

難病に関するゲノム医療推進にあたっての統合研究（20FC2002）
難病のゲノム医療推進に向けた全ゲノム解析（先行解析）に関する説明会

① 全ゲノム実行計画等に基づく先行解析（難病） の概要、ゲノム基盤の構築

国立国際医療研究センター
ゲノム医科学プロジェクト
徳永勝士

令和2年10月28日、東京

先行解析のイメージ

政策研究班

①協力医療機関について

②同意書の検討、国民への普及啓発

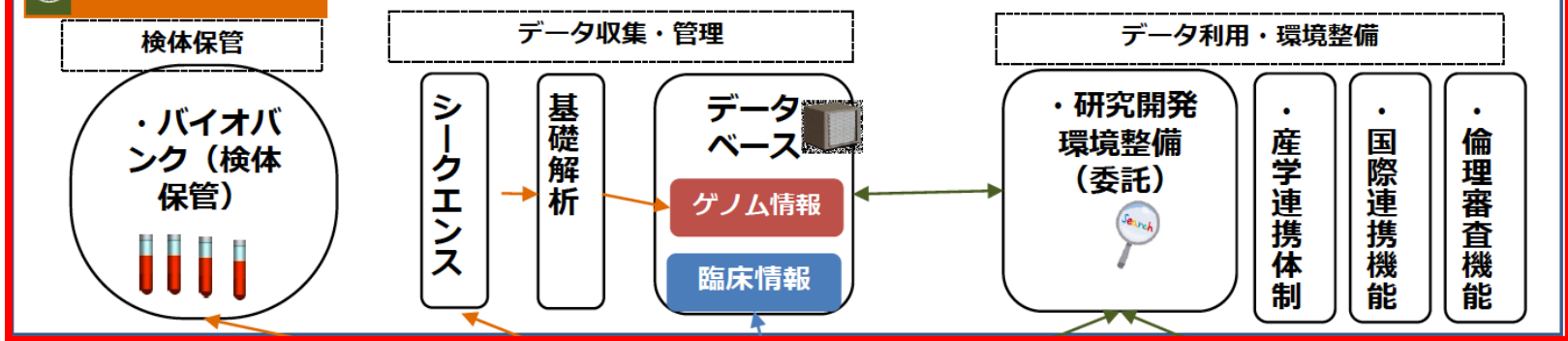
③臨床情報の検討

④ゲノム基盤の運営・管理方法について

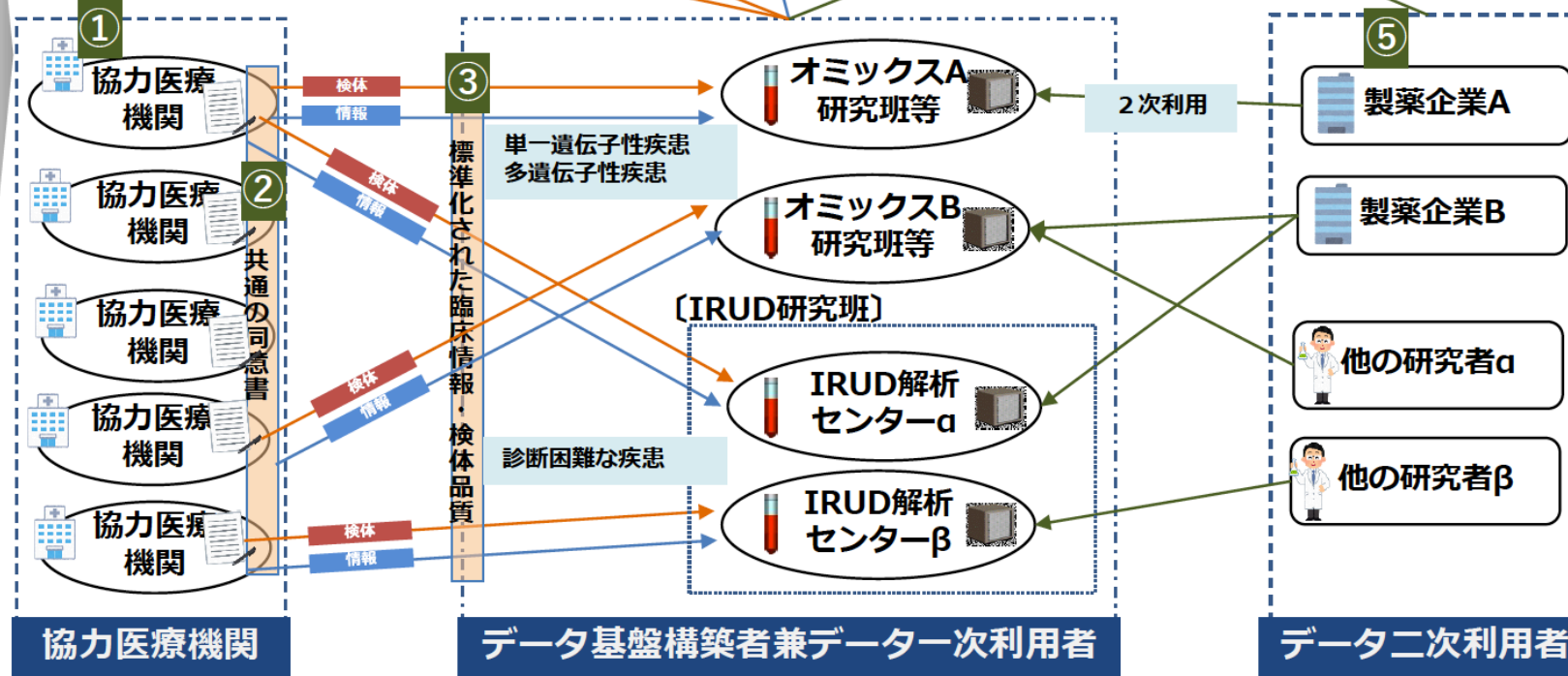
⑤医薬品開発の促進に向けたゲノムデータ基盤のあり方

⑥人材育成等

④ゲノム基盤



input

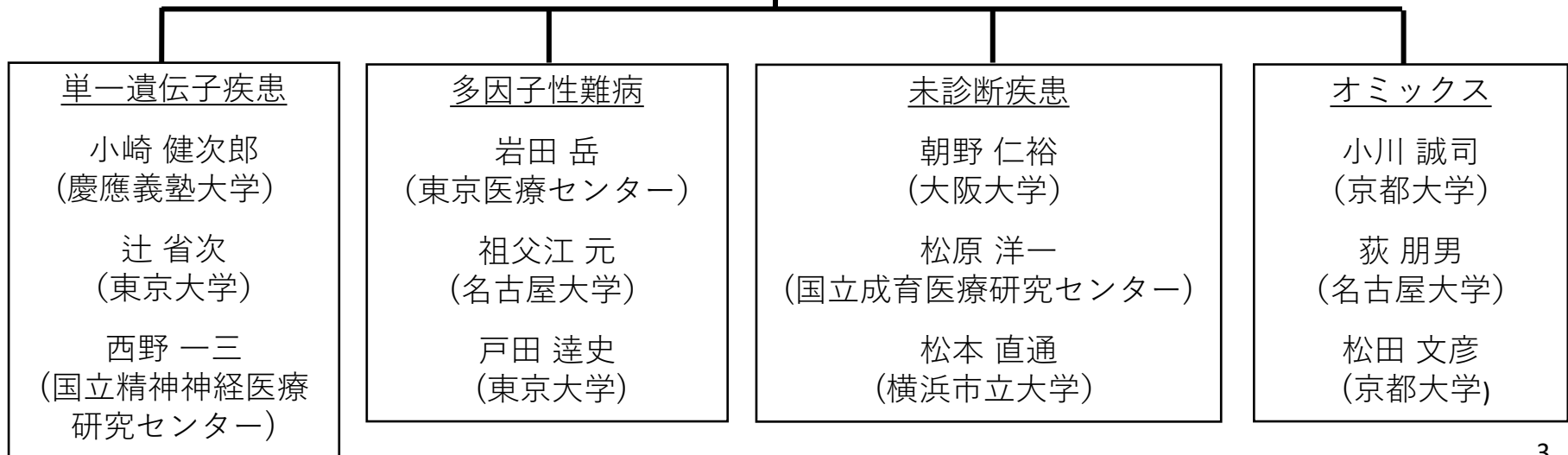


