

厚生労働科学研究委託費（革新的がん医療実用化研究事業）
委託業務成果報告（総括）

クリニカルシーケンスのための
実用的なバイオフィーマティクスプログラムの
開発および情報解析に関する研究

業務主任者又は担当責任者 加藤 護 | 国立がん研究センター研究所部門長

研究要旨

本研究では、臨床で使われる低腫瘍率かつFFPE 用の変異検出アルゴリズムを開発し、国立がん研究センターの臨床シーケンスに適用し、その結果をもって、日本人のがん変異に関する特徴を調べることである。

業務項目の担当責任者氏名・所属研究機関名及び所属研究機関における職名

国立がん研究センター研究所
部門長 加藤護

国立がん研究センター研究所
分野長 河野隆志

A . 研究目標

近年、国際がんゲノム・コンソーシアムといった大規模プロジェクトによって、様々ながんにおける体細胞変異が網羅的に同定されつつある。国際がんゲノム・コンソーシアムにおいては申請者も情報解析に参加し、次世代シーケンサーを使って肝がんの体細胞変異とその性質を解明しつつある(Totoki et al, 2011, Nature Genetics; Fujimoto et al, 2012, Nature Genetics; Totoki et al, 2014, Nature Genetics)。

これらの成果を臨床に応用する方法として、現在臨床シーケンスが実現されつつある。臨床シーケンスとは、患者のがん組織試料を次世代シーケンサーによって配列決定し、検出されたがん変異に対する作用薬の知識と突き合わせて、より適切な治療法や臨床試験のよりよい層別化を行うことである。アメリカのベンチャー企業Foundation Medicine社は既に商用運用を開始し、研究開発の成果をNature Biotechnology誌(Frampton et al, 2013)に報告している。国立がん研究センターでは、2013年度より臨床シーケンスの実施基盤を独自に開発し、その運用が開始されつつある。

臨床シーケンスは、次世代シーケンサーによる各個人の配列決定であり膨大なデータを生む。膨大なデ

ータを処理するにはコンピュータによる情報解析が不可欠であるが、研究用に開発されたアルゴリズムを臨床シーケンスにそのまま適用は出来ない。臨床用の情報解析には、研究用の情報解析では現れなかった問題が出てくるからである。申請者はこれまで、当センターの臨床シーケンス用に点変異・融合遺伝子検出のアルゴリズムを開発してきたが、これは研究用で使われる高腫瘍率、かつ、質の良い凍結試料をベースに開発されたものである。臨床で使われる低腫瘍率、かつ、質の悪いFFPE 試料での検討はまだ不十分である。また、コピー数異常に関しては、FFPE 試料に対する検出のアルゴリズムさえ定まっていない。さらに、アルゴリズムを適用した結果を整理して、日本人の診断に役立つ情報を抽出する分析も不十分である。

本研究の目的は、臨床で使われる低腫瘍率かつFFPE 用の点変異・融合遺伝子・コピー数異常検出のアルゴリズム群を開発し（平成26-27年度）、臨床シーケンスに適用して日本人がん変異のデータベースを作成し（平成27-28年度）、日本人の変異に関する情報解析を行うこと（H28年度）である。

B . 研究方法

研究協力者である国立がん研究センター病院の山本昇医師、田村研治医師がFFPE の患者臨床試料を取得し、同病院の担当病理医が腫瘍率推定を含む病理組織の評価を行う。分担研究者である河野隆志博士、研究協力者である市川仁博士が試料のDNAをキャプチャし、次世代シーケンサーによって配列決定する。2週間当たり約5名の患者の試料が、この流れでシーケンスされた。

配列決定されたデータ（fastq）を基に、低腫瘍率かつFFPE 用のアルゴリズム開発およびその情報解析を行う。

厚生労働科学研究委託費（革新的がん医療実用化研究事業）
委託業務成果報告（総括）

本研究は国立がんセンター倫理審査委員会にてすでに承認が得られている。試料採取に関しては、病理診断・検査の残余を研究に用いるため、提供者に新たに侵襲を与えず、また診断への影響や治療への介入はない。臨床試料の提供者には、試料が医学研究に使われることを文書および口頭で説明し、個別に同意を取得する。臨床試料は連結可能匿名化を行い、個人が特定されないようにして解析に用いる。

C. 研究成果

・FFPE 試料の収集およびシーケンス

国立がん研究センター病院の医師の協力の下、サンプルが収集され、病理医による確認、DNA 抽出と品質チェックを経て、約100 サンプルが次世代シーケンサーによって、シーケンスされた。

・点突然変異（SNV/indel）

研究用の場合、凍結サンプルというコストは高いが質は良いサンプルを使う。しかし臨床シーケンスの場合、低コストだがDNA が変性してしまうFFPE サンプルを使う。FFPE サンプルを使うとデータにエラーが多く、ミスアライメントを引き起こしやすい。また、臨床応用においては低腫瘍率のサンプルも使わざるを得ない。腫瘍率が低いサンプルでは当然感度が落ちる。

これらの問題に対処するために、今回ミスアライメントを特異的に除外するフィルターなど、11種類のフィルターを開発した。これらのフィルターにおいては、様々な種類のデータに対し頑健に対応するため、統計的手法が最大限利用されている。

この検出アルゴリズムを、上述したサンプルに適用し、変異を検出した。IGV やさまざまな統計数値を調べ、アルゴリズムの改良を行った。次に、検出された変異を質量分析法（MassArray）

によって検証した。その結果、SNV/indel 共に、非常な高精度で検出されていることを確認した。

・融合遺伝子検出

融合遺伝子検出では、これまで開発したsingle-end sequence read ベースのアルゴリズムに、pair-end sequence read ベースのアルゴリズムを組み合わせ、感度を高める改良を行っている。

