illumina®

MiSeq® システムで高速にマルチプレックス 16S アンプリコンシーケンス

わずか 2 日間のシンプルなワークフローで複雑な細菌叢を属レベルまで分類

はじめに

細菌叢は、深総土壌から海底の熱水噴出孔といったさまざまな環境下で生存しています。またヒトや動物には、体内、体外を問わず多様な種類の微生物が生息しています。メタゲノミクスとは、このような細菌叢から分離した遺伝物質、すなわちメタゲノムに関する研究です。メタゲノミクス調査は通常、原核生物の 16S リボソーム RNA（rRNA）遺伝子を対象に行われます。この遺伝子は全長が約 1,500 塩基で、9 つの散在した可変領域と保存領域から構成されています。この領域をシーケンスすることで、微生物を分類することができます。次世代シーケンサーを使い、16S rRNA 遺伝子の特定領域をターゲットとしてシーケンスすることで、多様な細菌叢を分類することが可能になります（図 1）。

この Application Note では、デスクトップ型の MiSeq システムを用いた 16S アンプリコンシーケンスの手法と、そこから得られる結果をご紹介します。研究者は、イルミナで確立されたアンプリコンシーケンス用のプロトコールに基づいて、ご要望の分類レベルに応じた、1 つ、もしくは複数のゲノム領域をターゲットとするプライマーを設計します。この手法はウイルス研究や変異の検出、また他の微生物関連の研究といった、どのような標的アンプリコンシーケンスでもお使いいただけます。

図 1：16S アンプリコンシーケンスのワークフロー

16S アンプリコンシーケンスは、MiSeq システムでお使いいただける確立されたプロトコールです。サンプル調製、シーケンスに加え、装置付属のソフトウェアによる自動データ解析からなっており、分析結果の検討まで行えます。

手法

ターゲットの増幅方法

イルミナがご紹介する 16S rRNA シーケンスのプロトコール [1] は 16S 遺伝子の V3 および V4 領域をターゲットとし、459 塩基対のアンプリコンが得られます（生物により異なる）（図 2）。プライマーはターゲット領域周辺の保存性の高い領域をもとに設計されています。MiSeq は 250 塩基のペアエンド解析ができるので、それぞれのリード末端をオーバーラップさせることで、非常に高品質の V3 および V4 領域の全長リードが、たった 40 時間のランで得られます。そして低サイクル PCR で、イルミナシーケンスアダプターとデュアルインデックスバーコードがアンプリコンに付加されます。このライブラリーを定量標準化およびプールしたのち、シーケンスを開始します。この手法は 16S 遺伝子 [2] のさまざまな領域、その他の興味ある遺伝子にもご利用いただけます。MiSeq システムでは、マルチプレックスで、96 サンプルを 1 ライブラリーにまとめ、シーケンスすることが出来ます [3]。

シーケンス

インデックスをつけたアンプリコンを MiSeq 試薬カートリッジに投入し、フローセルと共に MiSeq システムにセットします。その後、自動化されたクラスター形成とデュアルインデックスリードを用いたペアエンドシーケンスを行います。2 x 250bp を解析するには合計 40 時間かかります。

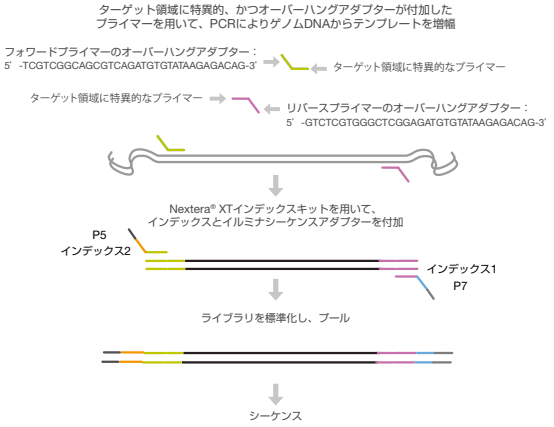
データ解析

1 次解析（画像解析およびベースコール）は MiSeq システム上で行います。ランの終了後は、イルミナのフリーのソフトウェア MiSeq Reporter（MSR）を利用して 2 次解析を実施することができます。MiSeq システム上またはイルミナの提供するクラウド型ソリューションの BaseSpace、その他ネットワークに接続していないコンピュータ上でもご利用頂けます。この MSR によりメタゲノム群集から属レベルのデータが得られ、それぞれの属が存在する割合まで把握できます。

