

4号：遺伝子解析特集

～次世代シーケンサーの基本原則と新たな技術の紹介～

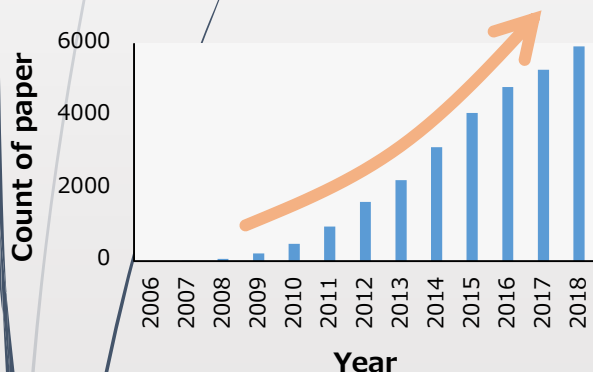
INDEX

- はじめに
- NGS解析全体の流れ
- 各ワークフローの基本原則
- 新たな遺伝子解析機器・技術の紹介

はじめに

2005年に454 Life Sciences社が世界初の次世代シーケンサー(NGS, Next-Generation Sequencing)をリリースしてから13年が経ちました。その間に各社が独自のNGSの開発と技術改良を繰り返し、測定コストが大幅に減少したことから、現在では幅広い分野の医薬生物学において欠かせないツールとなりました。実際にNGSを用いた論文数は年々増えており、その過程で“RNA-seq”や“ChIP-seq”などの代表的手法が確立され、多くの研究者により用いられております。本編では、まだNGSに馴染みのない方やこれからNGSを使ってみたいと考えている研究者を対象として、NGS解析の基本的な流れとその原理、新たな遺伝子解析技術を紹介いたします。

NGSを用いた研究報告



「PubMed」で“next-generation sequencing”で検索した論文数 (当社調べ)

NGSで用いられる代表的な手法

手法	目的
Whole Genome-seq	変異の探索 疾患原因遺伝子の特定 ゲノム多様性の解析
Exome-seq	
RNA-seq	網羅的発現量解析 新規転写産物同定
Small/micro RNA-seq	Small RNA, microRNAの定量解析
ChIP-seq	タンパク質結合領域の同定 ヒストン修飾状態解析
Bisulfite-seq	CpG領域のメチル化の網羅解析
Hi-C	ゲノム立体構造解析
16S rRNA amplicon-seq	細菌種の同定 メタゲノム解析

NGS解析全体の流れ

NGSの作業は基本的に4つの工程で構成されています。①サンプルから核酸(DNA/RNA)を抽出する工程、②シーケンスプラットフォームに持っていくために核酸を加工するライブラリー調製工程、③塩基配列を解読するシーケンス工程、④得られた塩基配列を解析用フォーマットに変換し、データベースを照合する解析工程です。

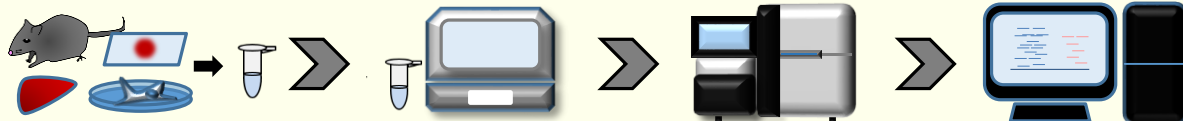
NGSを使ったアプリケーションとして発現解析、変異解析、メチル化解析、構造解析など様々ありますが、それぞれによって各工程での必要な作業と用いる試薬が異なってきます。

