用のデータとして用 いた。

2-8. コピー数多型

コピー数多型(CNV)の解析にはsaasCNV<sup>24</sup>を用いた。全エクソンシークエンスで得られた血液と腫瘍 の一塩基置換を、変異アリル頻度が90%以上のものをホモとして、それ以外をヘテロとして用いた。コピ ー数の増加(増幅)と減少(欠損)は以下の条件で抽出した。

- 増幅: P値<0.001、かつ、コピー数≥2.5
- 欠損: P値<0.001、かつ、コピー数≤1.5

2-9. 腫瘍含有量

腫瘍含有量の解析は、2種類の解析ソフトウェアFACETS<sup>25</sup>とSequenza<sup>26</sup>を用いて行った。いずれのプロ グラムも、全エクソンシークエンスあるいはがん遺伝子パネルシークエンスによって得られた正常と腫瘍 のペア解析の結果から腫瘍含有量を推定する手法である。ここでは、全エクソンシークエンスによって得 られた血液と腫瘍の結果を用いた。2手法で得られた値の平均を推定腫瘍含有量とした。なお、特に断りの ない限り、腫瘍含有量の推定値の如何に関わらず、すべての検体の解析結果を示した。

2-10. 生殖細胞系列SNPを用いた血縁解析

本研究の対象者は、その多くが静岡県民であるため血縁者が含まれる可能性が考慮された。そこで、同 一家系に由来する検体が、生殖細胞系列変異の結果解釈に与える影響を評価するために、生殖細胞系列SNP を用いて以下2通りの解析を行った。

- 全検体ペア間における生殖細胞系列SNPの一致率
  - ▶ 血縁関係にある検体のペアはそうでないペアに比べて一致率が高くなることが想定される
- ・ サンプル品質評価プログラムSeqSQC<sup>27</sup>の使用
  - SeqSQCでは親子関係(parent-offspring)、両親が共通の兄弟姉妹(full sibling)、片親が共 通の兄弟姉妹(half sibling)が検出可能である。

この解析には、全エクソンシークエンスの結果、血液検体中で検出された SNP を用いた。

2-11. ドライバー、アクショナブル、及びドラッガブルなゲノム変化の評価

- 本研究では、ゲノム変化を、
- 細胞がん化の要因となり得るドライバー変化
- 様々な開発段階における分子標的薬の標的となるアクショナブル変化
- 現在、がん種を特定し、臨床応用可能な分子標的薬・免疫チェックポイント阻害剤の対象となるド ラッガブル変化

という3種類の観点から評価した。ドライバー変化は、図1に示す手順に従って解析し、がん化との関連 度に応じて5段階に分類した。アクショナブル変化は、図2に示す手順に従って解析し、薬剤の承認状況 と適用対象がん種の情報に基づき5段階に分類した。評価には下記3種類のゲノム解析結果を用いた。

- 1. 遺伝子変異(全エクソンシークエンスとがん遺伝子パネルシークエンスの解析結果)
  - a. 一塩基置換
  - b. 挿入欠失
- 2. 構造異常(融合遺伝子パネルシークエンスの解析結果)

#### c. 融合遺伝子

- 3. 発現異常(全遺伝子発現解析とコピー数多型の統合解析の結果)
  - d. がん遺伝子の増幅
  - e. がん抑制遺伝子の欠損

本研究では、一塩基置換と挿入欠失を指す用語として「変異」を、上記5種類全てを指す用語として「変 化」を用いる。上記ゲノム変化のうち、体細胞および生殖細胞系列の遺伝子変異(一塩基置換と挿入欠失)

は、図3に示す手順に従って、細胞がん化との関連性が高い順に以下の5段階に分類した。

- Tier1 病原性変異としてデータベースに登録されている (pathogenic)
- Tier2 病原性変異の可能性有りとデータベースに登録されている (likely pathogenic)
- Tier3 予測プログラムによって病原性変異有りと予測された(predicted pathogenic)
- Tier4 非同義置換変異 (nonsynonymous)
- Tier5 同義置換変異 (silent)

変異の評価には、遺伝子変異データベース、遺伝子機能データベース、アリル頻度データベース、変異 が遺伝子機能に与える影響の予測結果を用いた<sup>28-37</sup>。本法では、がんドライバー遺伝子において検出され た腫瘍特異的な遺伝子変化をドライバー変化とし、また、人種特異的 SNP、特に日本人特有の SNP を評 価するために、日本人における変異頻度情報データベースを統合した。本研究では、Tier1 に分類された変 異をドライバー変異と定義し、以降の解析に用いた。

アクショナブル変化は、国立がん研究センターがんゲノム情報管理センターによって提案された基準に 基づき、図4に示す手順に従って、以下の5つのエビデンスレベルに分類した<sup>38</sup>。

A 当該がん種への適用が日本あるいは諸外国で承認されている薬剤が存在

- B 当該がん種の治験薬が存在
- C 他がん種への適用が承認されている薬剤が存在、あるいは他がん種の治験薬が存在
- D がん種を問わず小規模研究あるいは症例報告で有効性が報告された薬剤が存在
- E 前臨床試験の結果有効性が報告された薬剤が存在

さらに、本研究では、このうちレベル A として分類されたものをドラッガブルと定義した。オリジナル の分類案では、がん化と関連する変化を表すレベル F が存在するが、本研究ではドライバー変化を別途評 価しているため、アクショナブルの分類からは除外した。

変異の評価に際して、同定された変異が、(1)染色体番号、ゲノム座標、塩基置換パターンが全て一致、 (2)遺伝子シンボル、アミノ酸の位置、アミノ酸置換パターンが全て一致、のいずれかを満たした場合、 データベースに登録されている情報と一致すると判定した。



す。Tierの番号が小さいほど、がん化との関連を示す根拠となる情報の信頼度が高い、すなわちがん化との関連がより 確実であることを示す。





細胞および生殖細胞系列で検出された遺伝子変異におけるドライバー変異評価手順の詳細を示す。

Step1. Assi	gnment o	f drug status				
Database			Description in dat	tabase		
CGI	FDA/NCCI	N/CPIC/ELN guidelines	late trial, clinical trial	early trial, case report	pre-clincal	
CIViC	A : prover associatio	n/consensus n	B : clinical trial	C : case report	D : in vivo/	In vitro model
DEPO	FDA appro	oved	clinical trial	case report	pre-clinica	I
GDKD	FDA appro	oved, NCCN guidelines	late trial	early trial, case report	pre-clinica	I
OncoKB	1 : FDA re predictive 2 : standa predictive	cognized biomarker e of response rd care biomarker e of response	3 : compelling clinical evidence support the biomarker as being predictive of response		4 : compel evidence s biomarker of respons	ling biological support the as being predictive se
						•
Drug status	5	High	Medium	Low	Р	re-clinical
Step2. Assi	gnment o	f evidence level for a	ctionable alterations			
			Drug st	tatus		
		High	Medium	Low		Pre-clinical
Cancer	Same	A (druggable)	) В	D		E
type	Other	С	С	D		E
図4. 体細	抱アクシ	ョナブル変化の分	類手順			
アクショナン	ブル変化に	は、2 段階のステッフ	<b>プで分類した。薬剤の</b> 酒	承認状況に応	じた分類	(ステップ 1)、適
分類(ステン	ップ2)の	結果に基づいてAオ	からEの5段階に分類	した。		

