

カスタムLibraryを デザインするには

酒井名朋子

Sr Technical Applications Scientist

イルミナ株式会社

Outline

- ▶ 本セッションの注意点ときっかけ
- ▶ カスタムLibraryの作成に当たりご注意ください点
- ▶ Illumina TruSeq DNA HT libraryを理解する
- ▶ Case Study: 論文通り設計してもサンプルリードが取れない場合
- ▶ まとめ

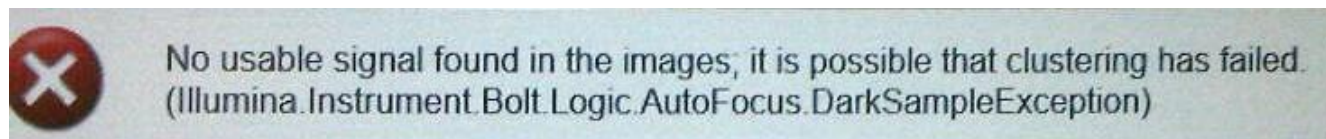
本セッションのご注意事項

- ▶ 本セッションは必ずカスタムLibraryのRun成功を保証するものではありません
 - あくまでご留意いただく点についてお伝えするセッションになります。
 - カスタムLibrary（弊社サンプル調製キットをお使いいただかないLibrary)についてはサポート外とさせていただきます。
 - 考慮すべきParameterが多く、本社でValidationを行っていないため
 - カスタムLibraryを使用したRunの装置トラブルシュートはお手伝いいたします。
- ▶ 本セッションの内容は、以下の論文をもとに作成しております
 - *Bentley et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53-59*
 - Supplementary Informationを中心にご参照ください。
 - Illumina Adapter Sequence Letter
 - Illumina HP SupportページよりDownloadしてください
http://support.illumina.com/documents/MyIllumina/6378de81-c0cc-47d0-9281-724878bb1c30/2012-09-18_IlluminaCustomerSequenceLetter.pdf
- ▶ 2週間後、4月19日の以下のセミナーも合わせてご参照ください
 - 「塩基に偏りのある配列をMiSeqでシーケンスするには」

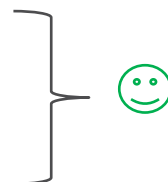
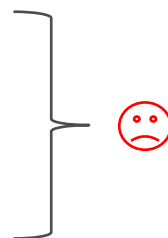
本セッションの開催のきっかけ

-こんなことでお悩みでないでしょうか？

- ▶ カスタムLibraryを作成したいが、どの配列があれば十分かわからない
- ▶ カスタムPrimerを作成しMiSeqでPCR Ampliconを流したが、データが出ない
 - サイクル1でシグナルが出ないエラー

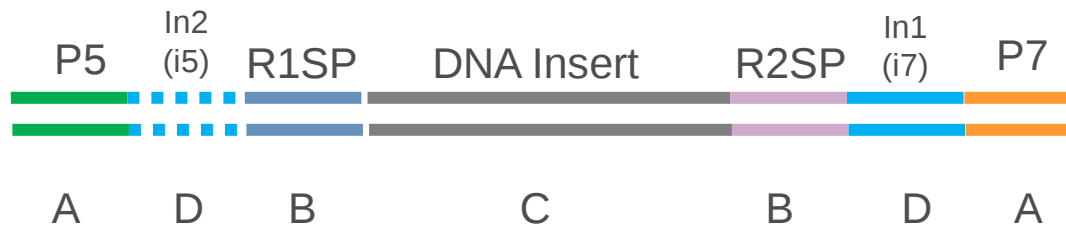


- FCにクラスターが存在する？
 - オリゴ配列は問題ない？
 - Denature時のNaOHのpHは>pH13？
 - NaOHのCarry Overは終濃度<1mM？
- クラスターにはSeq PrimerがHybしている？
- MiSeq装置本体の状態は問題ない？
 - 試薬Deliveryに問題がある？
 - 温度制御に問題がある？
 - Focus取得部に故障が起きている？



