

# 次世代シーケンサーによる 微生物ゲノム解析のご紹介

ゲノム解析をより早く・安く・大量に!!!

タカラバイオでは、454 Life Sciences社にて開発された次世代シーケンサーGenome Sequencer 20 (GS20)を用いる「高速シーケンス解析受託」サービスを昨夏よりご提供しています。微生物ゲノムのドラフト解析をはじめ、本法を用いてmicro RNA、EST、ChIPなど各種アプリケーションへの応用も承っています。今回は微生物ゲノム解析に焦点を絞って、本法の有用性をご紹介いたします。

\*GS20はロシュ・ダイアグノスティクス(株)AS事業部が販売しています。

## ■ 微生物ゲノム解析の動向

近年、メタゲノム解析の進展もあって、これまで解析対象とされてきた微生物種は実は微生物全体の1%にも満たないということが明らかとなっています。この膨大な微生物の理解のためにはこれまでに得られているゲノム情報では不十分であるため、さらに数多くのゲノム解析が進行中です。一方、GS20をはじめとする次世代シーケンサーの登場など、シーケンスの技術向上もゲノム解析の増加を後押ししています。従来の解析法(サンガー法)ではコスト的に難しかった、変異株・派生株との全ゲノムシーケンスでの比較解析なども現実のものになっています。

弊社の受託サービスにおいても、GS20導入後の半年間で既に10件以上のご依頼を頂いており、1ゲノム/1研究室という時代の到来を感じます。

## ■ 微生物ゲノム解析例

高速シーケンス解析では、通常、ゲノムサイズの15~20倍程度の冗長度でシーケンスを行います。図1にゲノムサイズが5 Mbの微生物のゲノム解析結果を示します。この生物種では冗長度が15程度(赤点線)でゲノムカバー率、コンティグ数がほぼ飽和に達しています。

表1は、弊社で解析した4種類の微生物ゲノムのドラフト解析結果を示しています。サイズ、配列の特性による結果のばらつきは従来法同様に見られますが、得られるコンティグ数、配列の品質も従来法と同程度であり、十分に高精度なドラフト配列情報が取得できています。

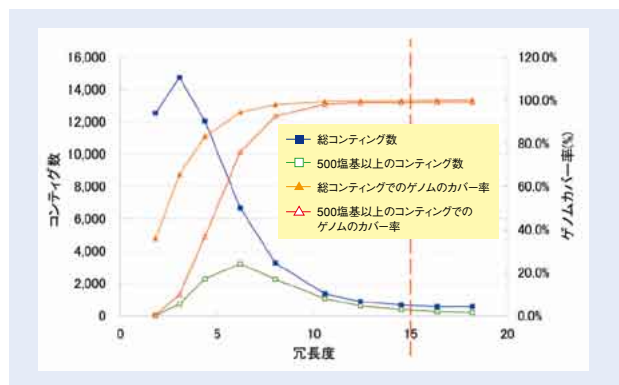


図1 ゲノムサイズ5Mbでの解析結果

表1 GS20によるドラフトシーケンス解析結果

	バクテリア			
	A	B	C	D
ゲノムサイズ	3.38 Mb	2.8 Mb	2.8 Mb	1.9 Mb
冗長度	36	21	19	15
500塩基以上のコンティグの数	51	102	107	40
500塩基以上のコンティグでのゲノムのカバー率	98.0%	98.5%	98.5%	98.8%
コンティグ中のQV40以上の塩基の割合	99.97%	99.86%	99.81%	99.76%
500塩基以上のコンティグの平均長	65.0 kb	27.8 kb	26.5 kb	47.8 kb
N50(500塩基以上のコンティグ)*	170 kb	51.2 kb	51.6 kb	170 kb
最も大きいコンティグのサイズ	280 kb	211 kb	176 kb	230 kb

\*500塩基以上のコンティグについて大きいものから塩基数を足し合わせ、その長さが500塩基以上のコンティグの総塩基長の50%となったコンティグのサイズ。このサイズが大きいほどデータに偏りがなく、長いコンティグが得られていることを示す。