核酸抽出・前処理

ライブラリー調製

シーケンス

解析



NGSワークフロー

●ライブラリー調製

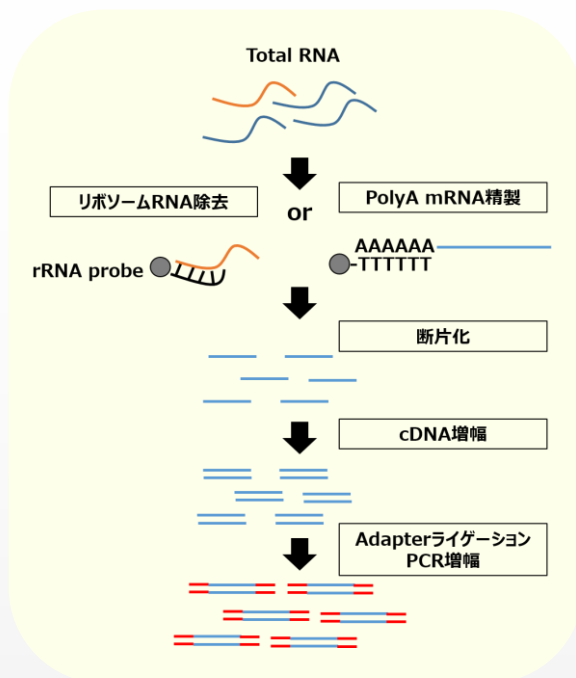
ライブラリー調製の目的は、シーケンスプラットフォームで解析できるよう核酸を加工することです。この作業は解析手法によって異なります。今回は真核生物における発現解析（トランスクリプトーム解析、RNA-seq）を例に紹介します。

① まずサンプルからtotal RNAを精製し、mRNAリッチな核酸を準備します。主に2通りの手法があり、1つめはrRNAを除去・枯渇する方法（ビーズによる分離や、酵素など）、2つめはmRNAのポリAを利用した精製方法です。

②次に精製したmRNAを数百bp程度に断片化します。これには酵素を用いる方法、機械的断片化する方法があります。いずれにしてもランダムに断片化し解析にバイアスがかからないようにすることが重要です。

③断片化したRNAを逆転写によりcDNA化し、NGSプラットフォームに合わせたアダプターをライゲーション反応により付加します。

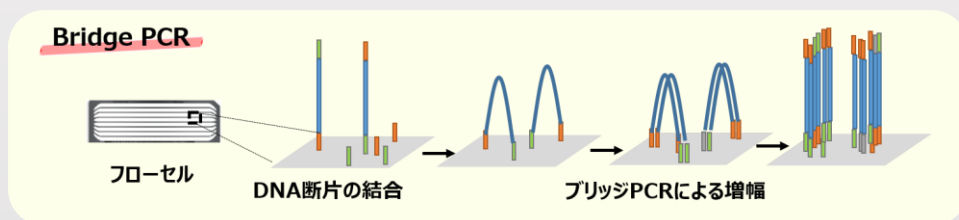
各NGS試薬メーカーから様々な試薬キットが販売されており、各試薬によって反応の原理は異なります。解析目的に合ったキットを選定する必要があります。



●シーケンス (Illumina社)

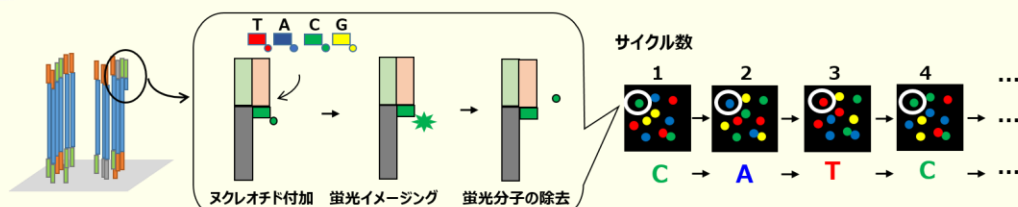


Illumina社のHiSeqやMiSeqでは、ブリッジPCRという方法を用いています。フローセルと呼ばれるスライドガラス上には、あらかじめアダプターと相補的に結合するプライマーが高密度に配置されており、サンプルのDNAはアダプターとフローセル上のプライマーと相補的に結合して橋（ブリッジ）がかかったような構造になります。この状態でDNAポリメラーゼによる伸長反応の後に変性させ、この反応を繰り返すことで、狭い面積の中で1本鎖DNAを固定しながら増幅することができます。これが塩基配列の解析の鋳型となります。