イルミナでトラブルシュートのお手伝いします

ライブラリデザインにあたり、どの配列が必要か？



A: Flowcellとのハイブリ

B: シーケンスプライマーのハイブリ

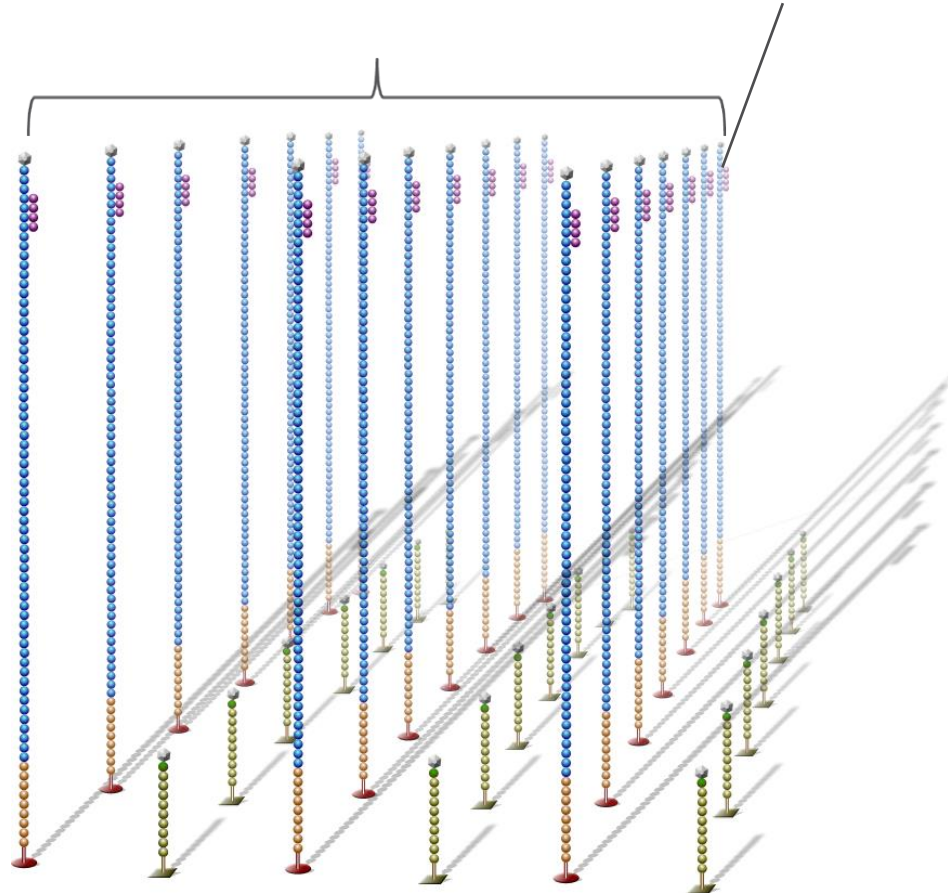
C: シーケンスされるDNA

D: Index (サンプル区別のためのバーコード)

ライブラリデザインにあたり、どの配列が必要か？

クラスターが形成されているか

Seq PrimerがHybしているか



Illumina FC上のCluster模式図

FC上のオリゴ配列

▶ SR用 FC: HiSeq, GAのSR FC上の配列

SR P5/oligo 'A': 5'-PSTTTTTTTTTT-(diol)3-AATGATACGGCGACCACCGA

SR P7/oligo 'B': 5'-PSTTTTTTTTTTCAAGCAGAAGACGGCCATACGA

▶ PE、MiSeq用FC: HiSeq, GA のPE FC上の配列、MiSeq

P5/oligo 'C': 5'-PS-TTTTTTTTTTAATGATACGGCGACCACCGAGAUCTACAC



P7/oligo 'D': 5'-PS-TTTTTTTTTTCAAGCAGAAGACGGCCATACGA^{Goxo}AT



カスタムLibraryを作成する際の注意点

- ▶ Cluster が形成されるか？
 - P5/P7 配列を持つか？
 - 最終LibraryのP7,P5 のFw Rv配列の組み合わせが正しいか？
 - P7- NNN- rP5
 - rP7-NNN-P5
- ▶ Seq PrimerがHybするか？
 - Seq Primerの特異性に問題はないか？
 - 1分子に2つSeq PrimerがHybするような可能性はないか？
 - Hyb効率は十分か？（縮重Primerに注意）
- ▶ Custom Oligoの精製は十分か？

