



システム癌  
新次元

## がんシステムの新次元俯瞰と攻略

文部科学省科学研究費補助金 新学術領域研究 (研究領域提案型)(複合領域:4701)

Newsletter No. **13**

2017年6月

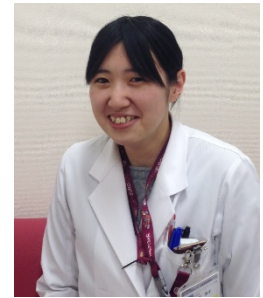
CONQUERING CANCER THROUGH  
NEO-DIMENSIONAL SYSTEMS  
UNDERSTANDING

公募研究

# ナノポアシーケンサーによるがん細胞の変異検出およびフェーズ情報解析手法の確立

公募研究 A02-4-16 研究代表者

国立がん研究センター先端医療開発センター 研究員 鈴木 絢子



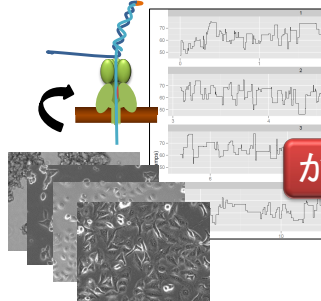
がんは遺伝子変異の病気であり、実際にかん細胞のゲノムには多数の異常が生じています。近年の大規模シーケンス技術の発展に伴って、それぞれのがんではどのような異常が生じているのかを明らかにするために、世界的に多種多様ながんゲノムのシーケンス解析が進められてきました。特に肺がんでは、がんの発生・進展に直接寄与する遺伝子異常(ドライバー変異)のカタログ化が進んでおり、

EGFR や ALK 遺伝子といった原因遺伝子をターゲットにした分子標的薬の開発がおこなわれています。しかし一方で、症例によってはゲノム解析を行ったとしても、ドライバー変異が不明、もしくは、ドライバー変異が分かってもそれに適した分子標的薬がない、などといった問題点も挙げられています。実際に、日本人肺腺がん患者の約 2~3 割は明確なドライバー遺伝子がわかっていません(Kohno et al. 2013

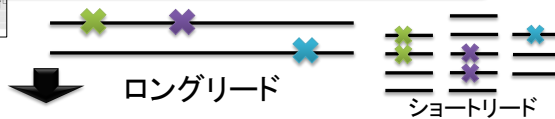
Cancer Science)。大規模シーケンサーによってがんゲノムからたくさんの変異を見つけることができましたが、解析の仕方によっては見逃されてしまっているゲノム異常があるとも考えられます。今後もさまざまな角度からがんゲノムを再評価する必要があると考えられます。

## ナノポアシーケンサー MinION

- ロングリード(>10kb)
- 安価なシーケンサー
- 手のひらサイズ
- 短時間(~48h)