難病のゲノム医療推進に向けた全ゲノム解析基盤 に関する研究開発 (AMED) の研究体制

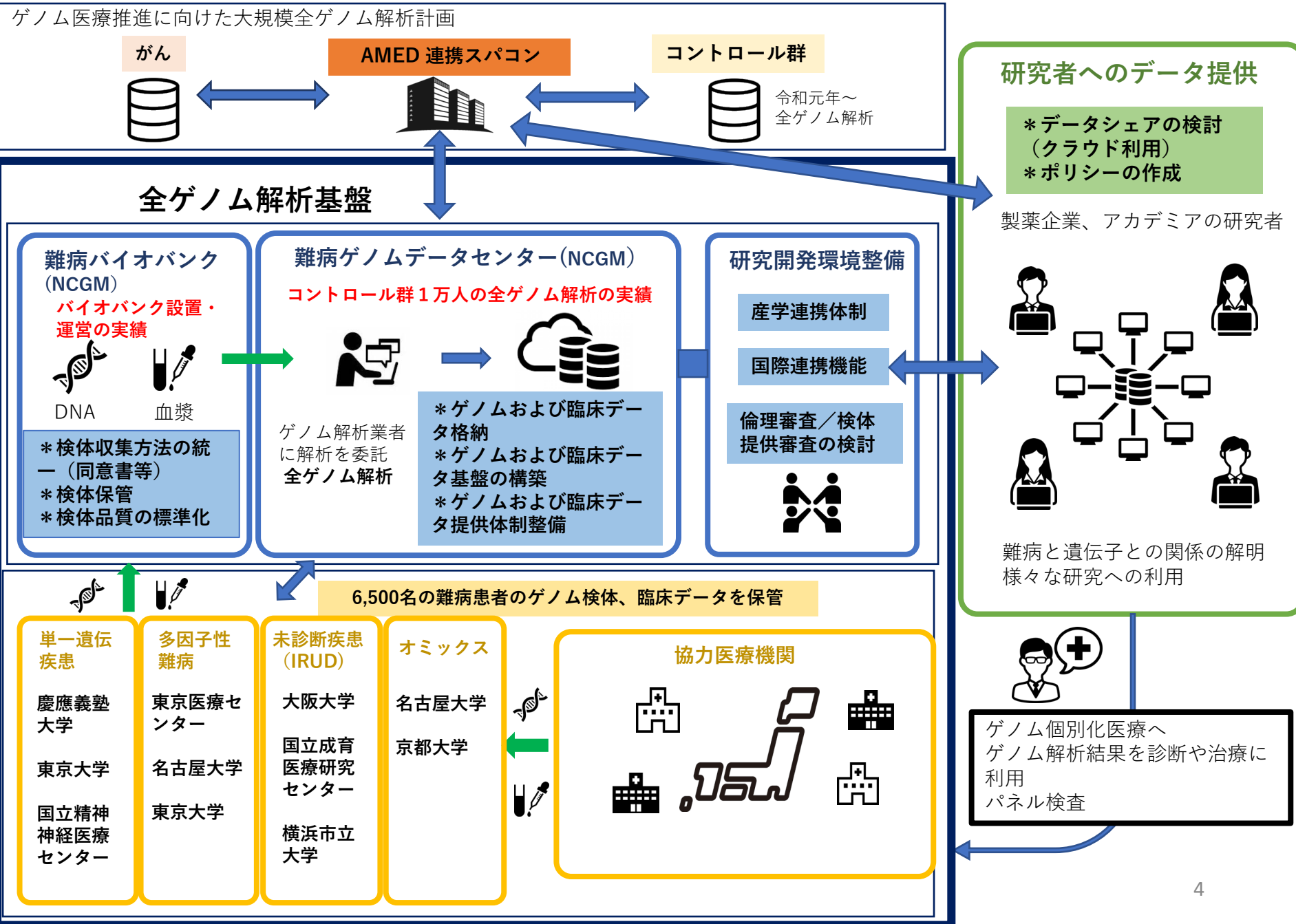
研究代表者
國土 典宏
国立国際医療研究センター(NCGM)

研究副代表
水澤 英洋
国立精神神経医療研究センター(NCNP)

ゲノム基盤
徳永 勝士
国立国際医療研究センター(NCGM)