2-12. がん関連遺伝子カタログの作成とパスウェイ分類

本研究では、細胞がん化の原因となるドライバー遺伝子変化が、どのような仕組みでがん化を引き起こ すのかを明らかにするため、これまでに知られている細胞がん化のメカニズムを、11 カテゴリー、27 パス ウェイに分類した。これらを構成する遺伝子群については、直接、細胞がん化を引き起こす可能性がある 遺伝子カタログを、以下2種類の遺伝子セットから作成した。(1)既知のがん遺伝子およびがん抑制遺伝 子、(2)既知のがん化の原因となる遺伝子変異を持つ遺伝子のうち現時点でがん遺伝子・がん抑制遺伝子 の分類が未確定であるもの。両者を合わせると 1087種が抽出されたが、ここから両者で重複している遺伝 子をまとめ、総計 914 種類のがんドライバー遺伝子として定義し、それらを対象にドライバー変化の評価 を行った。

- 1. がん遺伝子とがん抑制遺伝子(880 遺伝子)
  - COSMIC Cancer Gene Census<sup>39</sup>、OncoKB Cancer Gene List<sup>35</sup>、文献<sup>40-42</sup>から抽出
- がん化の原因となる体細胞病原性変異を持つ遺伝子(現時点でがん遺伝子・がん抑制遺伝子の分類 が未確定)(207 遺伝子)
  - CGI<sup>28</sup>、ClinVar<sup>36</sup>、DoCM<sup>32</sup>、OncoKB<sup>35</sup>、IARC-TP53<sup>34</sup>において病原性として登録されている 変異を有する遺伝子(ただし、IARC-TP53 は機能異常を意味する「非機能性」として定義さ れた変異のみを対象とした)

さらに、上記の 914 種類に留まらず、より広範ながん化に関与する可能性がある遺伝子を網羅的に解析 するため、各種がん遺伝子パネル、がん関連遺伝子リストを含む 27 のリソースから 1074 個の遺伝子を抽 出し、これをがん関連遺伝子とした。

こうして集められた 1988 種類のドライバー遺伝子、がん関連遺伝子について、その変化と細胞がん化の メカニズムの関係を追求する目的で、これらの遺伝子群を、KEGG<sup>43</sup>、UniProt<sup>21</sup>、文献情報<sup>241,42</sup>に基づく がん化パスウェイに割り当てた。パスウェイアノテーションの結果はマニュアルで確認した。その結果、 1462 遺伝子に対して、27 種類のパスウェイ情報が割り当てられた。パスウェイごとの遺伝子数、パスウェ イの機能カテゴリー分類を表1に示した。がんドライバー遺伝子 914 種とがん関連遺伝子 1074 種の合計 1988 遺伝子のリスト、ならびに各遺伝子に対する、がん遺伝子とがん抑制遺伝子の分類、パスウェイ情報 については原著論文のウェブサイトで公開されている表 S2 を参照されたい。

Category	Pathway		No. of gene	
		Cancer	Cancer	Total
		driver	related	
		gene	gene	
Cell cycle	Cell cycle	25	7	32
	Cell division	13	9	22
Cell death	Apoptosis	17	4	21
Cell growth	GPCR	12	4	16
	Нірро	26	1	27
	JAK/STAT	15	3	18
	МАРК	51	41	92
	MYC	7	0	7
	NFKB	12	6	18
	Nuclear receptor	6	4	10
	PI3K/Akt/mTOR	50	7	57
	RTK	61	33	94
Differentiation	Hedgehog	6	5	11
	NOTCH	19	30	49
	WNT	78	6	84
Epigenetic regulation	Chromatin modification	82	58	140
Genome maintenance	Core DDR	30	39	69
	DNA damage control	39	53	92
	TP53	8	0	8
Immune	Immune	47	49	96
Invasion	TGF-B	12	7	19
Metabolism	Drug metabolism	0	16	16
	Metabolic pathway	21	69	90
Protein homeostasis	KEAP1/NRF2	3	0	3
	Protein homeostasis	30	50	80
Transcriptional regulation	RNA metabolism	45	51	96
	Transcriptional regulation	132	63	195
Gene with pathway annotat	ion	847	615	1462
Gene without pathway anno	otation	67	459	526
	rotal	914	1074	1988

## 3. 結果

3-1. HOPE コホート

HOPE コホートを構成する症例・検体は、静岡がんセンター病院で 2014 年 1 月から 2019 年 3 月までの 期間に、外科的手術・治療を受けたがん患者 5143 人から採取された 5521 検体の新鮮凍結腫瘍組織である。 5521 検体のうち、5020 検体(90.9%)は原発性腫瘍、501 検体(9.1%)は転移性腫瘍であった。また、1 症例あたり 2 検体以上持つものが 336 症例 714 検体、残りは全て 1 症例につき 1 検体であった。

がん種ごとの検体数は図 5 に示すとおり、上から順に、結腸がん(1014 検体、18.4%)、肺がん(905 検 体、16.5%)、直腸がん(733 検体、13.3%)、胃がん(599 検体、10.8%)、頭頸部がん(344 検体、6.2%)、 乳がん(288 検体、5.2%)、肝臓がん(242 検体、4.4%)であり、これら上位 7 種が全体の 74.7%を占め た。一方、希少がん<sup>44</sup>は全体の 9.5%(523 検体)であった。このうち、消化管間質腫瘍(GIST)が最も多 く(86 検体)、次いで脳腫瘍(79 検体)、頭頸部がん(62 検体)、肉腫(58 検体)の順であり、合計 50 種 の希少がんが含まれていた。

HOPE コホートには病理学的に良性腫瘍と診断された 39 検体も含まれているが、これらについては以降の解析からは除外した。



HOPE データセットにおける各がん種の検体数を原発性腫瘍と転移性腫瘍のそれぞれについて示した。症例数は括弧内 に示した。検体数が10未満のがん種は「Other」としてひとまとめにした。本研究では、最終的に良性腫瘍と判定され た症例が38例含まれており、これらは検体数に関わらず「Benign」と表記した。

