

全ゲノム解析等実行計画（山本班） に係る研究報告

東京大学 医科学研究所

柴田 龍弘

○事業概要

事業の目的

本研究班は、がん全ゲノム解析計画のための体制構築として、インフォームドコンセント・患者返却方針、解析検体・臨床情報収集の検討、難治性がん（膵がん・白血病）並びに希少がんの試料を用いて、バイオバンク保管試料・シーケンス解析ワークフロー並びに新規技術について評価・検討し、一部パイロット解析を行うことを目的とする。

分担研究者

国立がん研究センター：柴田 龍弘、間野 博行、白石 友一、平田 真

京都大学：小川 誠司

東京大学先端技術研究センター：油谷 浩幸

東京大学 新領域：鈴木 穰

東京大学 医科学研究所：武藤 香織

早稲田大学：横野 恵

難治固形がん（膵がん・胆道がん）の全ゲノムパイロット解析

1. 解析数

全ゲノム解析（正常285症例、腫瘍部285症例）（KRAS VAFにより、腫瘍率>30%の症例を選別）

RNAシーケンス（腫瘍部 355症例）（低腫瘍率の症例も含む）

全ゲノムメチル化シーケンス（WGBS）（腫瘍部 160症例）

2. 解析状況

WGSシーケンスデータは、現在解析WGで開発した共通パイプラインにて解析中。RNAseq/ WGBSデータについては、解析WGでの議論も参考にしつつ、in-houseでの解析を実施する予定。

3. 今後の進捗見込み

臨床情報収集に関する方針・登録システム稼働に伴い、臨床情報の収集を行う。データ登録・共有に関する方針の決定に従いデータを登録する。共通パイプラインで解析された変異データの評価を行い、RNAseq/WGBSデータと統合的解析を行う。

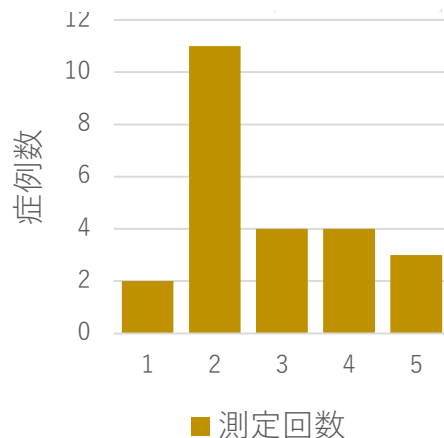
難治がん（白血病）経時的検体を用いた 全ゲノムパイロット解析

造血器腫瘍に対する経時的検体の全ゲノムシーケンスによるクローン構造の変化の解析

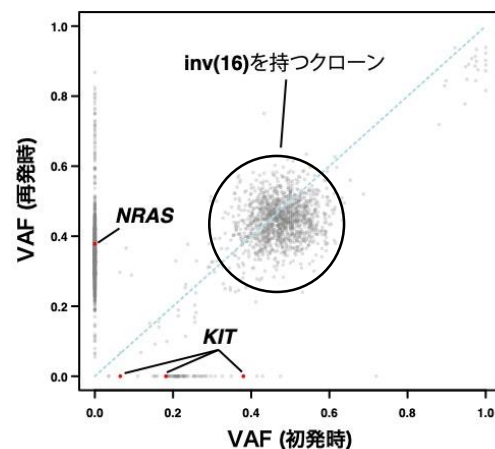
造血器腫瘍 25症例
(70腫瘍検体+25正常
コントロール)