・コピー数変化検出

COSMIC と DGV データベースを調べ、既知のがんでコピー数変動がない領域、日本人において生殖細胞系列でコピー数変動がない領域を選び出し、この領域をコピー数の基準領域とするアルゴリズムを開発した。また、データからコピー数変化領域を自動的に決定するsegmentation アルゴリズムを開発している。

D. 考察

FFPE 試料のシーケンスのさい、中には解析できないほどDNA の質が悪いものがあることが分かった。これらは現在、DNA の質を実際に測定することで解決している。点突然変異検出アルゴリズムは、様々な改良と確認実験を経て、実用レベルに達したと言える。

E. 結論

本年度の研究計画通り、FFPE試料のDNAを実際に約100 検体ほどシーケンスし、点変異検出アルゴリズムを実用化した。しかし今後も精度改良に努めていく予定である。今後は、融合遺伝子検出、および、コピー数変異検出アルゴリズムの改良に重点を移していく。これらの確認実験も行う。

F. 健康危険情報 特になし

1. 論文発表

（次のページに続く）

厚生労働科学研究委託費（革新的がん医療実用化研究事業）
委託業務成果報告（総括）

Yasushi Totoki, Kenji Tatsuno, Kyle R Covington, Hiroki Ueda, Chad J Creighton, Mamoru Kato, Shingo Tsuji, Lawrence A Donehower, Betty L Slagle, Hiromi Nakamura, Shogo Yamamoto, Eve Shinbrot, Natsuko Hama, Megan Lehmkuhl, Fumie Hosoda, Yasuhito Arai, Kim Walker, Mahmoud Dahdouli, Kengo Gotoh, Gen-ta Nagae, Marie-Claude Gingras, Donna M Muzny, Hidenori Ojima, Kazuaki Shimada, Yutaka Midorikawa, John A Goss, Ronald Cotton, Akimasa Hayashi, Junji Shibahara, Shumpei Ishikawa, Jacfranz Guiteau, Mariko Tanaka, Tomoko Urushidate, Shoko Ohashi, Naoko Okada, Harsha Doddapaneni, Min Wang, Yiming Zhu, Huyen Dinh, Takuji Okusaka, Norihiro Kokudo, Tomoo Kosuge, Tada-toshi Takayama, Masashi Fukayama, Richard A Gibbs, David A Wheeler, Hiroyuki Aburatani & Tatsuhiro Shibata

Trans-ancestry mutational landscape of hepatocellular carcinoma genomes.

Nature Genetics, 2014, 46, 1267-1273.

Tatsuji Mizukami, Kouya Shiraishi, Yoko Shimada, Hideaki Ogiwara, Koji Tsuta, Hitoshi Ichikawa, Hiromi Sakamoto, Mamoru Kato, Tatsuhiro Shibata, Takashi Nakano, Takashi Kohno,

Molecular mechanisms underlying oncogenic RET fusion in lung adenocarcinoma.

Journal of Thoracic Oncology, 2014, 9, 622-630.

加藤 護. 総説:「一細胞ゲノム解析」医学の歩み、2014, 249, 1088-1092. (2014)

翻訳: 加藤 護、「発がんドライバー変異の同定」 by David Tamborero, Abel Gonzalez-Per ez and Nuria Lopez-Bigas, 実験医学、2014, 32, 213-219

Nakaoku T, Tsuta K, Ichikawa H, Shiraishi K, Sakamoto H, Enari M, Furuta K, Shimada Y, Ogiwara H, Watanabe SI, Nokihara H, Yasuda K, Hiramoto M, Nammo T, Ishigame T, Schetter AJ, Okayama H, Harris CC, Kim YH, Mishima M, Yokota J, Yoshida T, Kohno T* (2014) Druggable oncogene fusions in invasive mucinous lung adenocarcinoma. *Clin Cancer Res.* 20(12):3087-3093.

Saito M, Ishigame T, Tsuta K, Kumamoto K, Imai T, Kohno T* (2014) A mouse model of KIF5B-RET fusion-dependent lung tumorigenesis. *Carcinogenesis.* 35(11):2452-2456

Saito M, Shiraishi K, Matsumoto K, Schetter AJ, Ogata-Kawata H, Tsuchiya N, Kunitoh H, Nokihara H, Watanabe SI, Tsuta K, Kumamoto K, Takenoshita S, Yokota J, Harris CC, Kohno T*(2014) A Three-microRNA signature predicts responses to platinum-Based doublet chemotherapy in patients with lung adenocarcinoma. *Clin Cancer Res.* 20(18); 4784-93.

2. 学会発表

加藤 護. A bioinformatics system developed for cancer clinical sequencing in National Cancer Center, Japan European Society for Medical Oncology, 招待講演, Madrid, Spain, September 26-30, 2014.

加藤 護. Development of a computational system in cancer clinical sequencing. 招待講演, 2014 Cancer Symposium, Mokpo, Korea, April 15-19, 2014.

加藤 護. A bioinformatics system developed for clinical sequencing in National Cancer Center, Japan (Symposia) 73rd Annual Meeting of the Japanese Cancer Association, Yokohama, Japan, September 25-27, 2014.

厚生労働科学研究委託費（革新的がん医療実用化研究事業）
委託業務成果報告（総括）

中奥敬史、蔦幸治、渡邊俊一、
軒原浩、金永学、三嶋理晃、横田
淳、河野隆志：Lung invasive mu-
cinous adenocarcinoma(IMA)にお
ける治療標的となる新規遺伝子融
合、第55回日本肺癌学会学術総会、
京都、第55回日本肺癌学会学術総
会抄録集、P116、11月、2014年

H . 知的財産権の出願・登録状況
（予定を含む。）