#### 3-2. 遺伝性腫瘍に対する生殖細胞系列変異

生殖細胞系列で生じた特殊な遺伝子変異が、がんを発症させ、遺伝性腫瘍の原因となることは広く知ら れている。本研究では、腫瘍組織における変異解析の比較対照として血液検体の全エクソンシークエンス を実施しており、生殖細胞系列変異の評価が可能である。そこで、血液検体の変異解析結果に基づき、遺 伝性腫瘍の原因となる病原性変異の有無を探索した。ここでは、最終的に全エクソンシークエンスで認め られた変異について、当該遺伝子に関する確認検査の実施や臨床情報との突合が終了した 3022 症例 3022 血液検体を対象に、遺伝性腫瘍の原因遺伝子として報告されている 49 遺伝子において検出された変異を検 索した。ここでは、遺伝子名、遺伝子変異の位置、塩基置換パターンの全てが病原性変異として変異デー タベースに登録されており、かつ集団内における変異アレルの頻度が 1%未満のものを生殖細胞系列ドラ イバー変異として抽出した(図 3B)。

解析の結果、遺伝性腫瘍の原因とされる病原性変異が、全体の 9.2%(279 検体)において、25 種類の遺 伝子で同定された(表 2)。上位 5 遺伝子は *MSH2、BRCA1、CDH1、SDHD、APC* であり、これらが全 体の 57.7%を占めた。

この 279 検体中 34 検体、12.2% (全体の 1.1%) はサンガー法ならびに MLPA 法により病原性変異の存 在が確認され、さらに 6.1%は家族歴などの臨床情報に基づき遺伝性腫瘍と診断された。これらの症例にお けるドライバー変異は 14 遺伝子で検出され、変異検出数ごとに *BRCA1*(9 検体)、*MLH1*(5 検体)、*BRCA2* (4 検体)、*CHEK2*(3 検体)、*PTEN*(2 検体)、*SDHB*(2 検体)であった。 なお、これらのデータが、ある地域において血縁関係にある症例によって頻度などに影響を与えている か否かを評価するため、病原性変異が確認された 34 検体(全体の 1.1%)について家族歴の精査を、また、 病原性変異が同定された 279 検体すべて(全体の 9.2%)について生殖細胞系列 SNP の解析を行った。そ の結果、前者においては同一家系由来の検体が存在しないこと、また、後者では 2 検体が同一家系に由来 するものと推定された。この 2 検体は結腸がん 1 検体と直腸がん 1 検体であり、ともに *APC*p.Gln2322Arg がドライバー変異として検出された。残り 277 検体においては同一家系由来と推定される検体は含まれて いなかった。

がんに関連する生殖細胞系列の遺伝子変化を日本人と欧米人とで比較するため、本研究結果と TCGA グ ループの結果を分析した。3022 症例中 279 症例(全体の 9.2%)で検出された遺伝性腫瘍の原因とされる 生殖細胞系列変異が認められた 25 遺伝子中 22 遺伝子は、本研究、TCGA コホートの両者で検出された。 この 22 遺伝子が両データセットにおいてカバーする変異の割合は、HOPE では 97.5%、TCGA では 39.5% であった。両者で共通して高頻度に病原性変異が検出された遺伝子は頻度順に BRCA1、CHEK2、BRCA2、 TP53、NF1、MSH6であった。一方、両者の差異は、データセット内のがん種構成ならびに解析手法の違 い(解析対象とする遺伝子数、マイナーアリル頻度のカットオフ値、変異の病原性評価に用いるデータベ ース)に由来すると考えられた。

2. 遺伝性腫瘍に対	する生殖細胞糸列ト	フイバー変異の割合
	No. of cases	Prevalence
Examined	3022	
Detected	279	9.2%
Confirmed	34	1.1%
Gene	No. of cases	Cancer type
BRCA1	9	Breast (3), Rectum (2), GIST(1), Lung (1),
		Ovary (1), Pleura (1)
MLH1	5	Colon (4), Stomach (1)
BRCA2	4	Breast (2), Colon (1), Head and neck (1)
CHEK2	3	Breast (1), Lung (1), Rectum (1)
PTEN	2	Breast (1), Pancreas (1)
SDHB	2	Breast (1), Head and neck (1)
MSH6	2	Colon (1), Uterus (1)
MSH2	1	Uterus (1)
EXT2	1	Colon (1)
CDH1	1	Breast (1)
NF1	1	Stomach (1)
NTRK1	1	Stomach (1)
TP53	1	Sarcoma (1)
VHL	1	Rectum (1)

### 3-3. 非がん性遺伝性疾患の生殖細胞系列変異

本解析では、米国臨床遺伝学会のガイドラインに従って、患者への開示対象として推奨されている 34 遺 伝子の変異が原因であるとされている 12 種類の非がん性遺伝性疾患を対象とし <sup>45</sup>、生殖細胞系列変異を評 価した。遺伝性腫瘍の解析で用いたものと同じ 3022 症例 3022 血液検体を対象とした結果、11 検体(全体 の 0.4%)で開示対象の非がん性遺伝性疾患に対する 11 の病原性変異を確認した(表 3)。その内訳は、家 族性高コレステロール血症 3 検体(*LDLR*)、肥大型心筋症 3 検体(*MYH7*)、マルファン症候群 1 検体 (*FBN1*)、Fabry 病 1 検体(*GLA*)、QT 延長症候群 1 検体(*KCNH2*)、家族性肥大型心筋症 1 例(*MYL2*)、 心筋症 1 検体(*TNNT2*) であった。

表 3. 非がん性遺伝性疾患の	に対する生殖細	胞系列ドライバー変異の割合
	No. of cases	Prevalence
Examined	3022	
Detected	33	1.1%
Confirmed	11	0.4%
Gene	No. of cases	Disease
LDLR	3	Familial hypercholesterolemia
MYH7	3	Hypertrophic cardiomyopathy
FBN1	1	Marfan syndrome
GLA	1	Fabry's disease
KCNH2	1	Long QT syndrome
MYL2	1	Familial hypertrophic cardiomyopathy
TNNT2	1	Cardiomyopathy

3-4. クローン性造血

クローン性造血は、造血幹細胞の生殖細胞系列変異ではなく、その体細胞変異に起因するもので、変異 を有するクローンは新たな変異が発生し、変異が累積することによって、骨髄異形成症候群および白血病 に進行するとされている<sup>46</sup>。本研究では、腫瘍特異性体細胞変異を検出する目的で、血液細胞をコントロ ールとして用いているため、クローン性造血を生じた血液細胞は体細胞変異の評価に影響を及ぼす可能性 があると考えられた。