疾患	症例数
ALL	1
AML	7
CML	1
MDS	14
MDS/MPN	1
MPN	1

25

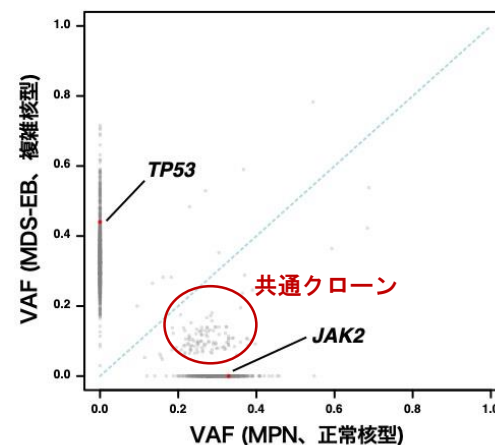


AMLの初発・再発時



inv(16)転座を含む
major cloneが初発時、
再発時を通じて存在し
ている。
初発から再発時に、3
つのKIT変異を含む
subcloneが消失し、
**NRAS変異を含む
subcloneが出現した。**

MPN症例に続発したMDS



JAK2変異陽性MPNから、
JAK2変異陰性、
TP53変異陽性のMDS
が発症した症例。
全ゲノム解析によって、
**MPN, MDSを通じて共
通して存在するクロー
ンが判明した。**

骨軟部腫瘍の全ゲノムパイロット解析

骨軟部腫瘍ゲノムコンソーシアム

国立がんセンター (NCC)
 神奈川県立がんセンター (KCC)
 栃木県立がんセンター (TCC)
 その他施設

160 症例

腫瘍組織(T): 凍結
 正常組織(N): 凍結 (手術時余剰検体)
 /血液 (バイオバンク)

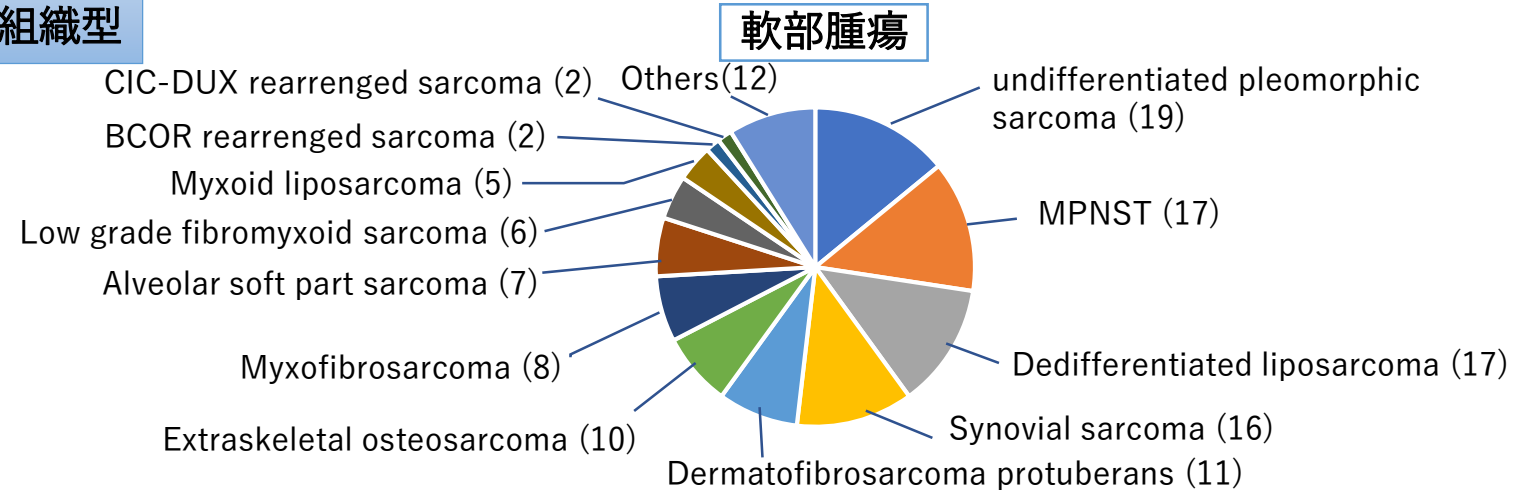
DNA(N, T pair)
 RNA(T)

