

## 新学術領域「ゲノム支援」

### 2011年度支援課題公募説明会

---

### 「ゲノム支援」発足の経緯

---

- 「ゲノム支援」は文部科学省科学研究費新学術領域研究(研究領域提案型)『生命科学系3分野支援活動』に採択されたものです。
- 提案課題名は「ゲノム科学の総合的推進に向けた大規模ゲノム情報生産・高度情報解析支援」ですが、以降「ゲノム支援」と略称します。
- この新たな仕組みは、特定領域研究の終了を受けて、生命科学系3分野(がん、ゲノム、脳)の今後の支援のあり方について科学技術学術審議会の部会で審議が行われ、「従来の3分野に関する特定領域研究における総括班・支援班の果たしてきた役割を何らかの形で継承し、新たに3分野を支援する仕組みを措置すべき」とのまとめに基づいて設計されたものです。

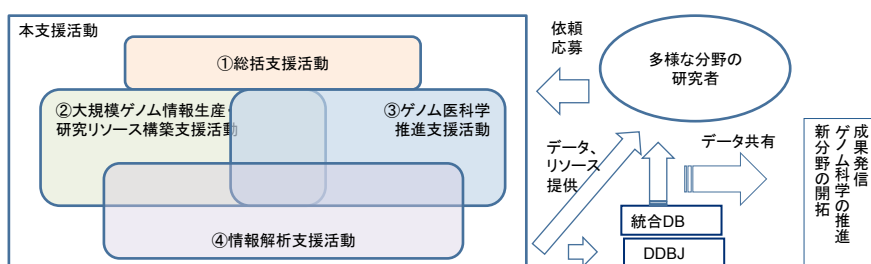


## 支援活動の基本的な考え方

高度かつ効果的な支援のために：

- 先端技術・システム開発研究を推進
- 人材育成(特に情報解析)に注力
- 分野融合の体制作り
- ゲノムを核とした研究コミュニティ作り
- オールジャパンの協力体制
- 国際的な窓口機能
- ELSI、社会との接点活動

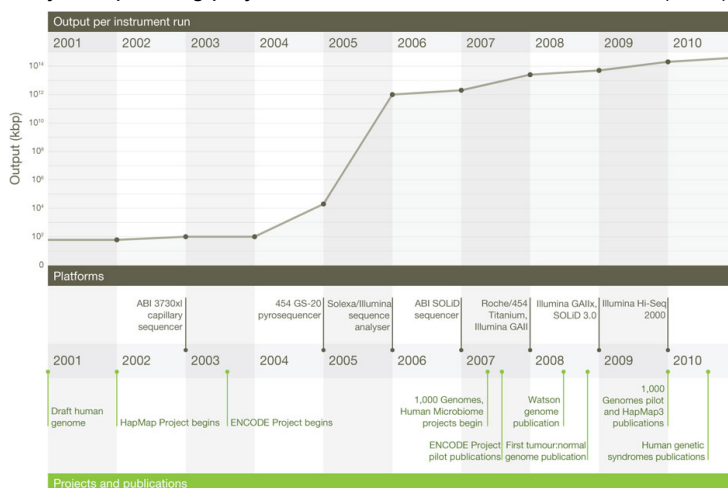
## 新学術「ゲノム支援」全体構成



- 4つの支援活動で構成
- 現ゲノム特定の基盤ゲノム、支援班を継承
- 支援方針、課題選考、全体戦略は総括支援活動
- 他3支援活動が相互に連携して支援
- 支援の成果は統合DB/DDBJ等を通して共有化を図る

## DNAシーケンサの進化

Changes in instrument capacity over the past decade, and the timing of major sequencing projects. ER Mardis. *Nature* **470**, 198-203 (2011)