そこで、血液循環がん細胞やがん細胞由来の血中遊離変異 DNA の影響を最小限にするため、転移症例 除き、また、重複がんなどがん既往歴のない原発性腫瘍のみを有する患者由来の 3751 症例 3751 血液検体 を対象に、既報の手法 47 を用いて、生殖細胞系列で検出された一塩基置換を分析した。本解析では、先行 研究にしたがって既知のがん遺伝子とがん抑制遺伝子(880 遺伝子)において同定された一塩基置換を用 いた。その結果、全体の 8.4%でクローン性造血が検出された(表 4)。より厳格な基準で評価した結果、41 検体(1.1%)は、他の研究で報告されたものとゲノム中での変異箇所および塩基置換パターンレベルで同 一であることが確認された<sup>7,47</sup>。これら 41 検体における変異遺伝子は表 4 に示す 11 種類であった。以前の 報告と同様に、クローン性造血は高齢者でより高頻度に観察された。

	No. of cases	Prevalence
Examined	3751	
Detected	316	8.4%
Identical to reported	41	1.1%
clonal hematopoiesis		
Gene with reported	DNMT3A (18), SRSF2	(5), TET2 (5), ASXL1
clonal hematopoiesis	(4), IDH2 (3), SF3B1 (3	3), GNAS (2), JAK2
	(2), NRAS (1), TP53 (1	), U2AF1 (1)

## 3-5. 腫瘍遺伝子変異総量と変異シグネチャ

腫瘍遺伝子変異総量は100万塩基当たりの腫瘍特異的体細胞変異の数で、近年、免疫チェックポイント 阻害剤の有効性を判定するバイオマーカーとして注目を集めている。*腫瘍遺伝子変異総量のがん種比較か ら、多くのがん種において高頻度変異症例が存在し、免疫チェックポイント阻害剤の治療効果を予測しう ることが、欧米人を対象としたがんゲノム解析の結果から報告されている。*ここでは全エクソンシークエ ンスによって腫瘍特異的体細胞変異の詳細が明らかにされた 5395 検体を対象に腫瘍遺伝子変異総量の解 析を行った。

全がん種を対象としたがん種横断的解析の結果、腫瘍遺伝子変異総量の二峰性分布が明らかとなった(図 6A)。そこで、二峰性分布の谷に相当する変異数をもとに、高頻度変異症例の分類基準を「腫瘍遺伝子変異 総量≥20 変異/Mb 以上(100 万塩基あたり 20 変異以上)」としたところ、全体の 5.4%(292 検体)が高頻 度変異症例と分類された(図 6A、B)。がん種ごとの比較解析からは、がん種によって腫瘍遺伝子変異総量 の数値分布が異なること、同一がん種内であっても検体間で腫瘍遺伝子変異総量が大きく異なることが確 認された(図 6C)。二峰性の分布は、脳腫瘍、胃がん、結腸がん、子宮がん、黒色腫を含むいくつかのが んで観察され、二つの峰の中間値(高頻度変異症例のカットオフ)はこれらのがん種で異なっていた。一 方で、他のがん種では明確な二峰性は見出されなかった。



次に、変異シグネチャの分析を行った。変異シグネチャは、喫煙、紫外線などの細胞がん化を引き起こ す外的要因や、POLE などの特殊な腫瘍特異性遺伝子変異の結果として、がん組織における特異な遺伝子 変化パターンが生じる現象である。変異シグネチャは、変異箇所の塩基置換パターンおよび前後の塩基を 含む合計 96 パターンの出現頻度によって定義され、COSMIC には 30 種の変異シグネチャが登録されてい る。代表的なものは、喫煙による C>A 置換、紫外線暴露による CC>TT 置換などで、変異惹起要因と一塩 基置換パターンの対応関係に基づき、塩基置換パターンから変異惹起要因を推測することが可能とされて いる。ここでは、高頻度変異症例の変異リストから、COSMIC に登録されている 30 種の変異シグネチャ それぞれの寄与度を各症例について算出することで、がんの要因の推定を行った。

その結果、一部のがん種では、明確ながん種別変異シグネチャが観察された(図7)。典型的な例は、肺 腺がんにおける喫煙シグネチャ(Sig.4)、結腸がんと子宮がんにおけるミスマッチ修復欠損シグネチャ (Sig.6)、黒色腫における UV シグネチャ(Sig.7) などである(図8)。これに加えて、極めて高い腫瘍遺 伝子変異総量と POLE シグネチャ(Sig.10)を同時に有する一群の検体が確認された(図7)。これら POLE カテゴリーの検体は、大腸がんおよび子宮内膜がんに限定されていること、DNA 修復酵素であるポリメラ ーゼイプシロンの機能喪失型変異により多数の突然変異が蓄積されていることが特徴的であり、これらの 表現型は DNA 修復酵素遺伝子の異常によるリンチ症候群のものとは異なると考えられた<sup>48</sup>。



上、中、下の各パネルに、がん種、腫瘍遺伝子変異総量、変異シグネチャの寄与度を示す。一つの細い縦列が個々の高 頻度変異症例を表す。寄与度>0 のシグネチャのみ、20 以上の検体が解析されたがん種のみを表示した。括弧内に記し た各シグネチャに対応する病因は COSMIC に登録されている情報に基づく。病因に関する情報がないものは「unknown」 と記載した。1 検体あたりの変異数が少ない場合、変異シグネチャの寄与度の推定精度が低下するため、腫瘍遺伝子変 異総量が 20 以上の 292 の高頻度変異症例を対象とした。



3-6. がん化の原因となる体細胞ドライバー遺伝子変化

細胞がん化の原因となるゲノム変化を特定するために、Shizuoka Multi-omics Analysis Protocol (SMAP、 静岡マルチオミックス解析プロトコル) (図1)を用いて、腫瘍特異的な体細胞遺伝子変化を評価した。こ の検討は、5521 検体のうち、下記3種類のゲノム解析結果すべてを得ることができた4131 検体を対象に、 以下の基準を満たす遺伝子変化に対して実施した。

- ▶ 一塩基置換と挿入欠失からなる変異データセット
  - > 図 3A に示す手順に従って解析し、Tier1 に分類された変異
- 融合遺伝子データセット
  - ▶ 検出された全ての融合遺伝子
- 発現データセット
  - がん遺伝子の増幅(正常に比べて腫瘍で発現量が5倍以上に亢進、かつコピー数が2倍以上 に増加)
  - がん抑制遺伝子の欠損(正常に比べて腫瘍で発現量が1/5以下に低下、かつコピー数が1/2 以下に減少)

この基準による検討の結果、対象とした 4131 検体中 2982 検体(72.2%)において、細胞がん化の原因 と考えられる 6817 個の体細胞ドライバー遺伝子変化が検出された(図 9)。これは、1 検体あたり平均 2.3 個の変化が存在したことを示しており、1 検体あたりのドライバー遺伝子変化は最少の 1 個であるものが 1120 検体で一番多く、一方、最大は 15 個で 1 検体において認められた。本研究では、体細胞ドライバー 遺伝子として 914 種類を想定していたが、今回の HOPE コホート 4131 検体中 2982 検体においてドライ バー遺伝子変化と判定された遺伝子は 362 種類であった。ドライバー遺伝子変化の検出率はがん種間で異