・ WGS解析 (N 160, T 196**)
 ・ RNAseq解析 (T 162**)
 ** 再発、転移検体等を含む

医療機関別症例数

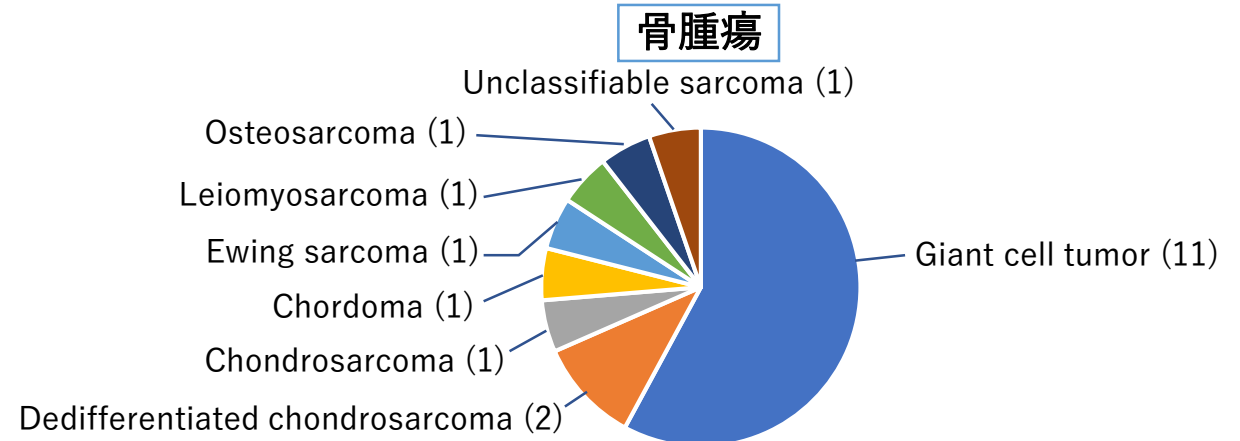
施設	症例数
NCC	122
TCC	7
KCC	9
others	22

組織型



臨床情報概要

臨床情報 (KCC含まず)	症例数
骨/軟部	骨 19 軟部 132
年齢 (中央値)	4-93 (49)
性別	男 71 女 48
経過観察機関 (中央値) (カ月)	1-163 (39)



その他の研究成果

ELSI WG 研究成果

(分担研究者：武藤香織・横野 恵)

● 検討項目

1. ICF案の作成
2. 既存試料・情報の利用に関する留意点
3. 解析結果の返却に関する留意点
4. 遺伝的特徴・情報に基づく差別禁止に関する法制度の検討

● 成果

- 上記検討に基づいてICFおよび留意点の案を取りまとめた。
- なお、昨年度の段階では事業計画が流動的であったため、上記項目1～3については対応方針を具体的に検討できなかった部分が残されている（とくに結果返却関連。下記参照。）。今後、事業計画の具体化にあわせてさらに検討を加え、対応方針を明確化する必要がある。
- 解析結果の返却方針に関しては、要検討事項が多岐にわたり、現実にどのような対応が可能かは、費用・人材等のリソースにもよる。研究での解析と医療目的の検査の違いを踏まえ、無理が生じない形の返却とすることが重要。これらの点を踏まえて結果返却に関するICF記載内容を慎重に検討する必要がある。また、既存試料・情報から得た結果の返却に関しては、新規試料・情報の場合とは異なる検討を要する事項も多いため、別途検討する必要がある。

バイオバンクWG 研究成果

間野博行[◎]、秋山靖人、石川俊平、織田克利、鈕持広知、土原一哉、南谷泰仁、野田哲生、松田浩一、三森功士、武藤学、谷田部恭、谷内田真一
([◎]WG長・班員)

2. 効率的かつ統一的なシーケンスや解析方法等についての検討 (バイオバンクWG)	・【臨床情報】 臨床情報の内容、収集方法
	・【臨床情報】 臨床情報の収集における現場負担軽減策の検討
	・【検体】 検体の処理・収集・保管等のワークフローを確立する

