

# 第10回全ゲノム解析等の推進に関する専門委員会

## (2) 全ゲノム解析等に係るAMED研究について

開催日：令和4年7月7日（木）

国立研究開発法人 日本医療研究開発機構（AMED）  
ゲノム・データ基盤事業部 ゲノム医療基盤研究開発課

### 目次

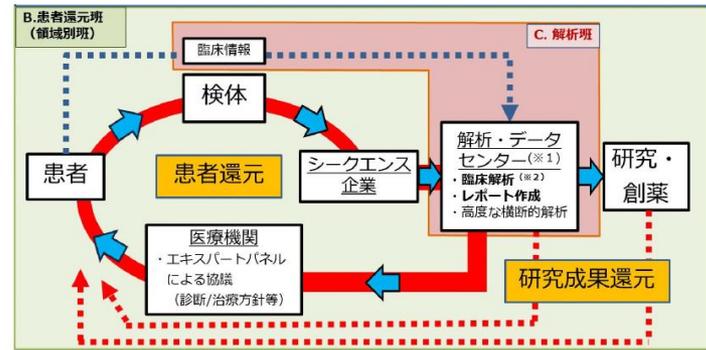
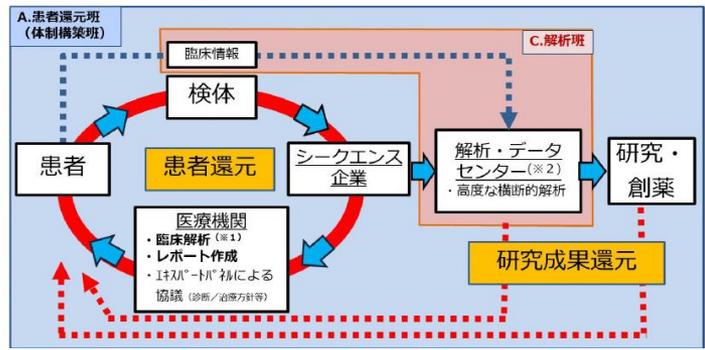
1. R4年度 AMED研究概要
2. WGS, RNA-seqの出検の進捗、データQC状況
3. 調整費によるAMED提案
4. 研究班とりまとめ
5. EDC入力の進捗
6. R4年スケジュール

# R4年度 AMED研究概要

## ● 患者還元体制について

(参考：専門委員会 (3/2) 資料1-1, p6, p10)

- ✓ 既存の3医療機関で400症例の患者還元を行う (A体制：自施設完結型体制) とともに、患者還元体制の整った医療機関を、専門性や地域性を考慮しつつ分担医療機関として追加する。
- ✓ R4年度は、A体制を維持するとともに、分担医療機関を追加し、B体制 (解析・データセンター体制) の整備に向けた比較検討を開始する。



## ● 全ゲノムプロジェクト症例内訳とR4年度実施内容

(参考：専門委員会 (3/2) 資料1-1, p6)

公募の種類	がん種	代表機関・代表者	令和3年度
A班：患者還元班 (体制構築班)	難治がん等	国立がん研究センター 角南久仁子	500症例 (内新規の患者200症例)
	難治がん等	静岡がんセンター 浦上研一	500症例 (内新規の患者200症例)
	難治がん等	がん研有明病院 上野貴之	500症例 (内新規の患者200症例)
B班 患者還元班 (領域別班)	消化器がん	東京大学 柴田龍弘	1,400症例
	血液がん	京都大学 南谷泰仁	1,400症例
	小児がん	東京大学 加藤元博	1,400症例
	希少がん	東京大学 松田浩一	1,400症例
	婦人科がん	がん研有明病院 森誠一	1,400症例
	呼吸器がん 他	国立がん研究センター 河野隆志	1,400症例
C班：解析班		東京大学医科学研究所 井元清哉	A班、B班併せて、9,900症例の解析

令和4年度体制・実施内容		
代表：国立がん研究センター 分担：国立がん研究センター東病院 分担：成育医療研究センター	600症例 + a	600症例の内訳は、新規の患者400症例と、分担医療機関の新規の患者200症例。また、進捗状況に応じて、+aとして、合わせて最大200症例を追加解析予定。
代表：静岡がんセンター 分担：近畿大学病院	600症例 + a	
代表：がん研有明病院 分担：慶應義塾病院 分担：大阪大学病院	600症例 + a	
臨床情報の登録を行うとともに、蓄積された全ゲノムデータを用いた研究を行う。また、A班とも連携しB班全体としての成果をまとめる。		
臨床情報を収集するとともに、統一パイプラインによる解析及びレポート作成を行う。また、集中管理システムの構築、全ゲノム解析結果に基づいた臨床応用のための出口戦略の構築を行う。		