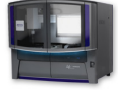
## 支援チームのデータ生産能力

### シーケンス拠点

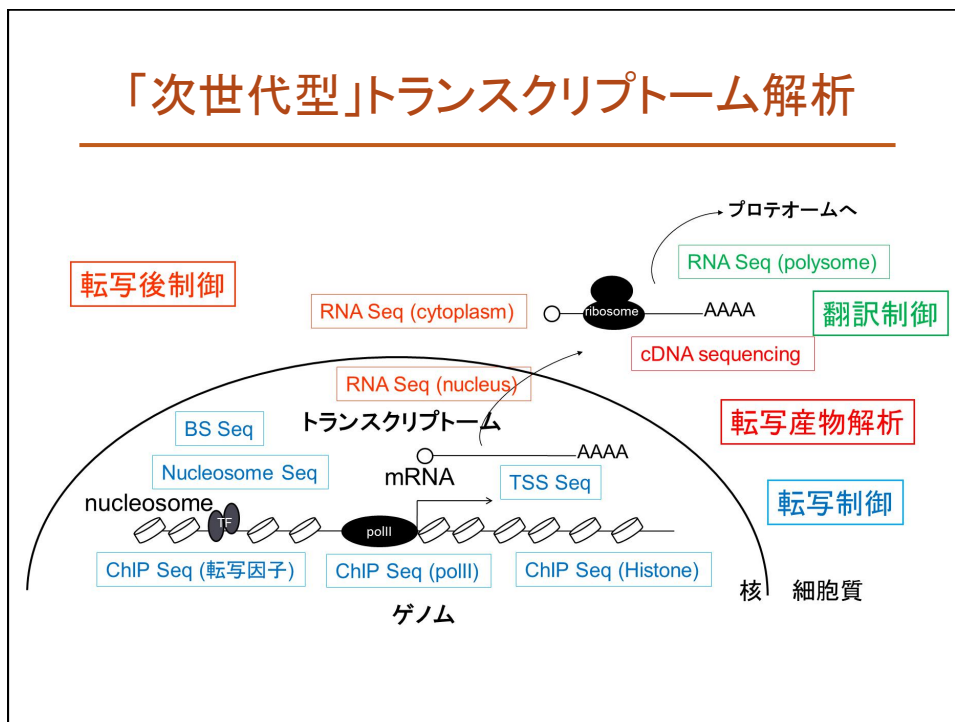
- ◆ 国立遺伝学研究所 多細胞（ゲノム）、ゲノム医学
- ◆ 東京大学新領域 微生物（メタゲノムを含む）、多細胞（RNA, ChIP-Seq 等）
- ◆ 宮崎大学 微生物（病原微生物）

機種名	解読長	データ量	ラン時間
Genome Sequencer FLX	400 bp	0.5 Gb	10時間
Genome Analyzer Iix	150 bp	90 Gb	14日
HiSeq2000	100 bp	200 Gb	8日
SOLiD4	50 bp	100 Gb	14日
5500xl SOLiD	75 bp	180 Gb	7日
ABI 3730xl	800 bp	77 kb	2時間

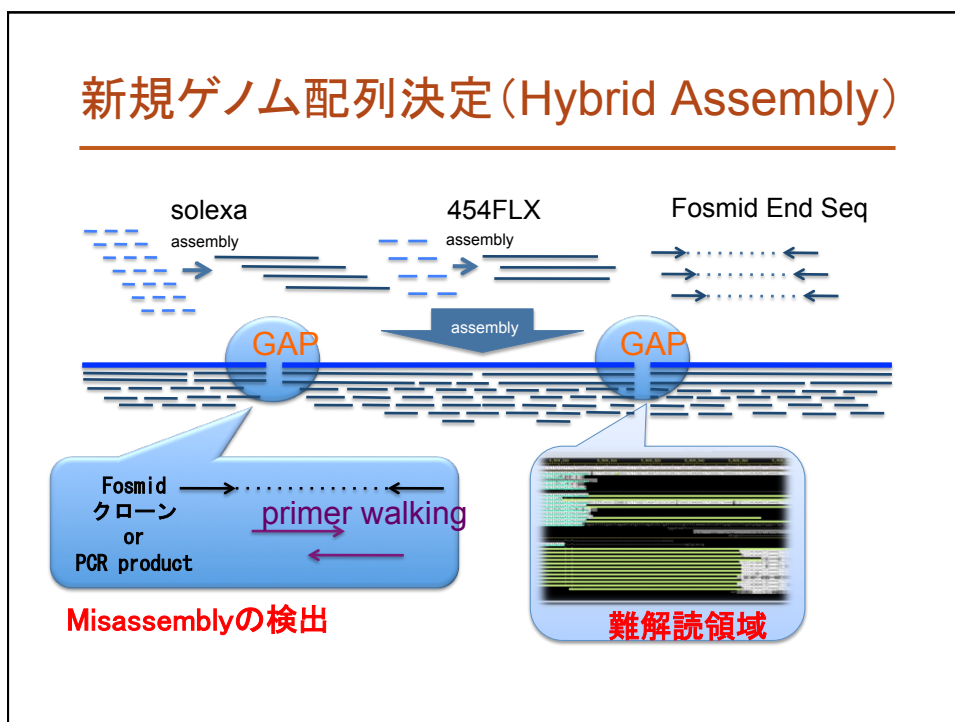
  

 <b>Illumina GAIIx</b> 3台	 <b>Roche GS FLX:</b> 4台
 <b>Illumina HiSeq2000:</b> 3台	 <b>ABI 3730xl:</b> 20台
	 <b>SOLiD4:</b> 1台
	 <b>5500xl SOLiD</b> 2台