- 1 収集臨床情報の内容・収集方法：固形腫瘍・造血器悪性腫瘍それぞれのサブWGにより**収集方法の確定**。
- 2 **現場負担軽減策**の検討：収集する臨床項目を厳選するとともに、選択方式を基本として定義の明確化を行い入力品質の向上を図る。また情報入力する人員の確保に留意。

がん全ゲノム解析プロジェクト
症例基本情報・同意情報

患者基本情報			
患者識別ID			
施設番号(任意)			
担当医師名			
性別	<input type="radio"/> 男 <input type="radio"/> 女 <input type="radio"/> 未入力・不明	生年月日	(カレンダー選択)
年齢	(自動計算)歳		
人種	<input type="radio"/> JAPANESE <input type="radio"/> ASIAN <input type="radio"/> OTHER		
身長	cm	<input type="checkbox"/> 未測定	
体重	kg	<input type="checkbox"/> 未測定	
BMI	(自動計算)		
腫瘍の別	<input type="radio"/> 固形腫瘍 <input type="radio"/> 造血器腫瘍		
これまで登録の有無	<input type="radio"/> あり <input type="radio"/> なし <input type="radio"/> 不明		
過去の登録ID			
症例関係区分	<input type="radio"/> 重複がん <input type="radio"/> 前医からの転院フォロー <input type="radio"/> 前医実施後再検査 <input type="radio"/> その他		

解析WG 研究成果 1

◎小川誠司、油谷浩幸、上田宏生、浦上研一、岡田随象、片岡圭亮、柴田龍弘、白石友一、鈴木穰、十時泰、中村祐輔、南谷泰仁、藤本明洋、森下真一

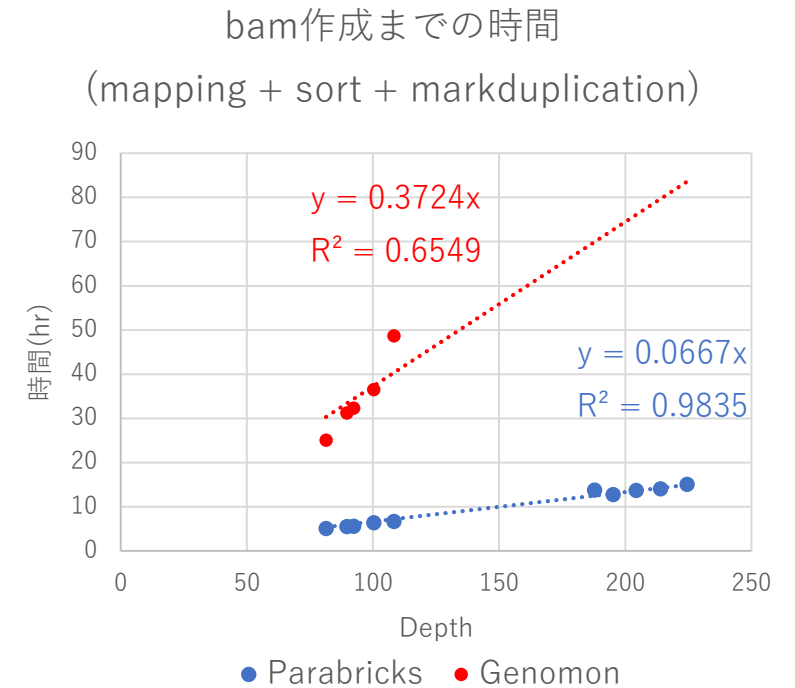
① 大腸がん5検体、正常5検体 → WGS (腫瘍200x、正常100x) => 返却済み、解析中