塩基配列の解析には、蛍光標識したdNTPの取り込みを、蛍光の色を読み取ることによって解析します。このdNTPは3'末端がブロックされており、1回の伸長反応で1塩基しか伸ばせず、1塩基ごとにどのdNTPが取り込まれたかを検出し、その後に蛍光物質とブロックを外して次の伸長反応を行います。このステップを繰り返し行い解析を進めていきます。このように断片の相補鎖を合成しながら配列を決定することからsequencing by synthesis (SBS)法と呼ばれています。

Sequencing by synthesis (SBS)法

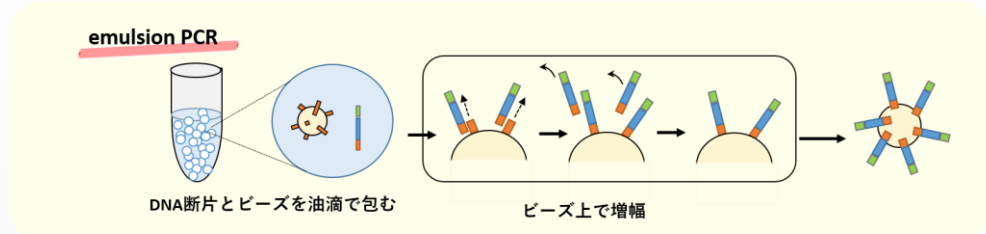


1塩基伸長反応と蛍光シグナルの取込み・検出のサイクル

●シーケンス (Thermo Fisher Scientific社)

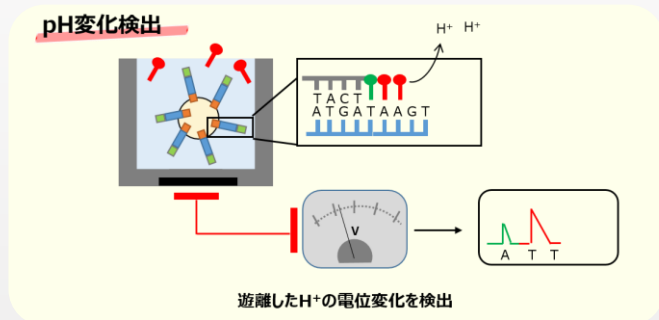


Thermo Fisher Scientific社のIonTorrentシーケンサー (Ion PGM、Ion S5など)の場合、増幅反応はエマルジョンPCRによって行います。アダプターと相補的な短いDNAが結合したキャプチャービーズと、サンプルの1本鎖DNAとが1:1で結合するように混合し、増幅試薬とともに油中水滴エマルジョンに内包させ、オイル中にビーズひとつとDNA断片ひとつだけを持つマイクロリアクターを形成します。これにより、各DNA断片は他の配列が混ざることなく、ビーズ上で増幅することができます。そしてエマルジョンを破壊してビーズを濃縮し、プレート上に載せて配列解析を行います。



IonTorrent技術は、解析に光を用いず、半導体による電位検出によって塩基を判定するシステムであることが特徴です。塩基の伸長反応の際に、水素イオン(H⁺)が遊離し、このH⁺が半導体に与える電位変化を検出しています。同じ塩基が連続する場合、その塩基の数だけ伸長反応が進みますが、H⁺は同じ塩基が連続する数だけ放出されるため、その分電位変化が大きくなり、1回のサイクルで増えた塩基数を判断できます。