- ✓ 各班は連携し、臨床情報等の収集及び高度な横断的解析等を行う。
- ✓ 各班は、実施状況について「全ゲノム解析等の推進に関する専門委員会」に報告し、当該委員会の方針に沿って解析等を行う。

## C班（解析班）の概要（がん領域）

	項目	担当	
解析班 (C班)	①集中管理システム (データ・検体)	新規の患者の検体およびゲノム・臨床情報の集中管理システム構築	○集中管理チーム（※1）
	②ゲノム解析・ クラウド基盤・監視	ゲノムデータベースやクラウド基盤・監視構築 統一パイプライン、アノテーション 高度な横断的解析（AI活用含む）	○ゲノム解析チーム
	③臨床情報等の活用	③-1. 臨床情報収集システム構築 自動収集API開発	○臨床情報チーム（※2）
		③-2. レポート作成システム	○レポート作成チーム
	④データ共有	データ共有・研究支援システム	○データ共有チーム
⑤出口戦略の構築	基本コホート+戦略コホートの構築	○出口戦略チーム（※3）	

※1：集中管理チームは、WGSデータ、臨床情報、検体および検体情報を紐付けて管理するシステムの構築（ID管理含む）を行う。また、検体管理については日本衛生検査所協会の専門家の協力を得て臓器別SOP詳細作成する。

※2：臨床情報収集チームは、基本コホートの全症例の基本的な臨床情報および、戦略コホート毎の付加的な臨床情報を統一的なシステムで収集し、データ共有、研究支援システムと連携させ幅広い利活用を目指す。その際は、データの相互利活用が可能となるように難病領域と連携する。なお、臨床情報収集項目については出口戦略チームが検討し、5年後には電子カルテ上のテンプレートからAPIを用いた臨床情報の自動収集を目指す。

※3：出口戦略チームは、臨床医を主体として構築する。また、出口戦略チームとして成果等を共有出来るシステムの構築や、集中管理チーム、臨床情報チームと協力し、全てのコホートにおいて統一的な臨床情報収集・集中管理、臨床試験支援システムの構築を目指す。

# データベース構築\_進捗状況

## A班：エキスパートパネル体制構築班・難治がん等

- 注1：FASTQデータは、症例IDに基づきペアをカウント  
 注2：QCは、研究班から提供されたペアリストを元に集計  
 (ペアリストが解析班に届いていない症例は集計には含まれていない)  
 注3：分割納品によるデータ統合が必要な症例の統合前のものが一部症例に含まれる  
 (それらの症例は塩基数が不足)

(令和4年6月20日時点)

		R3年度				R4年度			
代表機関・代表者		出検数		FASTQ データ 受領数 (T/Nペア確定数)	データ確定数 (QC完了数) 塩基数, QV%	出検数		FASTQ データ 受領数 (T/Nペア確定数)	データ確定数 (QC完了数)
国立がん 研究セン ター	角南 久仁子	WGS	500 (出検完了)	499	N:500/500,500/500 T:501/501,501/501	WGS	0		
		RNA-seq	478 (出検完了)	0		RNA-seq	0		
静岡がん センター	浦上 研一	WGS	500 (出検完了)	450 ※データ取込み中含む	ペアリスト受領。デー タ取込み後実施	WGS	110		
		RNA-seq	500 (出検完了)	0		RNA-seq	0		
がん研有 明病院	上野 貴之	WGS	535 (出検完了)	488	N:314/314, 314/314 T:299/314, 314/314	WGS	118		
		RNA-seq	535 (出検完了)	0		RNA-seq	118		