## 「次世代型」トランスクリプトーム解析



## 新規ゲノム配列決定 (Hybrid Assembly)



## 新学術領域「ゲノム支援」

### 2011年度支援課題の公募要領

---

### 支援の対象となる研究課題

---

- ・2011年度文部科学省・科学研究費補助金(新規・継続)に採択されている研究課題を対象とします。
- ・本ゲノム支援活動により、当初の計画を上回る重要な研究成果を得ることが期待でき、その結果、我国のゲノム科学自体の発展を促すことにもつながるような支援申請を優先します。
- ・支援の申請内容が科研費の研究当初計画に含まれていることは必要ありませんが、密接に関連することは必須です。
- ・科学研究費のすべての研究種目を対象とし、その研究課題の研究代表者あるいは研究分担者からの申請を受け付けます。
- ・個人・機関からの応募件数の制限はありませんし、科研費の”重複応募制限”とも無関係です。
- ・新学術研究「ゲノム支援」に配分された科研費で、「ゲノム支援」班員が行うもので、支援依頼者に研究費を配分することはありません。

## 支援によりえられた 研究リソースとデータの公開、共有

- 研究の結果得られるデータ量が大規模化してくると、個々の研究者が解析するだけでなく、研究コミュニティ全体で共有することにより有効活用が図れます。
- 支援活動による解析結果は、論文・データベースなど適切な方法で公表してください。塩基配列データは、適切なデータベースに登録・公開することとします。
- ライブラリなどの研究リソースは、論文発表後、研究コミュニティへの配布体制の整備、レポジトリへの登録等により、公開することとします。
- 個人ゲノム情報を含まない塩基配列データは、支援依頼者に提供すると同時に、DDBJあるいはDRA(DDBJ Read Archive; 新型シーケンサーのデータ)に仮登録します。そして、論文発表後は直ちに、未発表の場合は事前に協議した時期(支援終了後原則として1年以内)に公開することとします。
- 個人ゲノム情報についても、その質に応じた適切な個人情報の保護の仕組みの下で、研究コミュニティでの共有を図ります。

## 支援申請で留意していただきたい点

- 原則として、支援依頼内容は単年度で実施可能なものとなるようにしてください。また、不必要に過大な内容の申請は避けてください。なお、採択された場合でも、翌年度にさらなる支援の申請を行うことは可能です。
- 支援活動は、支援依頼者と支援担当研究者の共同作業と考えて下さい。支援活動により得られた研究成果について、支援担当者への迅速なフィードバックをお願いします。
- 大型ゲノムの新規配列決定など、支援をする側にとってもチャレンジングな課題についても挑戦したいと考えています。しかし、そうした課題の数は限られるものとなります。
- 少しでも多くの課題を支援するために、依頼者に消耗品などの一部負担をお願いすることがあります。
- 支援に当たっては最善の努力を払いますが、成功を保証するものではありません。当初に期待したとおりのデータが得られない場合もあります。
- 支援グループから提供するデータや研究用リソースの内容について、必ず支援依頼者側で最終的な確認をして下さるようお願いいたします。

## 大規模ゲノム情報生産支援活動

### A. ゲノム配列解析

- ・新型シーケンサによる配列決定
- ゲノム配列の新規シーケンスまたは再シーケンス(Sequence Capture法による部分ゲノム再シーケンスも含む)
- メタゲノム解析(メタ16S rRNA解析を含む)
- エピゲノム解析(メチル化解析、ヌクレオソーム解析等)
- ChIP-seq解析
- ・その他
- BAC/fosmidライブラリの末端配列決定
- 個別クローンの全配列決定(BAC/fosmidクローンに限ります)