がんゲノムのロングリードデータ



同一ゲノムDNAコピー上にある変異・多型の相互関係を維持したままゲノム解析

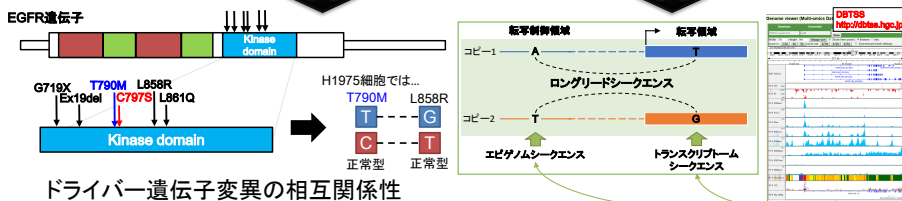


図1: ロングリードシーケンサー MinION とがんゲノムフェージング解析

がん現在、ゲノム解析に使用されている多くの大規模シーケンサーは、数十～数百塩基対の短い DNA 断片の配列（ショートリード）を解読して大量のデータを出力します。これらデータをスーパーコンピュータなどの大規模コンピュータシステムを用いて情報解析を行うことで、がんゲノムに特異的に起きている変異が数多く同定されてきました。しかし、これらの短い配列データから、本来は長い DNA 分子上に生じている異常を、相互の関係性（フェーズ情報）を保持したまま検出するのは難しいと考えられます。そこで本研究では、10kb を超える長い DNA 配列（ロングリード）を解析することが可能なナノポアシーケンサー MinION を用いて、肺がんゲノムをより詳細に解析するための情報解析手法を確立することにしました（図 1）。MinION は、DNA 分子がナノポアチャネルを通過する際の電気シグナルのパターンから塩基配列を読み取ります。ロングリードが解析できることもさることながら、初期投資が安価でポータブル性が高く、短時間でデータを得られることから、新しいゲノム解析技術として注目されています。データ産出量と精度は従来型のショートリードシーケンサーのほうが高いといわれていますが、MinION もここ 2 年ほどで飛躍的に改良されてきました。

ロングリードシーケンサーを用いてがんゲノムを解析することで何を知ることができるでしょうか。ヒトの細胞にはゲノム DNA が 2 コピーあり、それぞれ父方と母方に由来するとされています。しかしがん細胞ではゲノムの不安定性がおき、遺伝子のコピー数が増加・減少していたり、染色体異常が見られたりします。つまり細胞、特にがん細胞がもつ DNA 配列は一

本ではありません。ゲノム変異を解析する際に、ロングリードで一氣に長い DNA 断片を読み取ることができれば、どの変異が同じ DNA 上に生じていたのかが比較的簡単にわかります。例えば、日本人肺腺がんでは EGFR 遺伝子に変異が頻出し、この変異をもつがん細胞には EGFR 遺伝子に対する分子標的薬が効果を示すことが知られています。しかし、治療を進めると薬剤に対して耐性を示すようになることがあり、耐性を示した症例の約半数が同じ EGFR 遺伝子上の異なる場所に新しく耐性変異を獲得していたと報告されています（Camidge et al. 2014 Nature Reviews Clinical Oncology）。これら耐性を示すがんの効果のある薬剤を開発してもまた次の耐性変異が出現することがありますが、変異の組み合わせによっては以前耐性を示した薬剤に再度感受性を示すこともあります（Shaw et al. 2016 N Engl J Med）。がんが進化していく過程でドライバー遺伝子上に蓄積していく複数の変異が同一 DNA コピー上に存在するのか、がん組織の中にあるたくさんのがん細胞の中でどの程度の頻度で存在するのか解析することは分子標的薬耐性獲得メカニズムの解明や耐性を示すがんに対する次の治療選択・開発への有用な情報となるのではないかと考えています。

また、タンパク質コード領域の変異ではなく、いわゆるノンコーディング領域の変異を解析していく上でもロングリードシーケンス技術は役立ちます。いくつかのがん種ではプロモーターやエンハンサーといった遺伝子の転写を制御する領域に変異があることがすでに報告されており、例として TERT 遺伝子のプロモーター領域におけるホットスポット変異があげられます

（Huang et al. 2013 Science）。実際に私たちが解析している肺腺がん培養細胞でも、このプロモーター変異がみられるものがありました。少なくとも 2 コピー以上あるゲノム DNA のうちの片方（アリル）にのみこの変異が見られ、エピゲノムやトランスクリプトームのデータを組み合わせたと、変異アリルが特異的に転写活性化され、発現していることがわかりました。転写制御領域の変異はがんゲノムにおいてこのほかにも数多く検出されるのですが、実際の遺伝子発現・転写制御にどの程度影響するのかわからないことが多いです。そこで、肺がん培養細胞をモデルに、制御領域の変異と同じゲノムコピー上にある下流の転写領域がどの程度遺伝子発現しているかどうかを明らかにすることによって、転写制御領域の変異の影響を解析することも現在行っています。肺がん培養細胞のゲノム、エピゲノムおよびトランスクリプトームといった大規模ショートリードデータは、連携研究者である東大の鈴木穰教授、中井謙太教授が構築・管理するデータベース DBTSS および DBKERO (<http://dbtss.hgc.jp/>) からすでに公開されています。将来的にはここにロングリードシーケンスデータを新たに加えることで、肺がんゲノムの転写制御異常をシングルアリルレベルでゲノムワイドに記載していくことも目指しています。