ドライバー遺伝子変化を、塩基配列変化、発現異常、構造異常などの種類ごとにみると、遺伝子変異(一 塩基置換と挿入欠失)の検出率が 62.2%、発現異常が 21.0%、融合遺伝子が 12.9%であった。13.0%の検 体では、これら 3 種の変化のうちの複数種を有していた。また、遺伝子変異、発現異常、融合遺伝子 3 種 の頻度もがん種間で異なっていた(図 10)。結腸がんと直腸がんにおけるドライバー遺伝子変化の 85%以 上は遺伝子変異であり、ミスセンス変異と短縮型変異(フレームシフト変異、ナンセンス変異、スプライ スサイト変異の 3 種)がほぼ同数であった。一方、肉腫、腎臓がん、肝臓がんにおけるドライバー変化の 70%以上は発現異常であった。

遺伝子変化の検出頻度においても、がん種に特徴的な遺伝子およびがん種横断的に検出される遺伝子の 双方が認められた(図11)。前者の典型例は、結腸がんと直腸がんにおける APC、黒色腫における BRAF、 肺腺がんにおける EGFR、GIST における KIT などであり、後者の代表例はがん種横断的に検出される KRAS と TP53である。

ドライバー遺伝子変化が検出された 362 種類の遺伝子をパスウェイレベルで見ると、その多くが細胞増 殖 (MAPK、PI3K、RTK パスウェイ)、分化 (WNT パスウェイ)、ゲノム安定性の維持 (TP53 パスウェ イ) に関与する遺伝子であった (図 12)。







本研究の対象とした HOPE コホートは、単一組織で実施された、日本人のがん種横断的ゲノムデータセ ットとして最大規模のものである。そこで、本研究による日本人のデータセットと、TCGA コホートから 抽出された白人由来のデータセットにおいて、共に 20 以上の検体が含まれていた 18 のがん種を対象とし て体細胞遺伝子ドライバー変化の人種間比較を行った。その結果、8 遺伝子で人種間の違いが検出された。 TCGA コホート中の白人由来データセットで頻度が高かったのは、肺腺がんと膵臓がんにおける KRAS、 胃がんにおける PTEN であった(図 13)。一方、本研究結果における日本人コホートでは、肺腺がんの EGFR、乳がんの AKT1、卵巣がんにおける CTNNB1、KRAS、PIK3CAの頻度が高かった。



次に、ホットスポットに位置する変異について両コホート間で頻度の比較を行った。TCGA コホートに おける白人由来データセットでは、胃がんの PTEN p.K267fs、肺腺がんの KRAS p.G12C の頻度が高い一 方で、本研究の日本人コホートでは、肺腺がんにおける EGFR p. E746\_A750del と p.L858R の頻度が高か った。また、KRAS p.G12D と pG12V、PIK3CA p.E542K は日本人コホート全体において高頻度であった。

膵臓がんおよび乳がんは、間質もしくは脂肪組織が豊富であることから腫瘍含有量が低い傾向があり、 腫瘍遺伝子変異総量や変異頻度を評価するうえでその影響を受ける可能性がある。そこで推定腫瘍含有量 が 0.25 未満の検体を除外し再度検討を行った結果、*KRAS* および *AKT1* 変異頻度に差は認められなかっ た。また、卵巣がんでは、コホート間で組織型の構成が異なっていた。TCGA コホートは漿液がんのみを 含む一方で、本研究における日本人コホートは明細胞がん、類内膜がん、粘液がんを含んでおり、卵巣が んにおいてみられた両コホート間の変異頻度の差には、組織型の構成比の違いが影響している可能性が考 えられた。

3-7. 分子標的薬の対象となるアクショナブルな遺伝子・ゲノム変化

ゲノム情報に基づいた薬剤選択は、がんゲノム解析によって得られる最も重要な結果の一つである。本 論文では、細胞がん化の原因となるドライバー遺伝子変化あるいはゲノム変化に対し、有効な薬剤が存在 あるいは研究されている場合を「アクショナブル遺伝子・ゲノム変化」とし、また、そのうち、当該がん種 に一致して、本邦あるいは海外において薬剤が承認されている場合を「ドラッガブル遺伝子・ゲノム変化」 と称することとしている。ここで、「遺伝子・ゲノム変化」としている理由は、ドライバー遺伝子変化が細 胞がん化の原因であるのに対し、アクショナブルあるいはドラッガブル遺伝子・ゲノム変化の場合、がん 化の原因である場合に加え、免疫チェックポイント阻害剤の場合のように、原因ではなく、変化が特定の 遺伝子のみならず全ゲノムに及び、それが薬剤選択のバイオマーカーとなる場合があるためである。さら に、腫瘍における遺伝子変異が、ある薬剤の有効性を減じる薬剤抵抗性変異も知られており臨床応用され ている。

上記で定義したアクショナブルな遺伝子・ゲノム変化の評価においては、遺伝子変異、遺伝子増幅と欠 損、発現異常、腫瘍遺伝子変異総量、マイクロサテライト不安定性が対象となる。本研究で得られた体細 胞遺伝子変化から有効な薬剤の有無を判定するために、SMAP 法(図 2)を用いた評価を行った。この検 討は、ドライバー変化の解析に用いたのと同じ 4131 検体を対象として実施した。

今回のデータを用いて、ドラッガブル遺伝子・ゲノム変化を持つ症例を抽出した結果、全体の 11.3%に あたる 467 検体において、現時点で、日本あるいは世界各国で当該がん種への適応が承認されている薬剤 が同定され(図 14 の赤いバー)、このうち、107 検体(全体の 2.6%)が免疫チェックポイント阻害薬の候 補として推定された。ドラッガブルな変化は、6 種類の遺伝子、融合遺伝子、マイクロサテライト不安定性 (MSI-high)に集中していた。このうち最も頻度が高かったのは*KIT*(GIST の 81.4%、黒色腫の 13.6%) であり、次いで *EGFR*(肺腺がんの 40.8%)、*BRAF*(黒色腫の 31.8%)、MSI-high(子宮がん 9.5%、結腸 がん 7.5%、胃がん 5.6%)、融合遺伝子(肉腫の 8.2%、肺腺がんの 3.7%)、*PDGFRA*(GIST の 2.9%)、 *ERBB2*(乳がんの 2.7%、胃がんの 2.4%)、*MET*(肺腺がんの 0.7%)の順であった(図 15)。

当該がん種に対応する治験薬が同定されたのは、全体の 18.7%にあたる 773 検体であった(エビデンス レベル B、図 14 の青いバー)。また、全体の 14.0% (577 検体) はエビデンスレベル C (図 14 の緑のバー) に分類され、分子標的薬の適用拡大の候補であった。小規模スタディあるいは症例報告において有効性が 確認されているエビデンスレベル D は、全体の 7.9% (326 検体) であった。さらに、前臨床の実験結果が 報告されているエビデンスレベル E は、全体の 3.2% (131 検体) であった。残りの 45.0% (1857 検体) では、可能性ある薬剤に関する情報は存在しなかった。