### B. トランスクリプトーム解析

- 新型シーケンサによるRNA-seq解析
- ESTライブラリの末端シーケンス決定

### C. ライブラリ作製(BAC/fosmidライブラリ作製)

- Sheared FOSMIDライブラリ
- オリゴキャップ法による完全長cDNAライブラリ

## ゲノム情報生産支援での留意点

- ・配列決定に用いるDNA試料の調製は、依頼者側で実施していただきます。解析に必要なDNA量とその質について、支援担当グループと協議していただきます。
- ・ChIP-Seq解析については、既知のタンパク質結合領域について、4-8倍程度以上のDNAの濃縮がreal time PCRで確認されていることが必要です。
- ・RNA-Seq解析のためのRNA調整は、依頼者側で実施していただきます。その後の、配列解析のための前処理については、支援担当グループと協議していただきます。なお、期待される転写産物が含まれていることのreal time RT-PCRによる確認などの、基礎的検討をお願いします。
- ・波形や画像データからの配列データへの変換と品質スコアの付記などの前処理、シーケンサ付属または支援活動拠点が保有する解析パイプラインによる定型的なマッピングやアセンブルなどの一次的解析までを行います。
- ・配列データの高度化(ギャップフィリングや高精度配列決定)、アノテーション、高度な統計処理などの個別研究的な要素が大きい解析は、希望に応じて共同研究として実施するか情報解析支援活動グループの班員を紹介します。

## 情報解析支援活動

### A. 一次的解析の支援

(一般的に利用可能なソフトウェアを既定のパラメータで使用する解析)

- 概要配列決定: マッピング、アセンブル、大規模相同性検索、アノテーション
- 微生物ゲノム解析、病原性微生物ゲノム解析: アセンブル、大規模相同性検索、アノテーション、大規模系統解析
- メタゲノム解析(メタ16S rRNA解析): メタゲノム遺伝子予測、大規模相同性検索、アノテーション、大規模系統解析
- 比較ゲノム解析: 自動オーソログ抽出、ゲノムアラインメント
- ゲノム再シーケンス、多型解析
- ChIP-seq解析: マッピング、領域 & ピーク検出

### B. より高度な情報解析支援

- 大規模な一次解析や複数のソフトウェアを組み合わせた情報解析について、解析パイプライン構築の支援を行います。
- 大規模相同性検索、アセンブル、遺伝子予測、マッピングなどのために新たにソフトウェア開発を必要とするもの、高度な配列解析、高精度アノテーション、解析結果の統計処理など、個別研究的な要素が大きい場合は、希望に応じて情報解析支援活動グループの班員との共同研究を紹介します。

## ゲノム医学支援活動

### A. ゲノムワイド関連解析(GWAS)

- ゲノムワイドSNP タイピング(一次スクリーニング)
- GWAS後の大規模2次スクリーニングあるいはreplication study

### B. 家系由来試料の連鎖解析

- パラメトリック連鎖解析
- モデルフリー連鎖解析(罹患同胞対解析など。タイピングおよび一連の連鎖解析を含む)

### C. 大規模ヒトゲノム再シーケンシング

- 超高速シーケンサを用いてexome/target sequence captureまたは全ゲノムを対象に大規模再シーケンシングを行い、希望に応じて、一次的な解析(マッピング、SNV、short in/del の検出、variationの一覧など)、さらに、variationについてのfunctional annotationなどの医学・生物学的な解釈、統計学的解析などを実施します。

### D. 新型シーケンサを用いたゲノム機能解析

- RNA-seq解析
- ChIP-seq解析
- エピゲノム解析

## ゲノム医学の解析結果の公開と共有

### A. GWASデータ

- サマリーデータ: SNP解析の結果について、頻度情報、統計解析結果などのサマリーデータは、バイオサイエンスデータベースセンターを通して公開します。
- サマリーデータ以外のSNP解析データ: 制限付きアクセスを条件に、バイオサイエンスデータベースセンターを通して他の研究者が共有できるようにします。共有できる内容は、SNP データ、疾患名(あるいはコントロール)とします。

### B. ヒトゲノムシーケンスデータ

- 頻度データ: 支援活動によって得られたすべてのゲノム配列から、variationについて頻度情報を整理し、この情報をバイオサイエンスデータベースセンター上で公開します。
- 日本人ゲノムの参照配列: すべてのゲノム配列から、日本人ゲノムのコンセンサス参照配列の作成を進めます。十分な数のゲノム配列が得られ、個人識別を可能にする情報は含まれないと判断される段階で公開します。
- 個別配列情報: 制限付きアクセスを条件に、バイオサイエンスデータベースセンターを通して、他の研究者が共有できるようにします。共有できる内容は、ゲノム配列、疾患名(あるいはコントロール)とします。