## ■ ATリッチなゲノムの解析例

高速シーケンス解析の特長のひとつとして、大腸菌によるライブラリー作製を行わない点が挙げられます。結果としてクローニングなどのバイアスが無いため、データの偏りが少ないと考えられます。特に大腸菌内での安定性が低いといわれているATリッチな配列についてはその傾向が顕著です。弊社で行ったATリッチな微生物ゲノム(GC含量:37%)の解析例を図2に示します。従来法ではゲノム配列が100個以上のコンティグに分断されていましたが、高速シーケンス解析結果ではコンティグ数は1/3以下に減少しています。また、両者のデータを統合することによってさらにコンティグ数は減少し、従来法では得られなかった超高精度なドラフト配列を得ることができました。

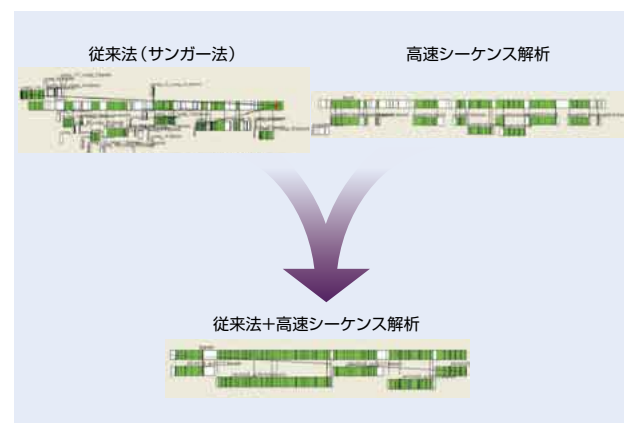


図2 ATリッチなゲノムでの解析結果概要図(GC含量:37%)

## ■ Fosmid, BACのマルチクローンシーケンス

高速シーケンス解析は全ゲノム解析だけではなく、クローンの配列解析にも適用可能です。複数のクローンを混合して解析することで、一度に多数のクローンのドラフト解析が可能となります。Fosmidクローンであれば24クローン程度を混合した解析が可能です。1クローンあたりのコストは20万円以下となり、非常にリーズナブルな価格となります。

## ■ 従来法との比較

従来法(サンガー法)との比較を表2に示します。全ゲノム配列の約98%以上の配列を決定するドラフト解析においてはコストが約1/2、時間が約1/3に軽減でき、非常に大きなメリットが得られます。

一方、精密解析(全長配列決定)においては依然としてサンガー法でのデータの投入が必須となりますが、従来法だけで解析する場合に比べると作業期間の短縮、コスト削減、バイアスの低減などを実現でき、効率的な解析が行えます。

なお、高速シーケンス解析ではピロシーケンス法を用いているため、ポリベース部位での配列の信頼性が低く、得られたゲノム配列を使用する際には注意が必要となります。

タカラバイオではご希望に応じてサンガー法での追加解析も可能であり、皆様の目的に即した解析をご提案いたします。

表2 従来法との比較

	高速シーケンス解析	従来の解析
解析手法	ピロシーケンス法	サンガー法
塩基長	約100 塩基	600 ~ 800 塩基
データ量	20万リード / Run (20Mb/Run)	96/384 リード/Run (50 ~ 200 kb/Run)
価格	サンガー法の約1/2	—
納期	1ヵ月	3ヵ月
鋳型調製工程	emulsion PCR	大腸菌へのクローニング
メリット	<ul style="list-style-type: none"> <li>より安価に大量、高速な解析が可能</li> <li>クローニングによるバイアスがない</li> <li>→特にATリッチな配列に有効</li> <li>・組換え実験が不要</li> </ul>	<ul style="list-style-type: none"> <li>・ベアリードシーケンスが可能 (bridgeクローンの取得が可能)</li> </ul>
検討課題	<ul style="list-style-type: none"> <li>・ベアリードシーケンスへの対応が不完全であり、GS20のみのデータからの全長配列決定は非常に困難</li> <li>・ポリベース(AAAAA等)の信頼性が低い</li> </ul>	—

## ■ タカラバイオのゲノム解析サービス

タカラバイオではゲノム解析をトータルでサポートしております。ドラフト解析後の追加配列解析、配列データ取得後のアノテーションなどバイオインフォマティクス解析、さらにはポストゲノム解析まで、多彩なメニュー、製品を取り揃えております。また、メニューにない解析についてもご希望に応じてサポートいたします。まずはお気軽にお問い合わせください。