難病のゲノム医療推進に向けた全ゲノム解析基盤に関する研究開発(概要図)



「ゲノム基盤」の担当課題

課題1 臨床検体の保存・管理の方針、運用方法

- 検討事項1 バイオバンクの構築方針について
- 検討事項2 (既存バンクに委託する場合) 当該バンクに求められる要件
- 検討事項3 バイオバンクに保存する検体の量・種類等について

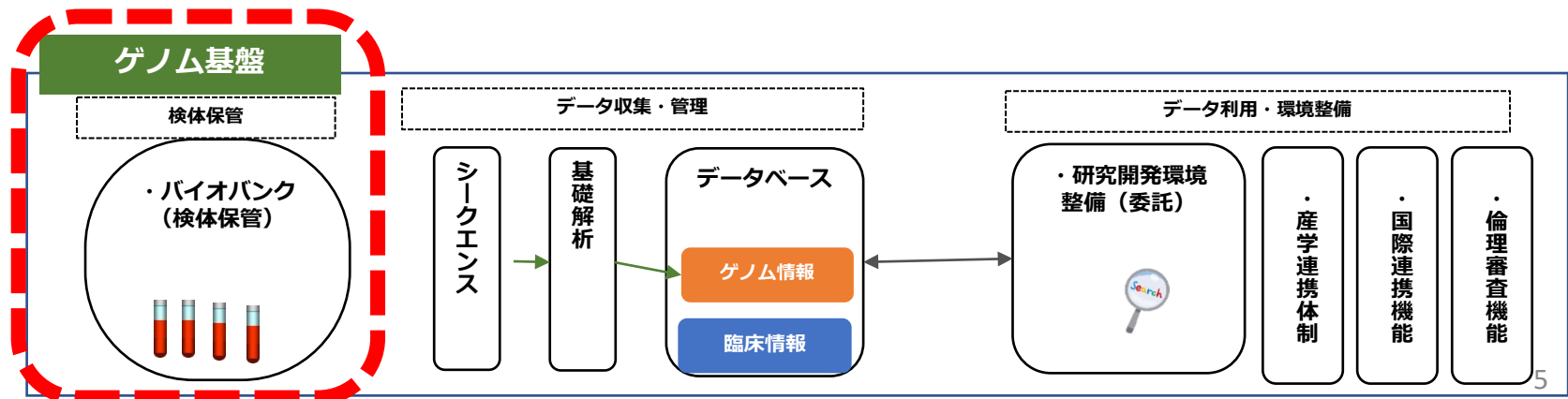
課題2 シークエンス・基礎解析の方針、運用方法

- 検討事項4 ゲノムデータ基盤のシークエンス解析部門に求められること
- 検討事項5 シークエンス解析に関する委託・自主運営の比較について

課題3 ゲノムデータ・臨床データの保存・管理の方針、運用方法

- 検討事項6 ゲノムデータ保管に必要とされる能力等について
- 検討事項7 パブリッククラウドとオンプレミスの比較について
- 検討事項8 臨床データの保存・管理について

課題4 その他、運営・管理に必要な事項



課題 1：バイオバンクにおける臨床検体の保存・管理の方針、運用方法

○検討事項 1 バイオバンクの構築方針

- 全ゲノム解析等実行計画（第1版）に基づき、保存が必要な検体数について、先行解析では約5,500人（6,500人）と想定。
- 下記のメリット、デメリットを踏まえ、NCGMバイオバンクに蓄積することを想定。
- 「主として運用するバンク」と「検体を保管するバンク（バックアップ）」との複数箇所に分ける必要がある。
- 統合IDを付与して管理

方針	既存バイオバンクへ業務委託
メリット	・新規建屋の整備が必要ない ・トレーニング済み作業員がいて、新たな雇用は必要ない可能性 ・既存の作業管理と動線が応用可能 ・バックアップも既存バイオバンクへ委託可能 ・倫理対応、施設規則も既存の小修正で可能
デメリット	・既存の作業管理と動線を修正する必要性が生じる可能性