また本プラットフォームでは蛍光や発光検出する従来の装置と異なり、スキャナやカメラ、光源などを必要とせず、さらにシーケンス反応を行う半導体内でデータ処理も同時に行えるため、生データ処理のためのコンピューターを別途用意する必要がなくなっています。これにより装置などのコストが安く抑えられていることもメリットと言えるでしょう。



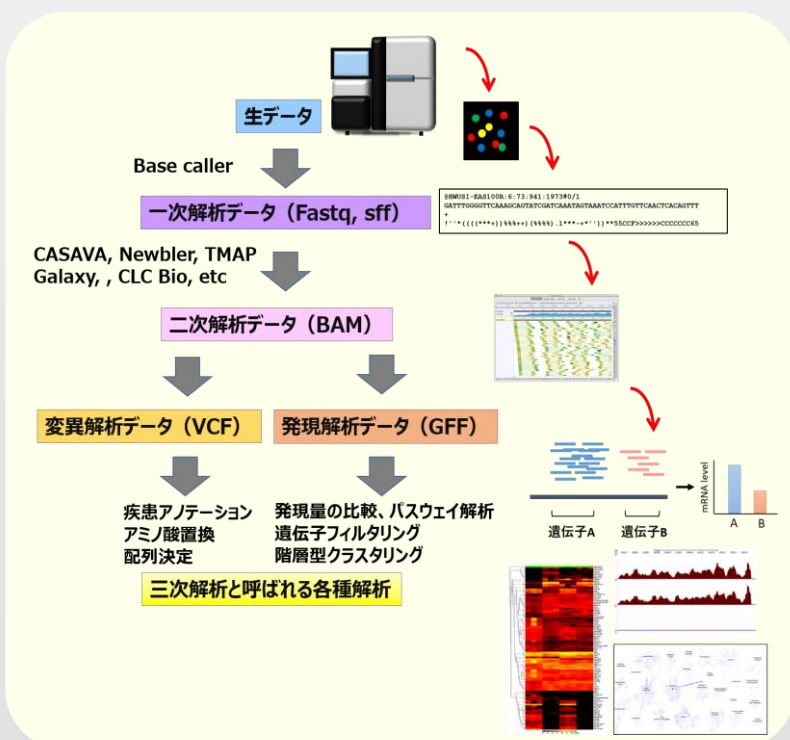
●解析

シーケンスで得られたシグナルは機器に搭載されているBase Callerによって塩基に変換され、FASTQというファイル形式にて保存されます

(Illumina社NGSの場合)。この状態では、ただ多量の塩基配列が並んだだけのデータであり、意味を成しません。この情報を元に公共データベースのリファレンス配列を照合(マッピング)することで解析を行っていきます(二次解析データ、BAM file)。この状態でマッピングがモニター上で可視化することができます。

ここからデータに意味づけ(アノテーション)を行います。

例えば発現解析では、マッピングされたリード数をカウントすることで遺伝子の発現量を比較し、変異解析ではマッピングされた配列情報とリファレンス配列との違いを判定することを行います。解析目的に合わせてアプリケーションを選択することが重要になります。



非常にざっくりと記載しましたが、解析は1ステップで完了とはいかず、解析に合わせたファイル形式の変換を行い、さらには各解析ごとに適した解析アルゴリズムの調整など操作が複雑であり、専門家と相談しながら進めていくことが推奨されます。

これまで紹介したNGS機器は、数百bpの塩基を並列解析することで大量の塩基配列情報を得る、いわゆるショートリードNGSです。それに対し、Oxford Nanopore Technologies(ONT)社の「MinION」「GridION」、PacBio社の「RS II」「Sequel」はリード長がONT社は数100kbp、Pacific Biosciences (PacBio)社では10kbp以上ものロングリードを可能とします。特にONT社は、膜上に配置されたタンパク質微細孔（ナノポア）をDNA分子が通過する際の電流変化によって配列決定を行うナノポアシーケンサーで、従来のシーケンサーのようにDNA伸長反応を利用しないという点で他と一線を画していると言えるでしょう。一方PacBio社の方は、wellに固定されたDNAポリメラーゼが鋳型DNAを複製する際、蛍光標識された塩基が取り込まれ、伸長反応ごとに標識リン酸基が遊離します。この遊離蛍光を検出することにより、伸長反応を進めながらリアルタイムに塩基の検出を行います（SMRTDNAシーケンシング技術）。