エビデンスレベル A-E 全てを対象とした場合、全体の 55.0%にあたる 2274 検体でアクショナブル変化 が検出された。検出率は、がん種間で大きく異なっていた。最も高かったのが GIST の 84.3%、次いで黒 色腫の 77.3%、子宮がん(75.0%)、結腸がん(71.6%)、肺腺がん(71.4%)、直腸がん(69.1%)であった (図 14)。

一方、腫瘍における遺伝子変化によって薬剤有効性が減じる場合が知られている。代表的なものに、大 腸がんにおける KRAS活性化型変異、肺がんにおける EGFR エクソン 20 の挿入がある。このような治療 抵抗性変異は、全体の 18.3%で検出され(756 検体)、最も割合が高かったのは直腸がんにおける 45.3%

(*KRAS*, *NRAS*)、次いで結腸がんにおける 41.9%(*EGFR*, *KRAS*, *NRAS*)、肺がんにおける 13.5%(*EGFR*、 *KRAS*)、GIST における 1.4% (*PDGFRA*) であった。



# 3-8. ドライバー遺伝子変化が認められない症例への対処

本研究においては、全がん種を合わせ 72.2%の症例で細胞がん化を説明しうるがん関連遺伝子変化を 検出し得たが、一方で、27.8%の症例については、細胞がん化を説明できる遺伝子変化を検出できなかっ た。そこで、現在、未検出例について全ゲノム解析 (Illumina, Novaseq 6000)を実施し、さらに、これ らの症例の追跡を目標としてリキッドバイオプシーのシステムを開発している (Illumina, Nextseq 550Dx)。

#### 4. 考察

静岡がんセンターでは、日本版がんゲノムアトラス JCGA を構築するために、2014 年 1 月より単一施設 におけるマルチオミックス研究「プロジェクト HOPE」を実施している。本研究では、日本人がん患者 5143 名から採取された 5521 検体の新鮮凍結腫瘍組織に対して、全エクソンシークエンス、がん遺伝子パネルシ ークエンス、融合遺伝子パネルシークエンス、マイクロアレイによる全遺伝子発現解析を行った。本研究 で用いた試料は、すべて静岡がんセンター病院で外科的に腫瘍切除を受けた患者由来のものであり、十分 な量の腫瘍組織が利用可能である。そのため、様々な技術による繰り返し研究が可能である。また、病態 生理学的特徴、あるいはがん種に関係なく、全ての検体を解析対象としたため、希少がんを含む様々なが ん種についての解析が可能である。加えて、単一医療機関における研究であるため、個々の症例について すべての臨床情報の突合が可能である点も特徴の一つである。

本研究における日本版がんゲノムアトラス JCGA の構築に合わせて、日本人がんゲノムにおける遺伝子 変化の情報から、細胞がん化との関係ならびに有効な薬剤を評価するための新しい解析方法として SMAP 法 (Shizuoka Multi-omics Analysis Protocol、静岡マルチオミックス解析プロトコル)を確立した(図1、 2)。本法では、既報に基づき 914 種類のがんドライバー遺伝子を想定し、そこで検出された腫瘍特異的体 細胞変異、融合遺伝子、がん遺伝子の増幅、がん抑制遺伝子の欠損を体細胞ドライバー遺伝子変化と定義 した。さらに、遺伝性腫瘍の原因となる生殖細胞系列の遺伝子解析においては、49 種類の遺伝性腫瘍原因 遺伝子を対象とし、ClinVar あるいは HGMD に病原性として登録されている変異を生殖細胞系列ドライ バー遺伝子変異とした。また、これらの遺伝子変化をパスウェイレベルで評価するために、ドライバー遺 伝子のパスウェイ分類を行った。

SMAP 法においては、日本人特有の SNP の取り扱いについても配慮した。全世界 26 集団の解析から、 人種特異的な SNP が存在し、SNP の 15%は日本人に特異的であることが報告されている。また、異なる 集団間における SNP の共有化率は 0.5-0.9 との報告もある <sup>49</sup>。このような人種特異的な SNP の存在は、 ゲノム解析結果の臨床的意義の評価を複雑にする一因となり得る。これは、以下に示す、現在の解析方法 における SNP 情報の取り扱いに起因する。

- 生殖細胞系列の変異解析において、SNP は非病原性変異として除外される。
- 体細胞変異解析において、病的意義が不明な変異のうち、SNP に該当するものは機能的影響が低いとみなされる。
- 一部のがんパネル検査では、腫瘍検体のみの体細胞変異解析が行われている。この場合、比較対象 として世界標準の SNP が用いられるため、検出された変異から SNP を除去したものが体細胞変異 とみなされる。

従って、ドライバー遺伝子変異の評価に際して、対応する人種の SNP 情報を用いない場合、偽陽性(日本人に固有の SNP が除去されずドライバー遺伝子変異として同定される)ならびに偽陰性(日本人に特有の病的意義がある変異を、日本人以外で SNP であることからドライバー遺伝子変異リストから除外してしまう)の可能性が高まるものと考えられる。後者は、腫瘍検体のみを用いた体細胞変異解析において特に注意すべき点である。以上のことは、日本人のがんゲノム解析における日本人 SNP 情報の重要性を明確に

示している。本研究では、腫瘍特異的体細胞変異解析においては、腫瘍組織と正常組織を対にした解析を 行うことで、また生殖細胞系列の解析においては、日本人における変異頻度情報データベースである iJGVD<sup>50</sup>と HGVD<sup>51</sup>を統合することで、上記問題への対応を試みた。

本研究で明らかとなった主要な所見を図16に要約した。体細胞ドライバー変化は、全体の72.2%にあた る2982 検体において、1 検体あたり平均2.3 個検出された(図9)。1 検体あたりのドライバー変化数の 最小は1 個であり1120 検体で認められ、複数の変化が1862 検体で検出され、最大は15 個の変化であっ た。これらのドライバー遺伝子変化は、362 種類のがんドライバー遺伝子において同定されており、ドラ イバー遺伝子変化の数、影響を受ける遺伝子、遺伝子変異・発現異常・融合遺伝子の割合は、がんの種類 によって異なっていた。個々の症例において、ドライバー遺伝子変化とドラッガブル遺伝子・ゲノム変化 の双方ないしは一方の変化のみが認められた検体には次のような特徴が見られた。

- ドライバー変化とドラッガブル変化双方の検出率が高い(GIST)
- 80%以上の検体でドライバー変化が同定された一方、そのほとんどが分子標的薬とは関連付かない (結腸がん、直腸がん、子宮がんなど)
- ドライバー変化の検出率が70%を超えており、その半数が分子標的薬の標的である(肺腺がん、黒
  色腫)
- その他