TruSeq Libraryを理解しよう

作製されるライブラリ構造

D501-508 Adapters

5' AATGATACGGCGACCACCGAGATCTACAC **TATAGCCT** ACACTCTTTCCCTACACGAC **GCTCTTCCGATC** T
P5配列 i5

D701-712 Adapters

5' **GATCGGAAGAGC** ACACGTCTGAACTCCAGTCAC **AATACTCG** ATCTCGTATGCCGTCTTCTGCTTG
Y字形成用配列 i7 rP7配列

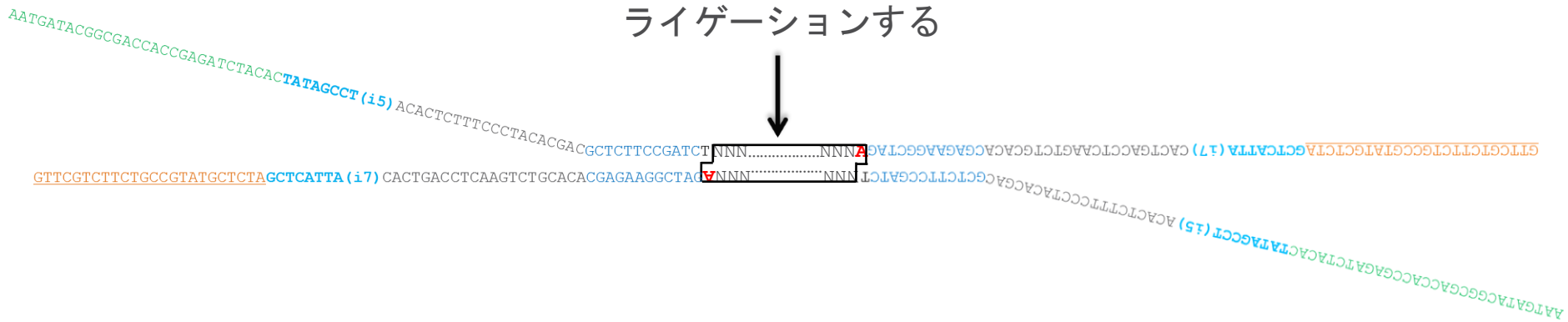


Y字アダプター構造を作る

5' AATGATACGGCGACCACCGAGATCTACAC **TATAGCCT (i5)** ACACTCTTTCCCTACACGACGCTCTTCCGATCT
GTTCTGTTCTGCCGTATGCTCTA **GCTCATT (i7)** CACTGACCTCAAGTCTGCACACGAGAAGGCTAG 5'

TruSeq DNA Sample Prep kit 作製されるライブラリ (2本鎖DNA) 構造

インサートDNA
末端平滑化の後に
3'末端A突出で作製し
Y字アダプターと
ライゲーションする



TruSeq DNA HT library

1, Adapter Ligation

5'-AATGATACGGCGACCACCGAGATCTACAC-TATAGCCT(i5)-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-[N]A- GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-ATTACTCG(i7)-ATCTCGTATGCCGCTCTTCTGCTTG-3' (1)
3'-GTTCTGTCTTCTGCGGTATGCTCTA-GCTCATTA(i7)-CACTGACCTCAAGTCTGCACACGAGAAGGCTAG-A[N]-TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-TCCGATAT(i5)-CACATCTAGAGCCACCAGCGGCATAGTAA-5' (2)

2-0, Possible PCR Primer (PCR Primer Cocktailは1種類、AdapterにIndexがついている。)

5' AATGATACGGCGACCACCGAGATCTACAC
5' CAAGCAGAAGACGGCATACGAGAT

2-1, PCR 1st cycle

5'-AATGATACGGCGACCACCGAGATCTACAC-TATAGCCT(i5)-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-[N]A- GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-ATTACTCG(i7)-ATCTCGTATGCCGCTCTTCTGCTTG-3' (1)
3'-GTTCTGTCTTCTGCGGTATGCTCTA-GCTCATTA(i7)-CACTGACCTCAAGTCTGCACACGAGAAGGCTAG-A[N]-TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-TCCGATAT(i5)-CACATCTAGAGCCACCAGCGGCATAGTAA-5' (2)
5' CAAGCAGAAGACGGCATACGAGAT =>

3, Amplicon from 1st cycle 1a=2a, 1b=2b

5'-AATGATACGGCGACCACCGAGATCTACAC-TATAGCCT(i5)-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-[N]A- GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-ATTACTCG(i7)-ATCTCGTATGCCGCTCTTCTGCTTG-3' (1a)
3'-TTAGTATGCCGCTGGTGGCTCTAGATGTG-ATATCGGA(ri5)-TGTGAGAAAGGGATGTGCTGCGAGAAGGCTAGT-[N]T-CTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG-TAATGAGC(ri7)-TAGAGCATACGGCAGAAGACGAAC-5' (1b)
3'-GTTCTGTCTTCTGCGGTATGCTCTA-GCTCATTA(i7)-CACTGACCTCAAGTCTGCACACGAGAAGGCTAG-A[N]-TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-TCCGATAT(i5)-CACATCTAGAGCCACCAGCGGCATAGTAA-5' (2a)
5' CAAGCAGAAGACGGCATACGAGAT-CGAGTAAT(ri7)-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T[N]-AGATCGGAAGAGCGTCGTAGGGAAAGAGTGT-AGGCTATA(ri5)-GTGTAGATCTCGGTGGTGGCGGTATCATT-3' (2b)