ゲノム科学、生物情報学、および腫瘍学の研究者の協力を得て、ゲノム生物学一般に対する技術基盤の確立、がんゲノムに対する新規概念の創出および実際の治療開発・診断への応用と、多方面にわたり研究を推進していく予定です。

## 公募研究

# 難治性肺がんに対する術後再発リスクや治療応答性に関わる HLA アレルの同定

公募研究 A02-5-16 研究代表者

国立がん研究センター研究所 ユニット長 白石 航也



本邦における肺がんは男女ともがんに死因上位を占める難治がんであり、罹患数は年々増加傾向にあります。近年のゲノム研究の進歩により、ある一部のがん遺伝子の内、がんの発生・進展において直接的に重要な役割を果たすドライバー遺伝子(例えば、EGFR(上皮増殖因子受容体)変異、ALK 融合、ROS 融合、RET 融合遺伝子等)の体細胞変異を治療標的とした複数の薬剤が開発されてきました。しかしながら多くの患者さんは、一時的にがんの進行を押さえられますが、再び増悪することが知られており、進行性肺がんに対する根治はきわめて難しいのが現状です。そのため、治療効果が得られない原因を解明することが急務となっています。また、進行性肺がんの 5 年生存率は

20%以下であるのに対して、早期肺がんの 5 年生存率は 70%以上であることから、肺がんに対する術後再発高危険度群を捕捉し、術後再発を抑制することができれば、肺がん死抑制のための有益な手段となりうると考えられます。我々は、肺腺がんのなかでも日本人に多い EGFR 変異陽性肺腺がんについて、罹りやすさを決める遺伝子領域を明らかにしました。その中には、免疫反応の個人差の原因となる HLA(Human Leukocyte Antigen:ヒト白血球抗原)クラス II 遺伝子領域が含まれていました(Shiraiishi et al., Nat Commun. 2016; 7:12451)。特に HLA クラス II 遺伝子産物のうちの一つである HLA-DPB1 タンパク質の 57 番目のアミノ酸の置換を起こす多型が、EGFR 変異陽性の肺腺がんの

罹りやすさを決める原因多型のひとつと考えられます。HLA 遺伝子群の個人差は臓器移植における適合性など、免疫反応の個人差の原因となるものです。また、その個人差の分布は、人種によって大きく異なっています。よって、EGFR 遺伝子に変異を起こした細胞に対する免疫反応の違いなど、いくつかの遺伝子の個人差による生体反応の個人差が、EGFR 変異陽性肺腺がんへの罹りやすさを決めていると考えられます。これらの研究成果は、悪性度の高いがん細胞が生き残るために必要ながん組織中で認められる遺伝子異常の他に、自己と非自己を認識するために必要な HLA 遺伝子の多様性が、がん免疫において重要な役割を果たしている可能性を示しています。

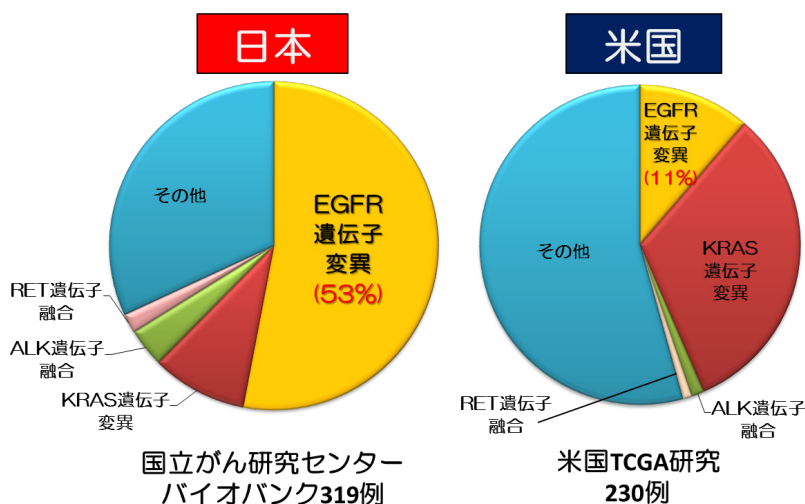


図1: 日本人と欧米人の肺腺がんに生じるがん細胞由来の体細胞変異の違い

EGFR 体細胞変異は日本人の肺腺がんの 50% に生じ、欧米人よりもはるかに高頻度です。

一方で、HLA 領域は非常に多様性に富むため、ゲノム解析は積極的に行われてきませんでした。しかし最近のゲノム解析技術の進歩により、日本人集団における HLA 領域のレファレンスパネルが構築され、それらを基に HLA imputation 法が開発されました(A02-1 計画研究代表者: 岡田随象教授の研究成果)。この解析技術を用いることで、既存の SNP チップのデータがあれば、HLA アレルの推定が可能になりました(Okada et al., Nat Genet. 2015;