のようにがん種間で異なっていた(図17)。





全がん種を対象とした腫瘍遺伝子変異総量の解析(図6)により、全体の5.4%にあたる292 検体が高頻 度変異症例(腫瘍遺伝子変異総量が20変異/Mb以上)に分類された。変異シグネチャの解析から、高頻度 変異症例の要因として、DNA ミスマッチ修復の欠損(44.4%)、喫煙(10.9%)、POLE 異常(8.9%)、紫 外線照射への暴露(2.1%)が同定された(図7)。ミスマッチ修復欠損シグネチャを有する高頻度変異症例 は免疫チェックポイント阻害剤の対象となり得る。これらの結果は、全エクソンシークエンスによって得 られる腫瘍遺伝子変異総量と変異シグネチャの情報に基づいて、免疫チェックポイント阻害剤の対象とな る症例を特異的に選択することが可能であることを示している。また、「高腫瘍遺伝子変異総量」という情 報は、遺伝子変化の結果解釈を慎重にすべき検体を選別する際の指標としても有用である。このようなケ ースでは、腫瘍特異的遺伝子変異が多数、発生し、がんのドライバー遺伝子にも変異が多数、認められ、 細胞がん化の原因遺伝子を特定することが困難な場合がある。

ゲノム情報に基づいた分子標的薬と免疫チェックポイント阻害剤の選択は、がんゲノム医療における重要な知見の1つである。現在の解析方法に基づき有効薬剤の選別がどの程度、可能かを推定するために、 4131 検体におけるドラッガブルな変化を評価した。その結果、11.3%(467 検体)で現在臨床適用されている薬剤の情報が同定された。その内訳は、分子標的薬が全体の8.7%(360 検体)、免疫チェックポイント阻害剤が全体の2.6%(107 検体)であった。高頻度に検出されるドラッガブルな変化は、がん種ならびに遺伝子特異的であり、肺腺がんでは EGFR、黒色腫では BRAF、GIST では KIT に集中していた(図15)。 一方で、がん種ごとのドラッガブル変化の検出率に大きな差が見られた。これは各がん種に対し、現時点で利用可能な薬剤が限定的であることを反映していると考えられる。

血液検体を用いた生殖細胞系列の解析を通して、遺伝性腫瘍、非がん性遺伝性疾患、クローン性造血の ドライバー変異を評価した。遺伝性腫瘍の原因とされる病原性変異は全体の 9.2% (279/3022) で同定され た。そのうち 12.2% (全体の 1.1%) は病原性変異の存在が別な解析法でも確認され、さらに 6.1% (全体 の 0.6%)は家族歴などの臨床情報に基づき遺伝性腫瘍と診断された(表 2)。これらの患者とその家族に ついては予防的手術ならびに関連するがんの早期検出の可能性が示唆される。

非がん性遺伝性疾患に関しては、全体の 1.1% (33/3022) で、米国臨床遺伝学会のリストにある 11 疾患 の病原性変異が同定され、0.4% (11/3022) が非がん性遺伝性疾患と診断された(表 3)。

クローン性造血は全体の 8.4% (316/3751) で検出された。クローン性造血に関する情報は、臨床上、骨 髄異形成症候群や白血病への進展のリスク評価に有用であると期待される。一方で、リキッドバイオプシ ーの結果解釈へ大きな影響を及ぼすものと考えられる。これらの遺伝子変異が、循環腫瘍細胞またはクロ ーン性造血のいずれに由来するかを確定するためには、今後さらに注意深く検査していく必要があると考 えられる。

がんドライバー遺伝子変化のうち、体細胞変異について本研究の結果を TCGA コホートから抽出した白 人コホートと比較した場合、特定のがん種においてドライバー遺伝子変化の頻度に明確な差が認められた (図 13)。例えば、日本人肺腺がんにおける高頻度の EGFR変異と低頻度の KRAS変異などである。これ らは、変異のホットスポットレベルでの解析結果により日本人の EGFR p.E746\_A750del と p.L858R、白 人集団における KRAS p.G12C からも確認することができた。他のがん種においても、さらなる分析が現 在進行している。また、本研究の HOPE コホートでは、がん種による解析対象の絞込みを行わず全ての検 体を分析しているため、希少がんにおける遺伝子・ゲノム変化も検出されている。その結果、GIST におけ る KIT、皮膚がんにおける HRAS、十二指腸がんにおける KRAS など、白人集団では稀ながんにおけるド ライバー遺伝子変異の存在が確認された。これらを除くと、がんゲノムアトラスの一般的な特徴は、日本 人と白人集団で類似していることも示された。

本研究の結果は、将来のがんゲノム研究に対して新たな展望を与えている。中でも、ドライバー遺伝子 変化が検出されなかった症例の解析は重要な課題である。今回の解析では、27.8%の検体でドライバー遺伝 子変化が検出されなかった。このような症例における、がん化の原因変化を特定するためには、全ゲノム 解析による大規模な構造変化やエピジェネティックな変化など、さらなる解析が必要である。遺伝性腫瘍 に対するドライバー遺伝子変化の評価にもさらなる研究が必要である。全体の 9.2%で遺伝性腫瘍の病原性 変異が同定された一方で、遺伝性腫瘍と診断されたのはこのうちの 6.1%であった。残りの症例には、何ら かの理由で病原性変異が存在するにもかかわらず発症しなかった症例、家族調査が困難な症例、遺伝性腫 瘍の発端者であり家族歴が明確でない症例、*検出された変異が病原性を有していない場合*などが含まれる と考えられる。プロジェクト HOPE は単一施設で実施されており、全ての臨床データとゲノム解析結果を 突合できることから、今後の研究によって、病原性の遺伝子変化と遺伝性腫瘍発症との関連の解明が進む ものと期待される。

以上をまとめると、プロジェクト HOPE の成果として確立された日本版がんゲノムアトラス JCGA は、 日本人がんゲノム研究の基盤資源であり、日本人がん患者に対するゲノム診療実践における有用な情報と なり得る。本研究で構築された基盤情報を利用した患者転帰と薬物療法歴の統合解析は、将来の個別化治 療と薬剤開発をより加速していくものと期待される。