# データベース構築\_進捗状況

## B班：領域別班（R3年度）

注1：FASTQデータは、症例IDに基づきペアをカウント

注2：QCは、研究班から提供されたペアリストを元に集計

（ペアリストが解析班に届いていない症例はQCの計算は終えているが集計には含まれていない）

注3：分割納品によるデータ統合が必要な症例の統合前のものが一部症例に含まれる（それらの症例は塩基数が不足）

（令和4年6月20日時点）

がん種	代表機関・代表者		出検数		FASTQ データ受領数 (T/Nペア確定数)	データ確定数 (QC完了数) 塩基数, QV%	備考
			WGS	RNA-seq			
消化器がん	東京大学	柴田龍弘	WGS	1,400 (出検完了)	983	N: 799/799, 799/799 T: 782/799, 799/799	
			RNA-seq	1,400 (出検完了)	50ペア		
血液がん	京都大学	南谷泰仁	WGS	1,400 (出検完了)	862	N: 854/859, 859/859 T: 817/859, 858/859	RNAは検体品質不良分を除いた出検数。 RNA-seq58検体のデータ提供予定
			RNA-seq	747 (出検完了)			
小児がん	東京大学	加藤元博	WGS	1,472 (出検完了)	1,420	N:1285/1292, 1292/1292 T:1267/1292, 1292/1292	解析データが確認できれば出検完了。 (データ不足等の場合は再度差し替えなどを検討)
			RNA-seq	1,015 (出検完了)			
希少がん	東京大学	松田浩一	WGS	1,558 (出検完了)	1,408	N:1017/1026, 1026/1026 T:1017/1026, 1026/1026	予定数の出検終了。 QC不良による追加分も出検済み。
			RNA-seq	1,396 (出検完了)	0ペア (T: 105)		
婦人科がん	がん研有明病院	森誠一	WGS	1,400 (出検完了)	1,379	N: 731/734, 734/734 T: 682/734, 734/734	組織で1,239検体 DNAで388検体を出検済み
			RNA-seq	1,366 (出検終了)			
呼吸器がん他	国立がん研究センター	河野隆志	WGS	1,400 (出検完了)	1,058	N: 1008/1008, 1008/1008 T: 1028/1028, 1028/1028	QC不良症例の再選択・再提出を完了。
			RNA-seq	1,149 (出検完了)			

### ● 提案コンセプト

**大規模なロングリード解析による全ゲノム解析を主軸としたマルチオミックス解析実施のための実行可能性、実施意義の検証**

→ **全ゲノム解析プラットフォームの高度化、創薬研究の基盤整備**

### ● 世界情勢

#### <ロングリード解析の技術革新>

- ✓ 従来のWGS（ショートリード）と同等の精度での解析、ウルトラロングリード解析の登場など技術革新がめざましく、その結果、ロングリードシーケンスにより、これまでテロメアやセントロメアの領域に多く存在する繰り返し配列等のため未解読であった領域を含むヒトゲノムの完全長が解読され、令和4年4月にScience誌に成果論文の特集号が報告された。完全長の8%に相当する領域が未解読であったこと、新たに解読された領域には99個の新規遺伝子が存在すること、などが示された。（4/1, science）
- ✓ 従来から、大きなゲノムの構造異常やウイルスゲノムの挿入、がん特有の病的なRNAアイソフォームの同定など、ロングリード解析のがん全ゲノム解析への寄与は大いに期待されていた。これに加え、今回のロングリード解析の技術革新によりロングリード解析でしか正確に解析出来ない領域から数万、数十万の変異（挿入/欠損、点変異）が検出されることが報告された。（5/11, Cell Genomics）

#### <がんゲノム解析>

- ✓ 全ゲノム情報とエピゲノム情報等を組み合わせた統合解析が、がん種横断的かつ大規模（19がん種、3,949検体）に実施（4/8, science）などマルチオミックス解析の大規模解析が開始されている。



**がん領域における研究開発は、ロングリードによる全ゲノム解析を主軸としたマルチオミックス解析が主流になると予想される。**

● 調整費実施内容

	班	ロングリード解析の症例数と対象の例	提案コンセプト・出口戦略	C班
B班	柴田班	50例：ウイルス発がん症例等	AIを用いた簡便な画像診断技術開発	<ul style="list-style-type: none"> <li>✓ ロングリード解析等の解析パイプライン構築</li> <li>✓ 大量のデータ受け取り、整合性確認等データセンターシステムの自動化</li> <li>✓ 統一解析パイプラインの高速化</li> </ul>
	南谷班	50例：染色体転座症例等	ゲノム異常による影響（異常なRNA産生）解明	
	加藤班	60例：セントロメア領域に異常のある症例等 遺伝的素因の関与が想定される症例	小児がんの原因や遺伝的素因解明	
	河野班	70例：7がん種の比較、薬剤抵抗性獲得症例等	がんの多様性や、がん種横断的解析を見据えた検討	
	森班	50例：薬剤奏効性、トランスポゾンの関与等から選択	バイオマーカー(免疫療法、標準治療)の開発	
A班	角南班	15例：再発難治症例	手術適応が無い再発難治症例に全ゲノム解析の恩恵を届ける。	
	浦上班	15例：間質の多いがん種	間質が多く精度の高い解析が困難な症例に全ゲノム解析の恩恵を届ける。	

\* ロングリード解析 計310症例（予定）：WGS（ショートリード）ではドライバー異常が検出されなかった症例を中心に解析。  
 \* 上野班、松田班（希少がん）は、多数症例を用いたマルチオミックス解析を提案予定。