## ゲノム医学支援の要件 (1) IC取得と倫理審査委員会審査

### A. GWAS、ゲノム再シーケンシング

- 研究協力者の方への説明同意文書において、通常の項目に加えて、全ゲノムを対象とする解析を行うこと、解析結果を公的データベースに登録し、当該研究だけでなく様々な疾患の研究に役立てることを記載してください
- 所属機関の倫理審査委員会へは、共同研究機関としてのゲノム支援側情報、解析結果のデータベース登録とその公開・共有方針の詳細等を記載した、研究計画書を提出し、承認を受けてください。
- 説明同意文書、研究計画書の要件の詳細については、公募WEBページに掲載します。また、説明同意文書の雛形も掲載します。
- すでに収集された試料を解析する場合、GWASのため取得された説明同意文書において、データベースにおける公開・共有が明記されていない場合、少なくとも、改めて所属機関の倫理審査委員会でデータの公開・共有について承認を得る必要があります。ゲノム再シーケンシングの場合、説明同意文書が上記の項目を反映していない場合、原則として、再同意の取得をお願いします。

## ゲノム医学支援の要件(2)

### B. 家系由来試料の連鎖解析

-説明同意文書は、通常の疾患遺伝子探索研究として取得してください。なお、データベースへの登録は想定していません。

### C. 新型シーケンサを用いたゲノム機能解析

-発現プロファイル解析等は、本来、三省ゲノム指針の対象ではありません。しかし、例えば、新型シーケンサを用いたトランスクリプトーム解析では遺伝子領域の配列情報がデータに含まれるようになるため、ゲノム配列解析に準じた扱い(研究協力者への説明と同意、倫理審査委員会の承認)をお願いします。

なお、支援の実施にあたり、支援実施者所属機関の倫理審査委員会の承認も要であり、支援の実施は、その承認後になります。

## ゲノム医科学支援の体制

検体や解析に必要なデータは、支援拠点における解析を行う必要上、「連結可能匿名化」の状態を受け取ります。ただし匿名化の対応表は、依頼機関において適切に保管をしていただき、支援側では対応表は受け取りません。

個人情報の保護が必要なシーケンシングは、生体認証によって入室管理が行われている場所に設置され、インターネット接続されていないシーケンサを使用します。データ解析については、入退室の管理が行われている部屋に設置され、インターネットに接続されていないコンピュータを使用します。

ヒトゲノム配列データについては、一定の管理体制(※)が整備している場合にのみ、支援依頼者に提供します。それ以外の場合は、サマリーデータのみを提供し、解析は支援拠点で行います。

(※)ヒトゲノム配列データの管理体制: 個人ゲノムを解析する場合、入退室の管理が行われており、データ解析にはインターネットに接続されていないコンピュータを用いる体制

## 支援申請の方法

---

2011年度第1回の募集期間:2011年4月1日(金)~4月25日(月)

### 1. 申請者情報のWEBでの入力

電子メールでIDとパスワードの取得方法をお知らせいたします。  
以後の手続きは、IDとパスワードで応募ページにログインして行ってください。

### 2. 科研費情報のWEBでの入力

### 3. 支援内容概要のWEBでの入力

### 4. 申請書のアップロード

### 5. 研究計画書、説明同意文章のアップロード

### 6. 申請内容の確認

## 申請書の作成について

---

### ・科研費課題の研究概要

支援申請内容と強く関係する**科研費研究課題**について、研究目的(概要)、研究目的を2ページ以内で記入してください(当該研究計画調書内容そのままでも結構です。**支援希望内容ではありません**)

### ・支援希望内容とその理由等

- ①「ゲノム支援」活動による支援を希望する具体的な内容  
対象とする生物のゲノムサイズ、ChIP-seq・RNA-seqについては試料数、ゲノム医学支援における検体数等、**なるべく定量的な記述をお願いします。**
- ②本支援の結果、どのような研究の展開が期待できるのか。そのための研究計画と研究体制
- ③ゲノム科学、ゲノム医学研究としての支援申請内容の重要性と国内外の研究状況
- ④支援を受けるための**材料等の準備状況**、**支援結果活用のための準備状況**  
シーケンス解析のためのDNA、RNA調整の準備状況、得られたシーケンスデータの情報学的解析やゲノム配列のアノテーションの計画等を記述してください。

皆様の申請をお待ちしています

---

新学術領域「ゲノム支援」事務局

<http://www.genome-sci.jp/>

お問い合わせ

<https://www.genome-sci.jp/contact/contact.php>