4, PCR 2nd cycle (usual)

5'-AATGATACGGCGACCACCGAGATCTACAC-TATAGCCT(i5)-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-[N]A- GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-ATTACTCG(i7)-ATCTCGTATGCCGCTCTTCTGCTTG-3' (1a)
3'-TTAGTATGCCGCTGGTGGCTCTAGATGTG-ATATCGGA(ri5)-TGTGAGAAAGGGATGTGCTGCGAGAAGGCTAGT-[N]T-CTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG-TAATGAGC(ri7)-TAGAGCATACGGCAGAAGACGAAC-5' (1b)
5'-AATGATACGGCGACCACCGAGATCTACAC =>

Cluster Generation

5-0, FC Grafted Oligo Sequence (PE FC)

P5|5'-Tx10AATGATACGGCGACCACCGAGAUCTACAC
P7|5'-Tx10CAAGCAGAAGACGGCATACGAGAT

5-1, Hyb to P7 (Read 1), 1a and 2a

P7|5'-CAAGCAGAAGACGGCATACGAGAT=>
3'-GTCGCTCTTCGCGTATGCTCTA-GCTCATTA(i7)-CACTGACCTCAAGTCTGCACACGAGAAGGCTAG-A[N]-TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-TCCGATAT(i5)-CACATCTAGAGCCACCAGCGGCATAGTAA-5' (a)

P5|5'- AATGATACGGCGACCACCGAGAUCTACAC=>
3'-TTAGTATGCCGCTGGTGGCTCTAGATGTG-ATATCGGA(ri5)-TGTGAGAAAGGATGTGCTGCGAGAAGGCTAGT-[N]T-CTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG-TAATGAGC(ri7)-TAGAGCATACGGCAGAAGACGAAC-5' (b)

6, Bridge Amplification

P7|5'-CAAGCAGAAGACGGCATACGAGAT- CGAGTAAT(ri7)-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T[N]-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-AGGCTATA(ri5)-GTGTAGATCTCGGTGGTCGCCGTATCATT-3'
3'-GTCGCTCTTCGCGTATGCTCTA-GCTCATTA(i7)-CACTGACCTCAAGTCTGCACACGAGAAGGCTAG-A[N]-TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-TCCGATAT(i5)-CACATCTAGAGCCACCAGCGGCATAGTAA-5'

P5|5'-AATGATACGGCGACCACCGAGAUCTACAC -TATAGCCT(i5)-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-[N]A-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-ATTACTCG(i7)-ATCTCGTATGCCGCTTCTGCTTG-3'
3'-TTAGTATGCCGCTGGTGGCTCTAGATGTG-ATATCGGA(ri5)-TGTGAGAAAGGATGTGCTGCGAGAAGGCTAGT-[N]T-CTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG-TAATGAGC(ri7)-TAGAGCATACGGCAGAAGACGAAC-5'

7, Primer Hyb

P7|5'-CAAGCAGAAGACGGCATACGAGAT-CGAGTAAT(ri7)-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T[N]-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-AGGCTATA(ri5)-GTGTAGATCTCGGTGGTCGCCGTATCATT-3'
=< TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA (possible R1)

8, Index, R2 Sequence primer hyb

P7|5'-CAAGCAGAAGACGGCATACGAGAT-CGAGTAAT(ri7)-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T[N]-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-AGGCTATA(ri5)-GTGTAGATCTCGGTGGTCGCCGTATCATT-3'
(i7)<=CACTGACCTCAAGTCTGCACACGAGAAGGCTAG (possible In1)

P7|5'-CAAGCAGAAGACGGCATACGAGAT-CGAGTAAT(ri7)-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T[N]-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-AGGCTATA(ri5)-GTGTAGATCTCGGTGGTCGCCGTATCATT-3'
(i5)<= CACATCTAGAGCCACCAGCGGCATAGTAA (possible In2)

P5|5'-AATGATACGGCGACCACCGAGAUCTACAC -TATAGCCT(i5)-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-[N]A-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-ATTACTCG(i7)-ATCTCGTATGCCGCTTCTGCTTG-3'
=< T-CTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG (possible R2)