課題 1 : バイオバンクにおける臨床検体の保存・管理の方針、運用方法
○検討事項 2 委託する既存バンクに求められる要件

項目	理由
施設: ・先行解析のために1万人分の検体を収納可能 ・自動保管システムを有する ・将来の検体移管の可能性にも対応可能 ・予算総額が新設バンクより安価	収容能力の必要条件 国際標準を満たすために必須 考慮すべき重要な条件 考慮すべき重要な条件
人材: ・トレーニング済みの作業員が従事(本事業のための作業をする余力がある)	作業員を新規雇用してトレーニングする必要がない
運営管理等: ・既存システムの他に本事業のシステムに対応した制度設計ができる ・研究班の運営方針に従う ・研究班のシステムに対応可能(チューブ、データシステム等) ・以上を安定して永続的に実施できる	運営管理上の必須条件 本事業のためにSOP等の精度管理ができる余力が必要 保管資材の規格やシステムを全施設であわせないと利活用に障害、臨床データと検体の紐付けが必要 運営管理上の必要条件

課題 1：バイオバンクにおける臨床検体の保存・管理の方針、運用方法

○検討事項 3 バイオバンクに保存する検体の量・種類等

- 新たに構築するゲノムデータ基盤は、データの利活用のみならず追加解析等に耐える拡張性の高い基盤であることが求められている。
- このため、複数の利用者に分譲することを想定すると一定の検体量を保管した方がいいが、患者への侵襲や保管・管理に関する負担の検討が必要。
- これらの状況を踏まえ、まずは、一般のバイオバンクにおいて備蓄されている検体量＊分譲を参考に、患者及びその両親一人あたりの検体必要量及び検体の種類については表 1 の通り。
- お願い事項：既存検体を提出する際には表 1 をご覧ください。

表 1 検体の種類と保管検体本数

検体	分注量	1検体の分注本数
血漿＊	1mL/本	4本
DNA	20ug/本	4本

＊用途：プロテオーム、メタボローム、遊離RNA、エクソソーム由来RNAなど
ただし、高品質で収納、保管する作業は容易ではない

「ゲノム基盤」の担当課題

課題1 臨床検体の保存・管理の方針、運用方法

- 検討事項1 バイオバンクの構築方針について
- 検討事項2 (既存バンクに委託する場合) 当該バンクに求められる要件
- 検討事項3 バイオバンクに保存する検体の量・種類等について

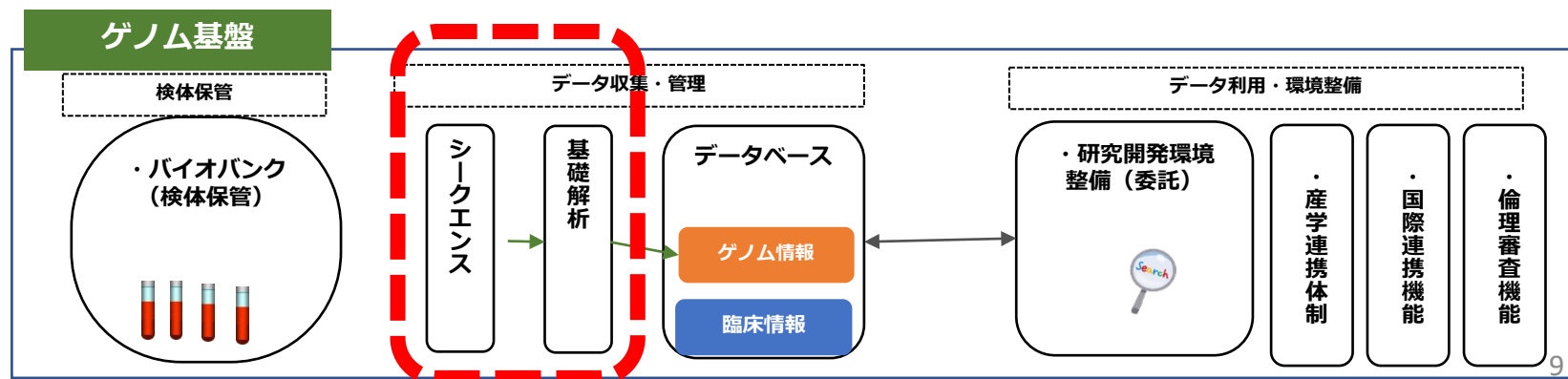
課題2 シークエンス・基礎解析の方針、運用方法

- 検討事項4 ゲノムデータ基盤のシークエンス解析部門に求められること
- 検討事項5 シークエンス解析に関する委託・自主運営の比較について

課題3 ゲノムデータ・臨床データの保存・管理の方針、運用方法

- 検討事項6 ゲノムデータ保管に必要とされる能力等について
- 検討事項7 パブリッククラウドとオンプレミスの比較について
- 検討事項8 臨床データの保存・管理について