47(7): p798-802.)。本研究では、既取得 SNP チップデータに基づき、HLA アレルを含む一塩基多型(SNP)を推定し、術後再発や治療応答性と関連する HLA アレルの同定を目指します。

昨年度は術後再発リスクや治療応答性との関連を検討するため、本研究に必要な症例数の確保を進めてきました。検出研究を拡張するため、一部検証研究に用いるサンプルより、Illumina OmniExpressExsome SNP chip を用いてタ

イピングを行いました。今年度は、これらの解析結果を統合して解析を進めてまいります。さらに強い関連が認められた HLA アレルについては、検証研究を用いて真の感受性遺伝子を同定する予定です。これらの成果は、HLA の多様性を含む胚細胞系列変異とがん細胞由来の体細胞変異情報を統合することで、がんの本態解明や効率的な治療法の開発に繋げられると考えられます。

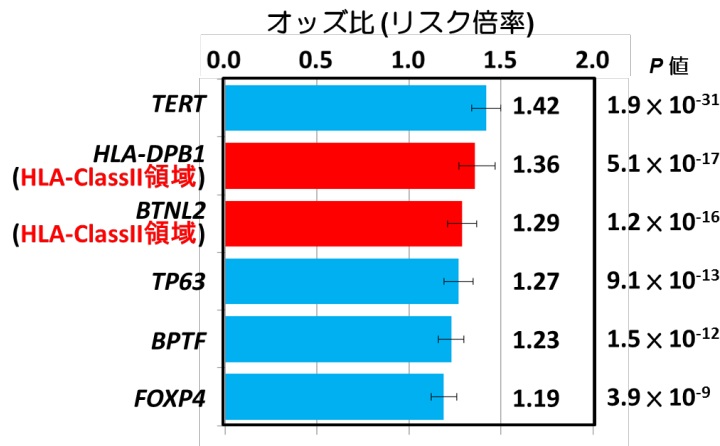


図2:

HLA の多様性が EGFR 変異陽性肺腺がんの発症要因に関わるある遺伝子の多型を持たない場合に対して、多型をひとつ持つ場合のリスクを倍率(オッズ比±95%信頼区間)で示しています。人は父、母由来の二つの遺伝子を受け継ぐため、両方が危険型の人の場合、オッズ比の 2 乗の危険度となります。例えば、HLA-DPB1 遺伝子の多型の危険型をひとつ持つと 1.36 倍、ふたつ持つと 1.9 倍 ( $1.36^2=1.85$ ) 肺腺がんにかかりやすくなると推定されます。

## Information



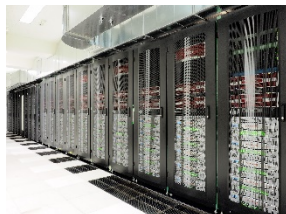
Human Genome Center

### 東京大学医科学研究所 ヒトゲノム解析センタースーパーコンピュータ

システム癌新次元で用いられるデータ解析ソフトウェアは、ヒトゲノム解析センターのスーパーコンピュータで動いています。スーパーコンピュータはどなたでもご利用になれます(有償)。



大容量のディスク装置



高密度の計算機

<http://supcom.hgc.jp>



システム癌  
新次元

<http://neosystemscancer.hgc.jp/>

新学術領域研究「システム癌新次元」ニュースレター No.13

発行日★平成29年6月6日(初版)

発行★がんシステムの新次元俯瞰と攻略

領域代表者★宮野 悟

- 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
- 〒108-8639 東京都港区白金台4-6-1
- TEL: 03-5449-5615 FAX: 03-5449-5442
- E-mail: miyanolab-jimu@edelweiss.hgc.jp