R1 の配列を読む際のLibrary構造

P5 | 5' -T_x10 AATGATACGGCGACCACCGAGAUCTACAC

P7 | 5' -T_x10 CAAGCAGAAGACGGCATACGAGAT

P7 配列

FC | 5' - CAAGCAGAAGACGGCATACGAGAT = CGAGTAAT (ri7) -GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T [NNN]-AGATCGGAAGAGCGTCGTGTAGGAAAGAGTGT-AGGCTATA (ri5) -GTGTAGATCTCGGTGGTCGCCGTACATT
<=TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA [possible R1 Sequence Primer]



<=TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA [possible R1 Sequence Primer]

TruSeq® DNA HT and RNA HT Sample Prep Kits^{2,4,5}

[possible R1 Sequence Primer]

D501–D508 adapters

AATGATACGGCGACCACCGAGATCTACAC [i5] ACACTCTTTCCCTACACGACGCTCTTCCGATCT



D701–D712 adapters

GATCGGAAGAGCACACGTCTGAACTCCAGTCAC [i7] ATCTCGTATGCCGTCTTCTGCTTG

Index1 配列を読む際のLibrary構造

P7 配列

FC | 5'-CAAGCAGAAGACGGCATACGAGAT= CGAGTAAT (ri7)-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T [NNN]-AGATCGGAAGAGCGTCGTAGGAAAGAGTGT-AGGCTATA (ri5)-GTGTAGATCTCGGTGGTCGCCGTACATT
(i7)<=CACTGACCTCAAGTCTGCACACGAGAAGGCTAG [possible In1 Sequence Primer]



(i7)<=CACTGACCTCAAGTCTGCACACGAGAAGGCTAG [possible In1 Sequence Primer]

TruSeq® DNA HT and RNA HT Sample Prep Kits ^{2,4,5}

D501–D508 adapters

AATGATACGGCGACCACCGAGATCTACAC [i5] ACACTCTTTCCCTACACGACGCTCTTCCGATCT

D701–D712 adapters

GATCGGAAGAGCAGACGTCTGAACTCCAGTCAC [i7] ATCTCGTATGCCGTCTTCTGCTTG

[possible In1 Sequence Primer]

i7 index
D701

i7 bases in D7xx adapter and
for entry on sample sheet
ATTACTCG

サンプルシートにはこのままの配列を記載する

FC | 5'-CAAGCAGAAGACGGCATACGAGAT= CGAGTAAT (ri7)-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T [NNN]

GCTCATTA (i7)<=CACTGACCTCAAGTCTGCACACGAGAAGGCTAG [In1 Sequence Primer]

Dual Index, Index2 のSRとPEでの違い

Figure 2 Dual-Indexed Single-Read Sequencing

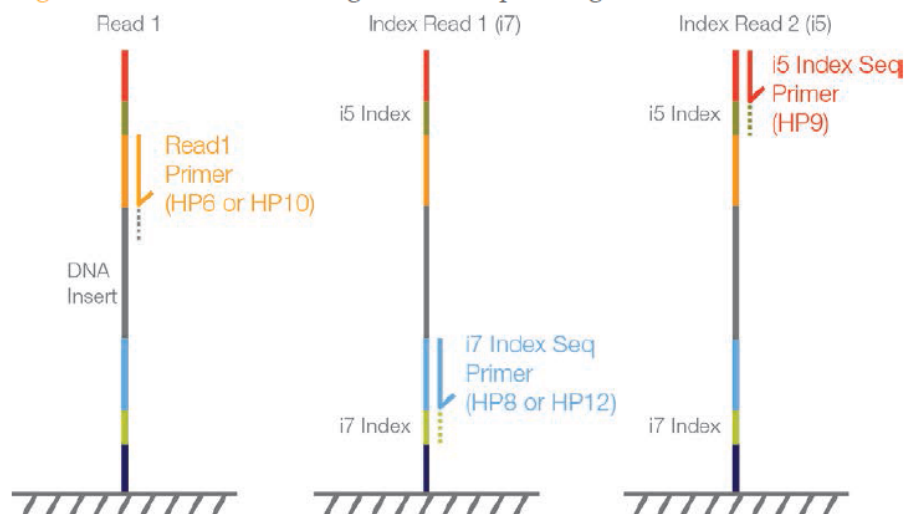
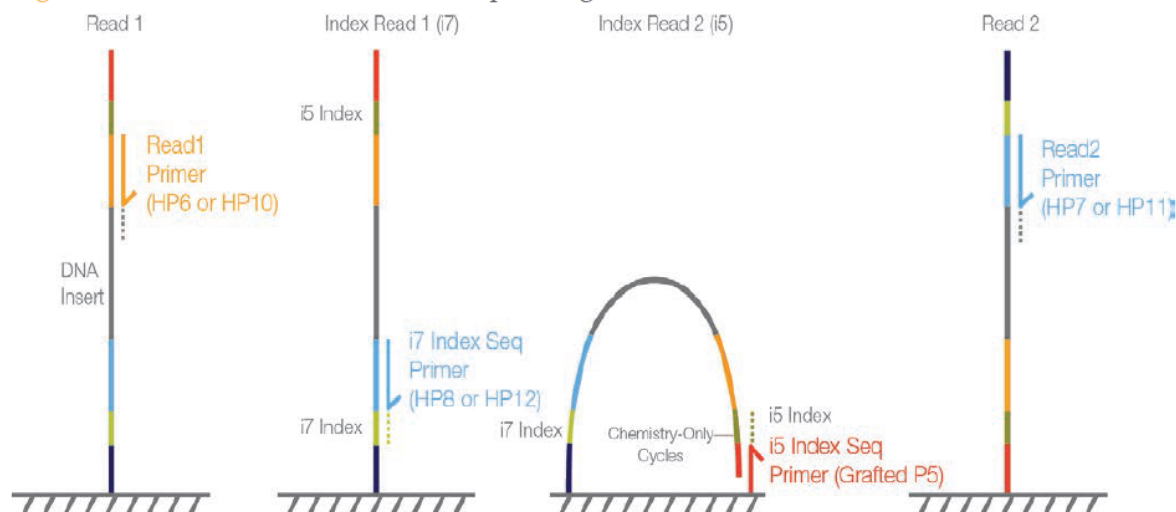