### 【参考文献】

- 1) Goldberg, SM., *et al.* :A Sanger/pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes. (2006) *Proc. Natl. Acad. Sci. USA.*, **103** (30), 11240-11245
- 2) Dirk Hofreuter, *et al.* :Unique Features of a Highly Pathogenic *Campylobacter jejuni* Strain. (2006) *Infection and Immunity*, **74** (8), 4694-4707.
- 3) Jung D. Oh, *et al.* :The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: Evolution during disease progression. (2006) *Proc. Natl. Acad. Sci. USA.*, **103** (26), 9999-10004.
- 4) Andries K, *et al.* :A Diarylquinoline Drug Active on the ATP Synthase of *Mycobacterium tuberculosis*. (2005) *Science*, **307**, 223-227.
- 5) Gregory J. Velicer, *et al.* :Comprehensive mutation identification in an evolved bacterial cooperators and its cheating ancestor. (2006) *Proc. Natl. Acad. Sci. USA.*, **103** (21), 8107-8112

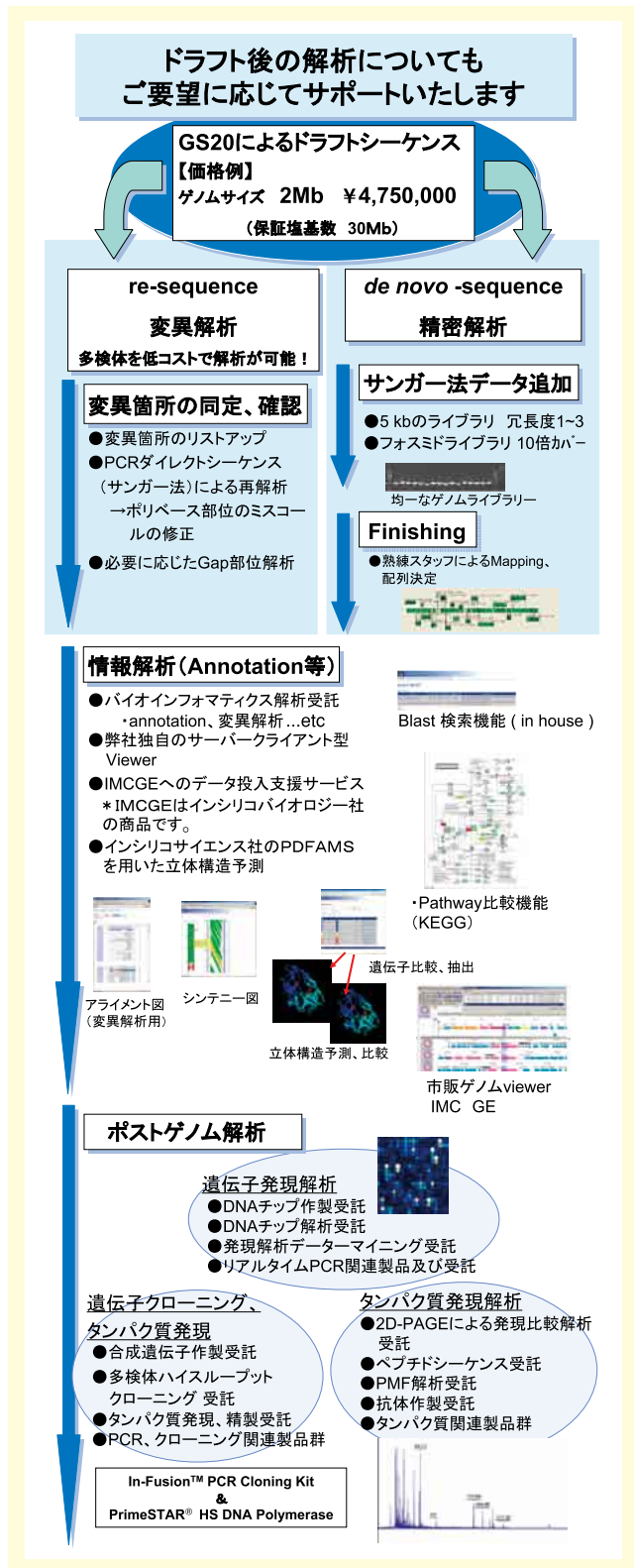


図3 タカラバイオでのゲノム解析サービス

### 【お問い合わせ先】

タカラバイオ(株)

ドラゴンジェノミクスセンター

TEL : 077-543-7331 FAX : 077-543-7225