結果

16S rRNA 遺伝子配列と高頻度可変領域は、数多くの生物種で既に決定されています。これらの配列は Greengenes [4] や Ribosomal Database Project [5] など複数のデータベースで入手可能です。分類においては、16S rRNA 遺伝子の全長の代わりに、高頻度可変領域のみをシーケンスすることで十分な情報が得られます [6,7]。MSR のメタゲノム 16S rRNA ワークフローでは、分類のために Greengenes データベースを使用しています。

図 2：シンプルなアンプリコンのワークフロー



ターゲット領域の上流および下流に相補的でオーバーハングアダプターを持つフォワードプライマーとリバースプライマーを設計し、ゲノム DNA からテンプレートを増幅します。次のステップの低サイクル PCR では、さらにマルチプレックスのためのインデックスとイルミナシーケンスアダプターが付加されます。ライブラリーは標準化後にプールし、MiSeq システムでシーケンスします。

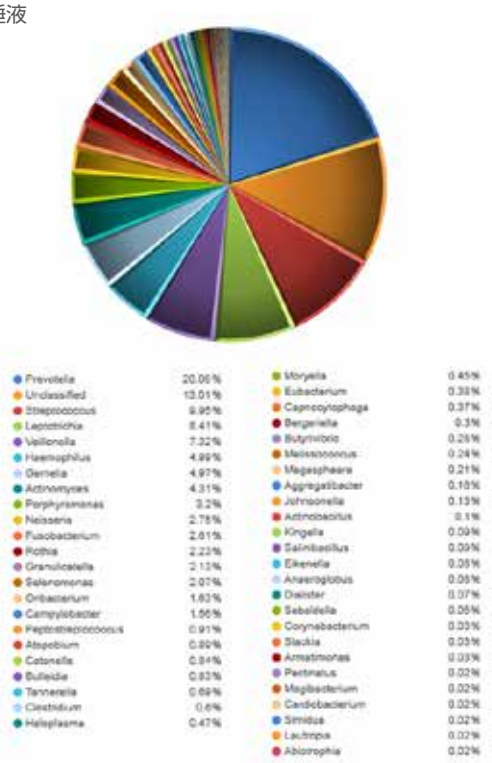
フォワードプライマー：
5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3'

リバースプライマー：
5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3'

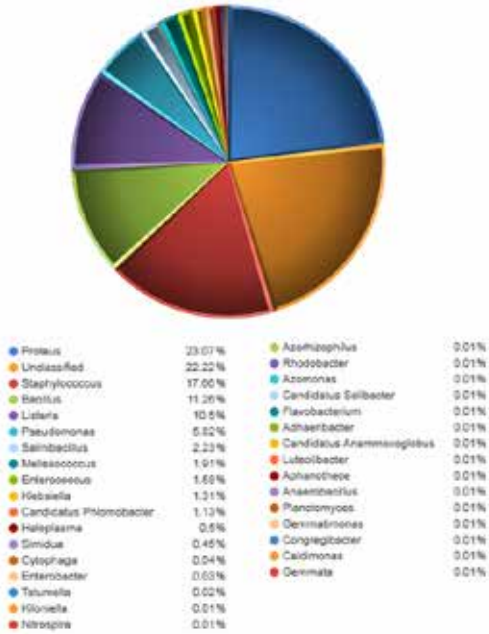
MSR は 16S のワークフローをサポートしており、属レベルでの分類を可能にしました。16S ワークフローは Bayesian 分類法 [8] に基づいて分類を行い、それぞれのリードに信頼度スコアを付けます。分類の際、リードが一定の信頼度スコアを満たさない場合は、そのリードは分類されずに「分類不可」のリードとなります。図 3 に示されるように、MSR からの図表アウトプットにはサンプル中のそれぞれの属の割合が表示されます。図 3A はヒトの唾液サンプルに含まれる生物種を属レベルで表示しています。また、図 3B はメタゲノムにおけるコントロールサンプルの分類結果で、以下の微生物由来の DNA が同等の割合で含まれています（*Bacillus cereus*、*Klebsiella pneumonia*、*Listeriainnocua*、*Listeriamonocytogenes*、*Proteus mirabilis*、*Proteus vulgaris*、*Pseudomonaaeruginosa*、*Pseudomonas stutzeri*、*Staphylococcusaureus*、*Staphylococcus epidermidis*）。