課題4 その他、運営・管理に必要な事項



課題2：全ゲノムシーケンス・基礎解析の方針、運用方法

○検討事項4 ゲノムデータ基盤のシーケンス解析部門に求められる要件

○検討事項5 シーケンス解析に関する委託・自主運営の比較

○ シーケンス解析部門の主な業務

- DNAサンプルを受領
- WGS解析（シーケンス＋基礎解析）の実施
- データの管理部門にゲノムデータを送る
- アウトプットはfastq

○ シーケンス解析部門に求められる能力

- プロジェクト全体で統一プロトコルに沿った解析を実施する能力
- 解析を十分量行う能力（2,000検体/月以上：年間10,000検体を想定してピーク対応）
- 検体・工程管理を着実に行う能力（システム導入等）
- 個人情報管理能力取り扱い（ポリシー・対策）を定め、実施している
- 継続的な運営できる経営的基盤

○ シーケンス解析の外部委託を原則とする

現状では、Illumina社のNovaseq 6000が最適プラットフォーム

統合IDを付与して検体管理

○ データ解析結果は拠点に報告する

「ゲノム基盤」の担当課題

課題1 臨床検体の保存・管理の方針、運用方法

- 検討事項1 バイオバンクの構築方針について
- 検討事項2 (既存バンクに委託する場合) 当該バンクに求められる要件
- 検討事項3 バイオバンクに保存する検体の量・種類等について

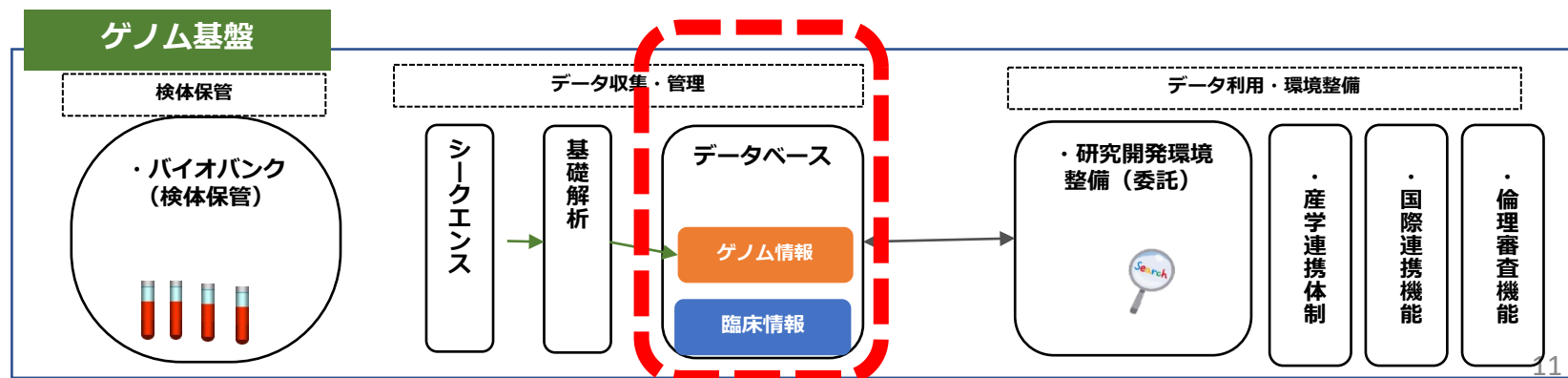
課題2 シークエンス・基礎解析の方針、運用方法

- 検討事項4 ゲノムデータ基盤のシークエンス解析部門に求められること
- 検討事項5 シークエンス解析に関する委託・自主運営の比較について

課題3 ゲノムデータ・臨床データの保存・管理の方針、運用方法

- 検討事項6 ゲノムデータ保管に必要とされる能力等について
- 検討事項7 パブリッククラウドとオンプレミスの比較について
- 検討事項8 臨床データの保存・管理について