Figure 3 Dual-Indexed Paired-End Sequencing

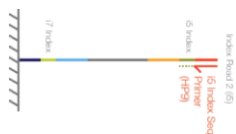


Index2 配列を読む際のLibrary構造

SR Dual Indexの場合

P7 配列

FC | 5'-CAAGCAGAAGACGGCATACGAGAT= CGAGTAAT (ri7) -GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T [NNN] -AGATCGGAAGAGCGTCGTGTAGGAAAGAGTGT-AGGCTATA (ri5) -GTGTAGATCTCGGTGGTCGCCGTACATT



(i5) <= CACATCTAGAGCCACCAGCGGCATAGTAA
[possible In2 Sequence Primer]

(i5) <= CACATCTAGAGCCACCAGCGGCATAGTAA
[possible In2 Sequence Primer]

PE Dual Indexの場合

P7 配列

FC | 5'-CAAGCAGAAGACGGCATACGAGAT= CGAGTAAT (ri7) -GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-T [NN

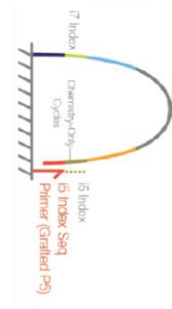
TTACATAGCCGCTGGTGGCTCTAGATGTG- (ri5) ATATCGGA- TGTGAGAAAGGGATGTGCTGCGAGAAGGCTAGA [NN

FC | 5'-AATGTATCGGCGACCACCGAGAUCTACAC=> (i5)



FC | 5'-AATGTATCGGCGACCACCGAGAUCTACAC=> (i5)

P5 配列



TruSeq® DNA HT and RNA HT Sample Prep Kits^{2,4,5}

D501–D508 adapters

AATGATACGGCGACCACCGAGATCTACAC [i5] ACACTCTTTCCCTACACGACGCTCTTCCGATCT

[possible In2 Sequence Primer]

D701–D712 adapters

GATCGGAAGAGCACACGTCTGAACTCCAGTCAC [i7] ATCTCGTATGCCGTCTTCTGCTTG

i5 index name

D501

i5 bases in D5xx adapter and
for entry on sample sheet

TATAGCCT

サンプルシートにはこのままの配列を記載する

R2 の配列を読む際のLibrary構造

P5 | 5' -T~~x~~10 AATGATACGGCGACCACCGAGAUCTACAC

P7 | 5' -T~~x~~10 CAAGCAGAAGACGGCATACGAGAT

P5 配列

P5 | 5' -AATGATACGGCGACCACCGAGAUCTACAC -TATAGCCT(i5) -ACACTCTTTCCCTACACGACGCTCTTCCGATCT- [N] A-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-ATTACTCG(i7) -ATCTCGTATGCCGTCTTCTGCTTG-3'
<= T-CTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG [possible R2 Sequence Primer]