図 3：微生物サンプルの解析

A. 唾液



B. コントロール



MiSeq Reporter（MSR）による 16S rRNA メタゲノミクスワークフローのアウトプットです。テスト（A）およびコントロール（B）の微生物サンプルに含まれているすべての属が表示されています。DNA サンプルはサンノゼ州立大学の Cleber Ouverney 博士に提供いただきました。

まとめ

イルミナの 16S rRNA プロトコールは、MiSeq システムで 16S rRNA 遺伝子の V3、V4 領域をシーケンスし、複雑な細菌叢を属レベルで分類するのに使用することができます。簡単なマルチプレックス法と、データのアウトプットの増加により、サンプルの調製からわずか 2 日で分類結果が得られます。またイルミナの MiSeq Reporter を使うことで、複雑な群集解析とすぐに論文投稿ができるレベルのデータを得ることが出来ます。MiSeq システムは多くのサンプル数、そして複数の 16S 領域を処理することができるため、大規模なメタゲノム解析をより精密に行うことが可能です。今後 MSR ソフトウェアのさらなる改善により、向上したデータのレポートや視覚化、さらには種のレベルでの分類が可能になることが期待されています。またプロトコールの中でも紹介されているこのワークフローは 16S アンプリコンのシーケンスのみならず、他のシーケンスプロジェクトにおいても応用が可能です。

また他社の解析ツール、例えば Quantitative Insights into Microbial Ecology（QIIME）もメタゲノムサンプルの解析にご利用になれます [9,10]。

MiSeq システムについての詳細は、
jp.illumina.com/miseq をご覧ください。

参考文献

1. Illumina-demonstrated protocol for 16S Amplicon Sequencing
2. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, et al.（2013）Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41（1）.
3. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, et al.（2011）Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci USA 108:4516–4522.
4. Desantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, et al.（2006）Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72（7）:5069–5072.
5. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, et al.（2009）The Ribosomal Database Project: improved alignments and newtools for rRNA analysis. Nucl Acids Res 37:D141–D145.
6. Liu Z, Lozupone C, Hamady M, Bushman FD, KnightR.（2007）Short pyrosequencing reads suffice for accurate microbial communityanalysis. Nucl Acids Res 35:18.
7. Liu Z, DeSantis TZ, Andersen GL, Knight R.（2008）Accuratetaxonomy assignments from 16S rRNA sequences produced byhighly parallel pyrosequencers. Nucl Acids Res 36:18.
8. Wang Q, Garrity GM, Tiedje JM, Cole JR.（2007） Naïve Bayesianclassifier for rapid assignment of rRNA Sequences into the new bacterial taxonomy. Appl Environ Microbiol 73（16）:5261-5267.
9. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman,FD, et al.（2010）QIIME allows analysis of high-throughput communitysequencing data. Nat Methods 7（5）:335–336.
10. http://qiime.org/tutorials/processing_illumina_data.html

追加情報

MiSeq についての詳細は、
jp.illumina.com/miseq をご覧ください。

この ApplicationNote は情報提供を目的としています。場合によっては、未承認の他社の試薬の購入が必要となる可能性がありますが、このような他社の試薬をイルミナ製品でご使用になる場合のテクニカルサポートについては保証しかねます。

イルミナ株式会社

〒108-0014 東京都港区芝 5-36-7 三田ベルジュビル 22 階
Tel (03) 4578-2800 Fax (03) 4578-2810
jp.illumina.com
www.facebook.com/illuminakk

代理店

本製品の使用目的は研究に限定されます。 販売条件：jp.illumina.com/tc

© 2016 Illumina, Inc. All rights reserved.

Illumina, BaseSpace, BeadArray, BeadXpress, cBot, CSPro, DASL, Design Studio, GAIIx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, NextSeq, NuPCR, SeqMonitor, Solexa, TruSeq, TruSight, VeraCode, the pumpkin orange color, the Genetic Energy streaming bases design は、Illumina, Inc. の商標または登録商標です。
その他の会社名や商品名は、各社の商標または登録商標です。予告なしに仕様および希望販売価格を変更する場合があります。

Pub. No. 1270-2013-J001 20DEC2016

illumina®