

科学研究費助成事業 研究成果報告書

平成 27 年 9 月 25 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2013～2014

課題番号：25830139

研究課題名(和文) がんの新規原因遺伝子特定に向けた変異解析パイプラインの構築

研究課題名(英文) Construction of mutation analysis pipelines for identifying novel cancer genes

研究代表者

上野 敏秀 (Ueno, Toshihide)

東京大学・医学(系)研究科(研究院)・助教

研究者番号：40381446

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：がんにおけるゲノム変異を解明するため、次世代シーケンサーで得られたデータを解析するコンピュータパイプラインの構築を行った。全エクソン解析、RNAシーケンス解析、またそれらの結果を統合するプログラムを開発し、さまざまながん種の変異リスト作成することができた。実際、作成した解析結果から疾患の原因遺伝子変異を見つけている。例えば、白血病患者への治療として実施された造血幹細胞移植後に発症する、移植ドナー細胞由来白血病の発症のメカニズムの解明となる原因遺伝子変異の同定に成功した。

研究成果の概要(英文)：To elucidate genomic mutation in cancer, we developed computer pipelines which analyzed data provided in a next-generation sequencer. For data of whole-exome sequencing and RNA sequencing, we were able to create the mutation profile list of various cancer tumors and succeeded in identification of novel cancer genes. For example, our pipeline results showed the identification of the genes which became the mechanism of leukemic evolution of donor-derived cells after allogeneic stem cell transplantation.

研究分野：バイオインフォマティクス

キーワード：バイオインフォマティクス 次世代シーケンサー 変異解析

1. 研究開始当初の背景

がんにおけるゲノム変異を解明することは、発がんの直接的な原因異常の同定に繋がり、それら異常遺伝子を標的とした新たな治療法開発のための重要な知見となる。それは、発がんに寄与するゲノム異常・エピゲノム異常(あるいはそのタンパク産物)を標的とし、その機能を制御する「分子標的治療薬」を開発することに繋がり、現在のがんのトランスレーショナルリサーチの主要な動向と言える。実際、BCR-ABL1 チロシンキナーゼ陽性慢性骨髄性白血病に対する ABL1 阻害剤(イマチニブ)や EML4-ALK チロシンキナーゼ陽性肺がんに対する ALK 阻害剤(クリゾチニブ)などは目覚ましい治療効果を上げることに成功しており、有効な分子標的治療薬法開発のためには、BCR-ABL1 や EML4-ALK のような発がんの直接的な原因となるゲノム異常・エピゲノム異常を同定することが重要である。このようなゲノム・エピゲノム異常を見つける一つの方法として次世代シーケンサーの使用が注目され、それらが普及し始めた時期であった。次世代シーケンサーの普及によりがんゲノム・エピゲノム解析を個々の研究室で遂行可能になったものの、研究開始当初は確立された解析ツールもなかったため、膨大な解読データの情報解析パイプラインの開発が求められていた。特に、当初は特定の遺伝子をターゲットにした cDNA を基質としてキャプチャーしていた方法を構築・改変を行っていたため、独自システムの開発が必要不可欠であった。

2. 研究の目的

次世代シーケンサーのデータから発がん原因となっている遺伝子変異を同定す

るためのパイプラインを開発し、疾患の原因となっている遺伝子変異を同定することが目的である。発がん原因の多くは、タンパクの配列変異による活性異常と予想されていることから、ヒトゲノム全体の約 1.5%に過ぎないエクソン領域のみを純化して配列決定する全エクソンキャプチャーが広くがん解析に用いられている。この方法は、ヒトゲノム約 3 Gbp すべてを解読する全ゲノムシーケンスと比較して解析できる領域が限定されてしまうものの、効率よくがんの体細胞変異を検出できることから「The Genome Cancer Atlas (TCGA)」や「International Cancer Genome Consortium (ICGC)」のプロジェクトにおいて主なゲノム解読手法として用いられている。一方、ゲノム構造異常から発生する遺伝子融合のほとんどは遺伝子内のイントロン領域で行われるため、エクソン配列を純化してしまう全エクソン解析ではイントロン情報が失われ遺伝子融合が検出できなくなってしまう。そのため、塩基置換・遺伝子融合・遺伝子の発現量を調べる方法として RNA(cDNA)シーケンス解析が用いられている。このように解析目的に応じて 2 つの基質(ゲノムと RNA(cDNA))を用いるため、それぞれに対応したプログラム(全エクソン解析と RNA シーケンス解析)の作成が必要となる。解析によって得られる変異候補は数多くリストアップされるが、それらの中にはシーケンサーの解読エラーなども多く含まれ偽陽性の要因になる。そのため、解析精度を上げるためのフィルター設計が重要である。そして実際に開発したパイプラインを用いて様々ながん種の変異プロファイリングデータを作成し、その中から原因遺伝子を特定することが目的である。