● 調整費の措置にかかわるスケジュール（仮）

調整費スケジュール



## 調整費措置によるマルチオミックス解析実施に向けて必要な検討事項

### 検討事項（統合解析のための解析要件の整理）

- ✓メチル化解析等、追加して実施する解析のQC要件
- ✓高額な高精度（網羅的）解析の実施要件（WGS（ロング）、メチル化解析（WGBS））
- ✓ロングリード解析とWGS（ショート）、メチル化解析の統合解析

### 検討事項（解析データセンター関連）

- ✓統一解析パイプラインの対象となる解析
- ✓メチル化解析等、委託解析データの納品の流れ

## 全ゲノム解析班 サブグループの整理（7月時点）

実施内容	AMED研究班			
	主担当	副担当	参加	情報共有
全ゲノム解析等の統合解析	B班 6班	C班 解析基盤	A班 解析等担当	C班 データ共有 C班 検体管理
臨床情報収集システムの構築（基本コホート用）	C班 臨床情報収集（DDC）	C班 出口戦略 -基本コホート	A班 3班 各がん種担当（臨床医）	C班 出口戦略 -戦略コホート
レポート作成（フォーマット・記載項目検討）	C班 レポート作成	C班 出口戦略 -基本コホート	A班 3班	C班 出口戦略 -戦略コホート C班 解析基盤
レポート作成（作成パイプライン）	C班 レポート作成	C班 解析基盤		A班 3班 C班 出口戦略

## ● 検討開始を要請するサブグループ

✓ 研究基盤整備：臨床情報収集システムの構築（戦略コホート用）

→臨床情報コードの統一、戦略コホート用DDC構築時の課題

✓ 参加施設拡大：中核拠点、拠点病院の参加

→検討を依頼した施設の意見、参加を検討する際にネックとなっている点

進捗状況（代表機関：達成状況、分担機関：項目達成の施設数/総数 6/8時点）

	①グローバルIPアドレス の登録依頼	②クライアント電子証明 書のインポート	③Webブラウザ詳細設 定のTLS1.2の有効化	④倫理審査	⑤臨床情報入力開始
柴田班					
食道がんG	1/5	1/5	1/5	5/5	0/5
肝臓がんG	1/1	1/1	1/1	1/1	1/1
胆道・膵G	4/9	4/9	4/9	0/9	0/9
胃がんG	1/1	1/1	1/1	1/1	1/1
加藤班					
代表	○	○	○	途中	×
分担機関	2/2	2/2	2/2	0/2	0/2
松田班					
代表	対象外 (代表機関である東大新 領域は統括班のみ)	対象外 (代表機関である東大新 領域は統括班のみ)	対象外 (代表機関である東大新 領域は統括班のみ)	対象外 (代表機関である東大新 領域は統括班のみ)	対象外 (代表機関である東大新 領域は統括班のみ)
分担機関	14 /40	8 /40	8 /40	9 /40	1 /40
森班					
代表	○	○	○	○	○
分担機関	3/5	2/5	1/5	5/5	0/5
南谷班					
代表	代表で患者情報を直接収 集しないため適応なし	代表で患者情報を直接収 集しないため適応なし	代表で患者情報を直接収 集しないため適応なし	○	×
分担機関	7/16 (39施設中16施設 がEDC使用)	6/16	6/16	33/39	0/39
河野班					
代表	○	○	○	○	○
分担機関	3/7	3/7	3/7	3/7	2/7

## ●EDC入力進捗(6/8時点)

- ✓ 代表機関等の主要な施設における準備は完了しつつある。
- ✓ 分担施設における準備状況に遅れがあったため、早急に準備を整え、倫理審査通過後の入力開始に影響が生じないように対応中。
- ✓ 倫理審査が完了しつつある。
- ✓ 準備が整った施設において入力開始。

## ●EDC入力開始に向けた調整事項

- ✓ 臨床情報のExcelでの取り込みシステムの検討（費用等）  
→費用対効果（要する期間、対象となる症例数など）
- ✓ 連携病院等で入力が難しい施設の対応（拠点となる施設でまとめて入力など）

R4年度-I			II			III			IV		
4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月	3月
						令和5年度公募					
						← 公募		← 書面評価		← ヒアリング	
										● 締結	
										● 紹介	
AMED研究班合同会議											
▲											
4/22	5/26	6/27	7/26	8/29	9/27	10月以降は日程調整中					
「全ゲノム解析等の推進に関する専門委員会」											
●											
(5/16)											
●											
(7/7)											

### ● 合同会議予定

1) AMED研究班合同会議：7月26日（火）14:00~15:00（最大延長15:30）

### ● 今後の予定

- ✓ 10月以降のAMED合同会議：日程調整を開始
- ✓ 後ろ向き検体の同意内容の確認：再同意の要件、使用する統一ICF、実施スケジュールなどの整理を開始。