5. 謝辞

検体収集に協力してくれた静岡がんセンター病院のスタッフと、意義深い議論をしてくれた静岡がんセンター研究所のメンバーに感謝する。

6. ディスクロージャー・ステートメント 著者は利害の対立がない。

# 7. 参考文献

- 1. Bailey MH, Tokheim C, Porta-Pardo E, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018;173:371-385.
- Ding L, Bailey MH, Porta-Pardo E, et al. Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. Cell. 2018;173:305-320.
- 3. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature. 2013;500:415-421.
- 4. Zehir A, Benayed R, Shah RH, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med. 2017;23:703-713.
- Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet. 2019;51:202-206.
- 6. Huang KL, Mashl RJ, Wu Y, et al. Pathogenic Germline Variants in 10,389 Adult Cancers. Cell. 2018;173:355-370.
- Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nat Med. 2014;20:1472-1478.
- Nagahashi N, Ling W, Hayashida T, et al. Actionable Gene Alterations in an Asian Population With Triple-Negative Breast Cancer. JCO Precis Oncol. 2018;2:1-13.
- Sato S, Nagahashi M, Koike T, et al. Impact of Concurrent Genomic Alterations Detected by Comprehensive Genomic Sequencing on Clinical Outcomes in East-Asian Patients with EGFR-Mutated Lung Adenocarcinoma. Sci Rep. 2018;8:1005.
- 10. Jia F, Teer JK, Knepper TC, et al. Discordance of Somatic Mutations Between Asian and Caucasian Patient Populations with Gastric Cancer. Mol Diagn Ther. 2017;21:179-185.
- The 1000 Genomes Project Consortium, et al. A global reference for human genetic variation. Nature. 2015;526:68-74.
- 12. Yamaguchi K, Urakami K, Ohshima K, et al. Implementation of individualized medicine for cancer patients by multiomics-based analyses -the Project HOPE-. Biomed Res. 2014;35:407-412.
- Ministry of Health, Labour and Welfare. Japanese ethical guidelines for human genome/gene analysis research. 2017. https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hokabunya/kenkyujigyou/i-kenkyu/index.html. Accessed September 17, 2019.
- 14. Shimoda Y, Nagashima T, Urakami K, et al. Integrated next-generation sequencing analysis of whole exome and 409 cancer-related genes. Biomed Res. 2016;37:367-379.
- 15. Urakami K, Shimoda Y, Ohshima K, et al. Next generation sequencing approach for detecting 491 fusion genes from human cancer. Biomed Res. 2016;37:51-62.
- Ohshima K, Hatakeyama K, Nagashima T, et al. Integrated analysis of gene expression and copy number identified potential cancer driver genes with amplification-dependent overexpression in 1,454 solid tumors. Sci Rep. 2017;7:641.
- 17. Nagashima T, Shimoda Y, Tanabe T, et al. Optimizing an ion semiconductor sequencing data analysis method to identify somatic mutations in the genomes of cancer cells in clinical tissue samples. Biomed Res. 2016;37:359-366.
- 18. Tan A, Abecasis GR, Kang HM. Unified representation of genetic variants. Bioinformatics. 2015;31:2202-2204.
- Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012;6:80-92.
- 20. O'Leary NA, Wright MW, Brister JR, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44:D733-D745.
- 21. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47:D506-D515.
- Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol. 2016;17:31.
- 23. COSMIC Mutational Signatures (v2). 2015. https://cancer.sanger.ac.uk/cosmic/signatures\_v2. Accessed

September 17, 2019.

- Zhang Z, Hao K. SAAS-CNV: A Joint Segmentation Approach on Aggregated and Allele Specific Signals for the Identification of Somatic Copy Number Alterations with Next-Generation Sequencing Data. PLoS Comput Biol. 2015;11:e1004618.
- 25. Shen R, Seshan VE. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. Nucleic Acids Res. 2016;44:e131.
- 26. Favero F, Joshi T, Marquard AM, et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Ann Oncol. 2015;26:64-70.
- 27. Liu Q, Hu Q, Yao S, et al. SeqSQC: A Bioconductor Package for Evaluating the Sample Quality of Next-generation Sequencing Data. Genomics Proteomics Bioinformatics. 2019;17:211-218.
- Tamborero D, Rubio-Perez C, Deu-Pons J, et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. Genome Med. 2018;10:25.
- Griffith M, Spies NC, Krysiak K, et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. Nat Genet. 2017;49:170-174.
- 30. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res. 2017;45:D777-D783.
- 31. Sun SQ, Mashl RJ, Sengupta S, et al. Database of evidence for precision oncology portal. Bioinformatics. 2018;34:4315-4317.
- 32. Ainscough BJ, Griffith M, Coffman AC, et al. DoCM: a database of curated mutations in cancer. Nat Methods. 2016;13:806-807.
- Dienstmann R, Jang IS, Bot B, Friend S, Guinney J. Database of genomic biomarkers for cancer drugs and clinical targetability in solid tumors. Cancer Discov. 2015;5:118-123.
- 34. Bouaoun L, Sonkin D, Ardin M, et al. TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data. Hum Mutat. 2016;37:865-876.
- Chakravarty D, Gao J, Phillips SM, et al. OncoKB: A Precision Oncology Knowledge Base. JCO Precis Oncol. 2017 Jul;2017. Epub 2017 May 16.
- Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016;44:D862-D868.
- Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, Cooper DN. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. Curr Protoc Bioinformatics. 2012;39:1-13.
- National Cancer Center, Center for Cancer Genomics and Advanced Therapeutics. https://www.ncc.go.jp/en/c\_cat/. Accessed September 17, 2019.
- Sondka Z, Bamford S, Cole CG, Ward SA, Dunham I, Forbes SA. The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat Rev Cancer. 2018;18:696-705.
- 40. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013;339:1546-1558.
- Sanchez-Vega F, Mina M, Armenia J, et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell. 2018;173:321-337.
- 42. Hovelson DH, McDaniel AS, Cani AK, et al. Development and validation of a scalable next-generation sequencing system for assessing relevant somatic variants in solid tumors. Neoplasia. 2015;17:385-399.
- 43. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45:D353-D361.
- Tamaki T, Dong Y, Ohno Y, Sobue T, Nishimoto H, Shibata A. The burden of rare cancer in Japan: application of the RARECARE definition. Cancer Epidemiol. 2014;38:490-495.
- 45. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med. 2017;19:249-255.
- 46. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. Science 2019 366, eaan4673.

- 47. Coombs CC, Zehir A, Devlin SM, et al. Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. Cell Stem Cell. 2017;21:374-382.
- 48. Hatakeyama K, Ohshima K, Nagashima T, et al. Molecular profiling and sequential somatic mutation shift in hypermutator tumours harbouring POLE mutations. Sci Rep. 2018;8:8700.
- 49. Henn BM, Botigue LR, Bustamante CD, Clark AG, Gravel S. Estimating the mutation load in human genomes. Nat Rev Genet. 2015;16:333-343.
- 50. Tadaka S, Katsuoka F, Ueki M, et al. 3.5KJPNv2: an allele frequency panel of 3552 Japanese individuals including the X chromosome. Hum Genome Var. 2019;6:28.
- 51. Higasa K, Miyake N, Yoshimura J, et al. Human genetic variation database, a reference database of genetic variations in the Japanese population. J Hum Genet. 2016;61:547-553.