3. 研究の方法

疾患原因遺伝子変異探索のためには、がん腫瘍部由来と正常組織由来の二つのサンプルをペアで解析し、体細胞変異を検出する。がん種やサンプルの状態によって異なるが、1000 人ゲノムプロジェクトで登録されている一塩基多型 (1000G-SNP) を除いたとしても体細胞変異の候補数は多いため、目的の変異を探し当てるのはなかなか困難である。そこで、ペアとなる正常サンプルの変異だけでなく、解析した全正常サンプルの結果や健常者を解析したデータベースを用いて体細胞変異候補を絞り込むシステムを構築した。一方で、ペアとなるがん腫瘍部と正常部のサンプルが同一患者とは考え難い場合があった。解析の精度を高めるためには、同一患者由来というのが重要なため、サンプルそれぞれの 1000G-SNP を使ってペアかどうかを確認する機構を作成した。実際、431 症例中 7 症例は同一患者由来のサンプルとは考え難いという結果であった。また、変異解析においては全エクソン解析と RNA シークエンス解析の結果をマージできるように設計している点も特徴である。RNA シークエンス解析で用いるリファレンスとゲノムポジションへの変換テーブルの作成により全エクソンと RNA シークエンスデータの相互比較を可能にした。

次世代シークエンサーのスループットは年々増大してきたため、パイプラインの高速化にも取り組む必要があり、一つの解決手法として並列コンピューティングの実装を目指した。各プログラムユニットを Sun Grid Engine 上で動作できるよう移植・マルチスレッド処理・効率的なデータの分割と結合を組み込んだ並列処理パイプラインシステムを

構築した。これにより、それまでより数倍から数十倍の高速化を実現した。

4. 研究成果

構築した変異解析パイプラインにより、様々ながん種の臨床検体や細胞株を 430 症例 / 検体以上を解析し、疾患の原因遺伝子変異同定を行った。ヒト線維肉腫細胞株 HT1080 から RAC1(N92I)、RAC2(P29Q)の遺伝子突然変異を発見し、これらが強力ながん遺伝子になることを示した。また健常者骨髄中に低頻度に存在する遺伝子変異 (IDH2(R140Q)、DNMT3A(V150Gfs)) から、造血幹細胞移植後の急性白血病が発症することをつきとめた。その他にも 20 がん種以上を解析し、新規がん原因遺伝子変異の特定の足がかりとなるリストの作成に成功した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 5 件)

Leukemic evolution of donor-derived cells harboring IDH2 and DNMT3A mutations after allogeneic stem cell transplantation.

Yasuda T, Ueno T, Fukumura K, Yamato A, Ando M, Yamaguchi H, Soda M, Kawazu M, Sai E, Yamashita Y, Murata M, Kiyoi H, Naoe T, Mano H.

Leukemia. 査読有 2014;28(2):426-8.

DOI: 10.1038/leu.

Recurrent CDC25C mutations drive malignant transformation in FPD/AML.

Yoshimi A, Toya T, Kawazu M, Ueno T,

Tsukamoto A, Iizuka H, Nakagawa M, Nannya Y, Arai S, Harada H, Usuki K, Hayashi Y, Ito E, Kirito K, Nakajima H, Ichikawa M, Mano H, Kurokawa M. Nat Commun. 査読有. 2014;5:4770.

DOI: 10.1038/ncomms5770.

Transforming mutations of RAC guanosine triphosphatases in human cancers.

Kawazu M, Ueno T, Kontani K, Ogita Y, Ando M, Fukumura K, Yamato A, Soda M, Takeuchi K, Miki Y, Yamaguchi H, Yasuda T, Naoe T, Yamashita Y, Katada T, Choi YL, Mano H.

Proc Natl Acad Sci U S A. 査読有. 2013;110(8):3029-34.

DOI: 10.1073/pnas.

Cancer-associated missense mutations of caspase-8 activate nuclear factor- B signaling.

Ando M, Kawazu M, Ueno T, Fukumura K, Yamato A, Soda M, Yamashita Y, Choi YL, Yamasoba T, Mano H.

Cancer Sci. 査読有.

2013;104(8):1002-8.

DOI: 10.1111/cas.

STK10 missense mutations associated with anti-apoptotic function.

Fukumura K, Yamashita Y, Kawazu M, Sai E, Fujiwara S, Nakamura N, Takeuchi K, Ando M, Miyazono K, Ueno T, Ozawa K, Mano H. Oncol Rep. 査読有.

2013;30(4):1542-8. DOI: 10.3892/or.

〔産業財産権〕

出願状況(計 0 件)

なし

取得状況(計 0 件)

なし

〔その他〕

ホームページ等

なし

6. 研究組織

(1)研究代表者

上野 敏秀 (UENO TOSHIHIDE)

東京大学・大学院医学系研究科・助教

研究者番号: 40381446

(2)研究分担者

なし

(3)連携研究者

なし

〔学会発表〕(計 0 件)

〔図書〕(計 0 件)