② 大腸がん5検体→WGBS (全ゲノムメチル化解析) ライブラリ作成 → シーケンス → 3/10HDD到着、解析中

③ 膵がんオルガノイド 高分子DNA 8ペア → Long read解析

GPU (Parabricks)を使用したmappingはCPU (Genomon)を利用した場合と比較して5.6倍高速であった。

全ゲノムシーケンスにおける情報解析 platform性能の比較



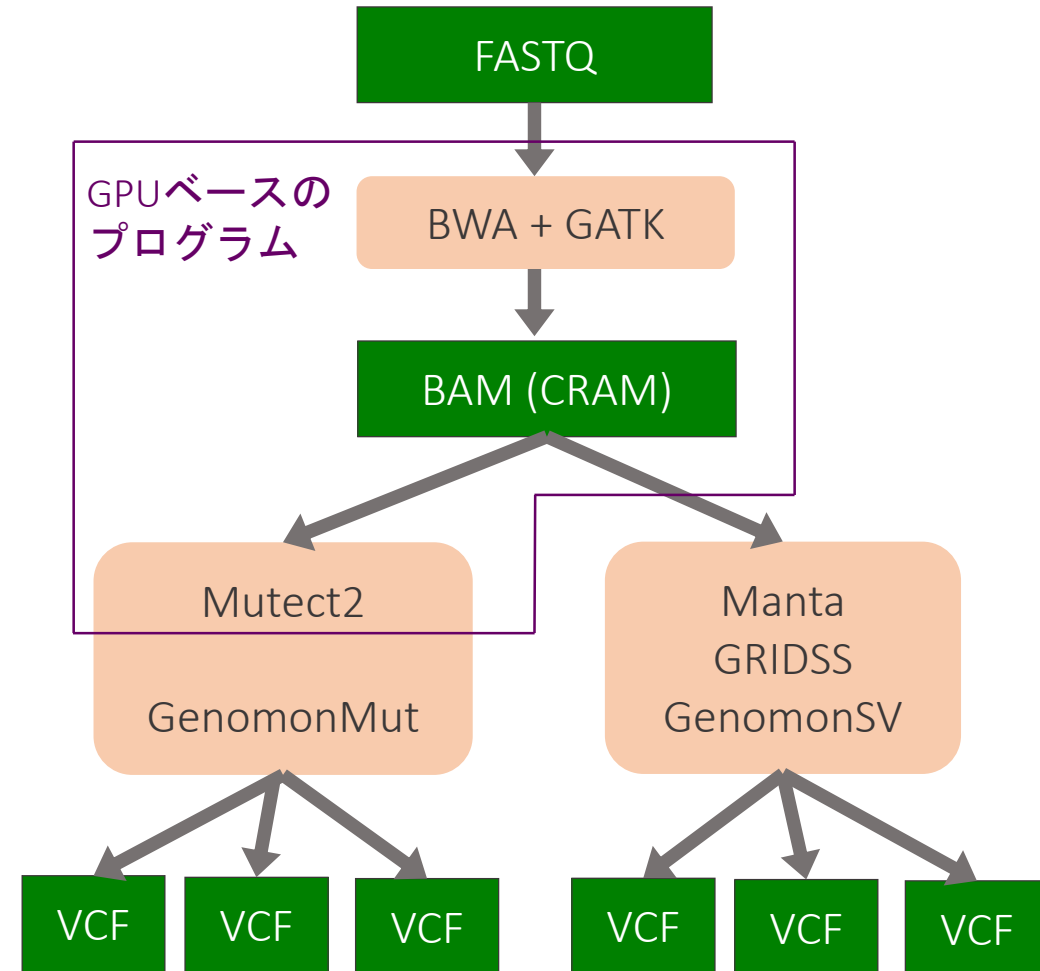
大腸がん5症例(腫瘍 200x・正常100x)を用いた比較

解析WG 研究成果 2

がん体細胞全ゲノム解析パイプラインの開発

- GPUを用いた処理の高速化。
 - 計算のアラインメント、変異検出の一部については、GPU上で実行可能なアクセラレーター（Parabricks）により高速化を行う。
- 予め主要なソフトウェアについての解析結果を配備する。
 - ユーザーが独自にプログラムを実行させる手間を省く。
 - 国際がんゲノムコンソーシアムなどでは、複数の変異検出ソフトウェアの手法を統合的に利用することが一般的である。
- **1台のGPUで6~7/dayの計算処理が可能。**
 - Tumor 100x + Normal 30x ペアでの見積もり。
 - GPUサーバーに加えて、小規模のCPUサーバー(総計で40コア、200GBメモリ程度)が必要。
- 今後の課題。
 - 数百、数千検体規模のデータで運用を行う。
 - 効率的な偽陽性フィルタリングの方法の開発。