このロングリードシーケンサーにより、ショートリードNGSでは解析が難しいとされるゲノムの大規模な構造変異（欠失・挿入・転座・融合）やフェーシング情報を得ることができます。

最近注目されている新たな技術としては、10x Genomics社が開発した「GemCode Technology」と呼ばれるNGS前処理技術があります。分子バーコードを持つオリゴがコートされたゲルビーズと高分子ゲノムDNAを数百万個の液滴(GEM)に分割する技術であり、GEM毎に分子バーコードが異なるため、バーコード配列を利用して、同じ分子由来のショートリードをアセンブルすることで、解析時に合成的にロングリードに変換することが可能です。

また、高分子DNAの代わりに細胞をGEMに分割することで、細胞集団の転写産物を1細胞ごとに網羅的に解析することが可能です（シングルセルRNA-seq解析）。そのため、細胞集団を構成する細胞がどう分類できるかが未知のままでも、細胞集団を亜集団にクラスタリングして特徴を抽出することが可能です。

KPSL 遺伝子受託サービス

KPSLではNGSによる発現解析、ChIP-seq、exome-seq、マイクロアレイ解析など様々な解析サービスのご案内が可能です。現在の研究で遺伝子解析を検討されている方がおられましたら、是非ともKPSLにご相談ください！！

RNA-seq	メッセンジャーRNAを網羅的に測定することで生体内の遺伝子発現変動を網羅的に解析します
ChIP-seq	クロマチン 免疫沈降（ChIP）により精製されたDNAの配列を解析し、転写修飾因子の結合領域の推定やヒストン修飾修を受けている領域の探索などの研究に用いられます
Exome-seq	ゲノムのうち、エクソン配列のみを網羅的に解析します。エキソンは全ゲノムの2%未満ですが遺伝性疾患の多くがエキソン領域の変異によるものと推定されており、効率的に疾患関連遺伝子を解析します。
ctDNA解析	血漿中cell free DNAを抽出し、腫瘍由来のDNAの変異を解析します

【参考資料】

- Illumina社ホームページ(<https://jp.illumina.com/>)
- Thermo Fisher Scientific社ホームページ(<https://www.thermofisher.com/jp/en/home.html>)
- Oxford Nanopore Technologies社ホームページ(<https://nanoporetech.com/>)
- Pacific Biosciences社ホームページ(<https://www.pacb.com/>)
- 10x Genomics社ホームページ(<https://www.10xgenomics.com/>)
- Cui, Peng, et al. *Genomics* 96.5 (2010): 259-265.
- Goodwin, Sara, John D. McPherson, and W. Richard McCombie. *Nature Reviews Genetics* 17.6 (2016): 333.
- Khouk, Mehdi, Jean-François Gibrat, and Mourad Elloumi. *Biology and Medicine* 9.3 (2017).
- Zheng, Grace XY, et al. *Nature communications* 8 (2017): 14049.

KPSL News バックナンバー

Vol.1 プロテオミクス特集

～次世代定量プロテオミクス

『iMPAQ』分析始めました～

Vol.2 メタボロミクス特集

～網羅分析からターゲット分析へ～

Vol.3 高感度マルチタイムノアッセイ特集

ホームページからダウンロード

または配布希望をご連絡ください



【お問い合わせ先】

九州プロサーチ有限責任事業組合

〒819-0388 福岡県福岡市西区九大新町4-1

TEL:092-805-3239 FAX:092-805-3239

MAIL: info@kpsl.jp

←URL: <https://kpsl.jp/>



KPSL
九州プロサーチLLP