<= T-CTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG [possible R2 Sequence Primer]

TruSeq® DNA HT and RNA HT Sample Prep Kits ^{2,4,5}

D501–D508 adapters

AATGATACGGCGACCACCGAGATCTACAC [i5] ACACTCTTTCCCTACACGACGCTCTTCCGATCT

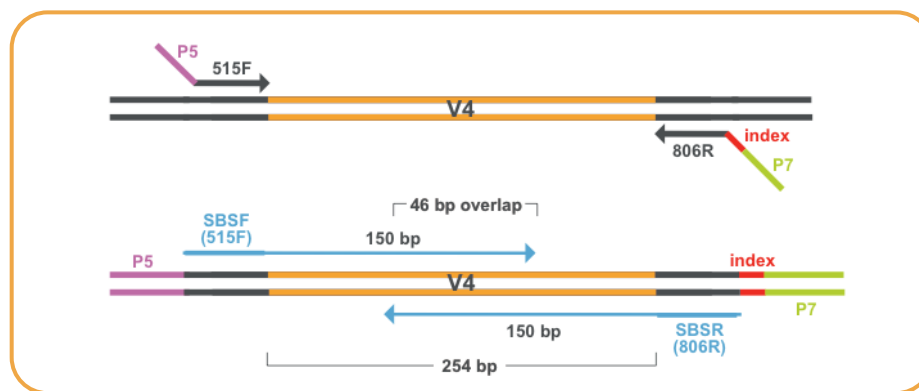
D701–D712 adapters

GATCGGAAGAGCACACGTCTGAACTCCAGTCAC [i7] ATCTCGTATGCCGTCTTCTGCTTG

[rev possible R2 Sequence Primer]

Case Study

- ▶ 論文通りに設計しても、サンプルからリードが取れない
 - 16S メタゲノム解析



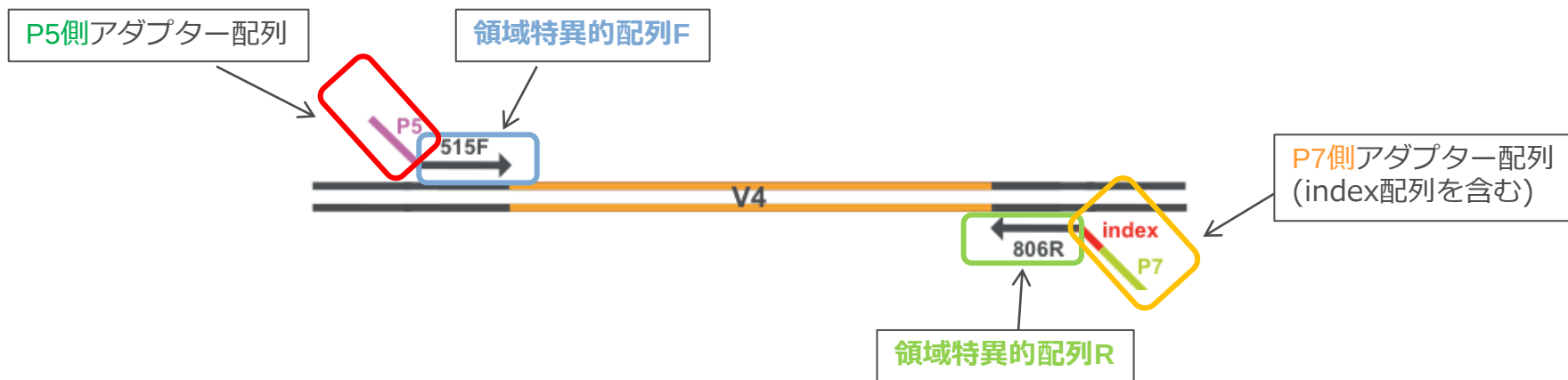
Caporaso JG, et al., Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 2012 Mar 8.