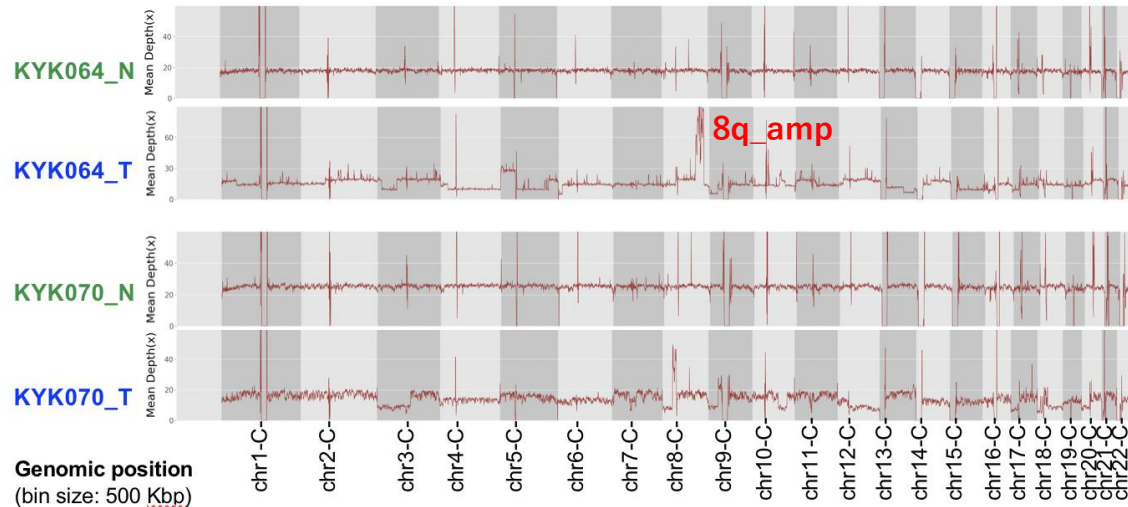
開発したパイプラインの処理フロー図



解析WG 研究成果 3

ロングリード解析(患者由来膵がんオルガノイド)の検討

- ショートリードではマッピング不可能なゲノム領域の情報取得可能
- とりわけ、構造変異の検出(増幅、欠失、転座、逆位など)
- パイロット研究として患者検体由来がんオルガノイドを用いて検証
- 1ラン(PromethION)で30Xゲノムの情報収集が可能
 - 腫瘍 51.4Gb (37.0~72.2) 配列長(N50)7.2kb
 - 末梢血 58.3Gb (30.7~110.4) 配列長(N50)9.1kb



case	type	Estimated Bases (Gb)	Estimated N50 (Kb)	Femto Pulse Peak (kb)
KYK015	tumor	46.75	7.28	
KYK019	tumor	45.19	8.61	20
KYK036	tumor	59.42	6.85	7.7
KYK054	tumor	47.98	7.72	9.3
KYK064	tumor	72.23	6.73	8
KYK067	tumor	40.83	7.49	8.2
KYK070	tumor	61.9	5.65	7.6
KYK084	tumor	37.02	7.5	8.2
KYK015	blood	52.55	7.53	8
KYK019	blood	30.66	19.89	33
KYK036	blood	66.68	7.31	7.7
KYK054	blood	56.32	10.18	9.5
KYK064	blood	60.05	7.27	7.7
KYK067	blood	48.35	6.83	8.5
KYK070	blood	110.45	6.68	7.9
KYK084	blood	41.63	6.94	8