課題4 その他、運営・管理に必要な事項



課題3：ゲノムデータ・臨床データの保存・管理の方針、運用方法

○検討事項6 ゲノムデータ保管に必要とされる能力等

○ データ収集・管理部門に求められる能力

- 2次解析はデータ収集・管理部門で実施
- fastqの2次解析（マッピング・バリエーションコール）により、CRAM・VCFなどの解析データを産出
- WGS解析の品質を管理し、必要に応じてWGS解析主体にフィードバック
- 解析数の増加とともに柔軟に容量を増やすことが可能なストレージシステム（1万検体あたり最低1PBは必要）
- データの分散配置やバックアップなど災害時にデータが失われない対策を施す
- 2次解析に必要な計算リソース（CPU/GPUクラスター）
- 個人情報の取り扱い（ポリシー・対策）を定め、実施
- 外部の解析リソースにネットワークで直結しているなど、保管データを円滑に解析に用いるための仕組みを備える
- 利用者が必要な検体を検索し、簡便に利用できるサービスを提供する

課題3：ゲノムデータ・臨床データの保存・管理の方針、運用方法
○検討事項7 パブリッククラウドとオンプレミスの比較

- ゲノムデータ保管先として想定されるパターンとしては、①パブリッククラウド、②オンプレミス、③双方の組み合わせの3つが想定される。
- 利用者が多いデータ提供後のサービスはパブリッククラウドの方が運用面のメリットが大きい。一方、1次的なデータの保管を自機関で行うためにはオンプレミスサーバーも必要。
- 難病の先行研究では1,000検体程度で試運用して比較することを想定

	パブリッククラウドを利用	オンプレミスサーバーを自主運営
メリット	・利用料に応じた出費で、支払先は単独なので管理が容易 ・必要に応じてスケールを調整可能 ・必要に応じて計算資源をほぼ無制限に利用可能 ・分散配置やバックアップが容易	・データ保管を機関内で完結させられる ・セキュリティなどの細部にわたりコントロールできる ・長期保管する場合はパブリッククラウドよりもコストを抑えることができる
デメリット	・クラウドのみにデータを保管することの可否 ・データの取り出しに費用が発生する	・初期費用が比較的高額 ・保守費用・人件費・光熱費など出費項目が多く管理が煩雑 ・バックアップ(特に複数の拠点に分散した場合)が困難 ・計算資源は別途調達が必要

課題3：ゲノムデータ・臨床データの保存・管理の方針、運用方法

○検討事項8 臨床データの保存・管理

○ 臨床情報の保管に必要な要件

- ・ 監査機能など高度なセキュリティを担保
- ・ （画像データを対象としない限り、サイズはゲノムデータほど大きくはない）
- ・ 個人情報の取り扱い（ポリシー・対策）を定め、実施
- ・ 高度な検索機能（データベース）を有する
- ・ データの分散配置やバックアップなど災害時にデータが失われない対策が望ましい

○ 臨床情報の保管のあり方

- ・ ゲノムデータとはデータ構造・サイズが異なり、個人特定性も大きく異なることから同時に参照できない異なるストレージに保管
- ・ 利用者側でIDによるゲノム情報との結合を行って解析を行う
- ・ ID体系・検索システムを整備

「ゲノム基盤」の担当課題

課題1 臨床検体の保存・管理の方針、運用方法

- 検討事項1 バイオバンクの構築方針について
- 検討事項2 (既存バンクに委託する場合) 当該バンクに求められる要件
- 検討事項3 バイオバンクに保存する検体の量・種類等について

課題2 シークエンス・基礎解析の方針、運用方法

- 検討事項4 ゲノムデータ基盤のシークエンス解析部門に求められること
- 検討事項5 シークエンス解析に関する委託・自主運営の比較について

課題3 ゲノムデータ・臨床データの保存・管理の方針、運用方法

- 検討事項6 ゲノムデータ保管に必要とされる能力等について
- 検討事項7 パブリッククラウドとオンプレミスの比較について
- 検討事項8 臨床データの保存・管理について

課題4 その他、運営・管理に必要な事項

各班員と連携して今後1年を目処に整備