5' - AATGATACGGCGACCACCGAGATCTACAC TATGGTAATT GT GTGCCAGCMGCCGCGGTAA
P5配列 16S V4領域に特異的な配列

5' - CAAGCAGAAGACGGCATACGAGAT TCCCTTGTCTCC AGTCAGTCAG CC GGACTACHVGGGTWTCTAAT
P7配列 Index12塩基 16S V4領域に特異的な配列
(2168種類)

改善方法: Tailed PCR

- ▶ 領域特異的配列の外側に、Illumina Libraryを参照したAdapter 配列を付ける



- ▶ Forward Primer : -80 mer

- 58 mer (P5側アダプター) + 領域特異的配列F

5' - AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT 領域特異的配列F

- ▶ Reverse Primer : -90 mer

- 68 mer (P7側アダプター) + 領域特異的配列R

- サンプルごとにIndex異なるIndexを使用

5' - CAAGCAGAAGACGGCATACGAGAT CGTGAT GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT 領域特異的配列 R

Index
48種類

配列情報詳細： MiSeqで使用するアダプター&プライマー配列

下記塩基配列が DNAインサート と書いた両端についたDNAを調製すれば解析可能

P5配列*

R1 Seq Primerが→向きに結合**

5' AATGATACGGCGACCAACGAGATCTACAC TCTTTCCCTACACGACGCTCTTCCGATCT -DNAインサート(下に続く)
3' TTACTATGCCGCTGGTGGCTCTAGATGTG AGAAAGGGATGTGCTGCGAGAAGGCTAGA -DNAインサート(下に続く)

DNAインサート(続き) AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC ATCACGATCTCGTATGCCGTCTTCTGCTTG 3'
DNAインサート(続き) TCTAGCCTTCTCGTGTGCAGACTTGAGGTCAGTG TAGTGCTAGAGCATACGGCAGAAGACGAAC 5'

Index1 Seq Primerが →向きに結合**

R2 Seq Primerが←向きに結合**

Index***

P7配列*

*P5, P7配列はフローセル上への結合とフローセル上でのDNA鋳型増幅（ブリッジPCR）に必要

** 実際のプライマー結合部位は多少前後する可能性がある。

***このインデックス配列をサンプルごとに变えることで、
MiSeqシーケンス後にサンプルごとにデータを分離できる

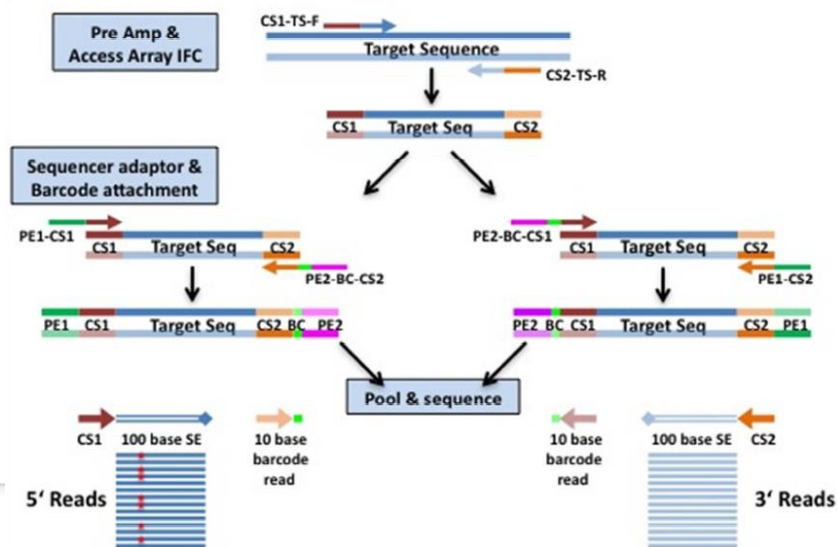
注意点:Tailed PCR

▶ 16S のアプリケーションの事例のご紹介

- 森永様のイルミナWebinarをご参照ください。
 - 「16S ribosomal RNA遺伝子から始める腸内細菌叢解析」
 - 森永乳業株式会社 食品基盤研究所 生物機能研究部 小田巻 俊孝 博士
 - http://www.illuminakk.co.jp/events/webinar_japan.ilmn
- High Fidelityな酵素の場合はLong Tail のPrimerを反応性が低く、PCRがかかりにくい場合がある。
 - TaKaRa Ex Taq HSを使用して解決
 - Earth Microbiome Projectでは5PRIME社製 5Prime Hot Master Mix を使用

▶ Tailed PCR がうまくいかない場合は2 Step PCRを使用

- 詳細は論文のCorresponding Author に直接お問合せください。



Forshe T, et al., Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci Transl Med*. 2012 May 30;4(136):136ra68.

まとめ

- ▶ カスタムLibraryデザインの際には以下の点に注意する
 - P5/P7配列の並び
 - Seq PrimerのHybサイト
 - 装置や試薬のバージョンによって、条件が異なる
- ▶ 一度で必ずうまくいくとは限らない。
 - 論文通り設計してもうまくいかない場合がある。
 - Tailed PCRを試す
 - うまくいかない場合はCorresponding Author に問い合わせ