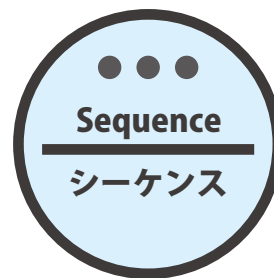
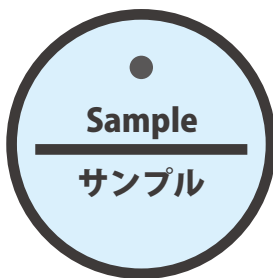


---

# 次世代シーケンス解析サービス I

---



\* データ解析に関しては、冊子「次世代シーケンス解析サービスⅡーデータ解析ー」をご覧ください。

# illumina HiSeq / MiSeq

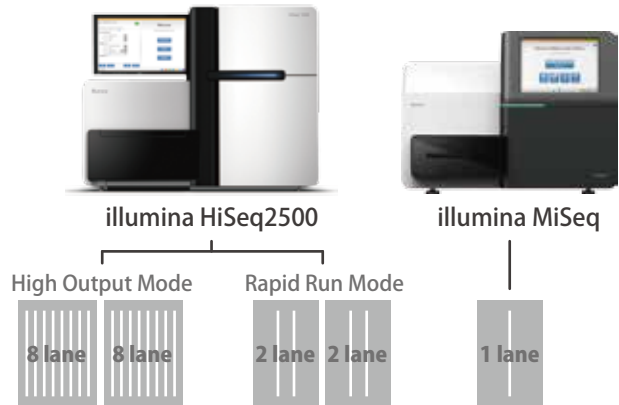
## illumina シーケンサー

### illumina HiSeq2500

100bp 程度のショートリードデータを、1 ランで最大 1Tb 出力します。マルチプレックス解析により、サンプルあたり任意のデータ量を取得する事が可能です。

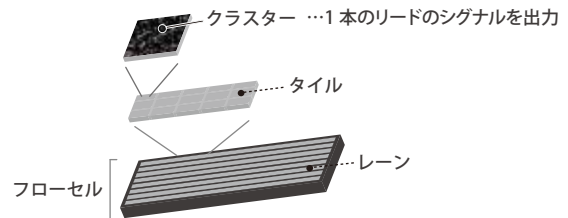
### illumina MiSeq

300bp ペアエンドシーケンスでは、～ 500 bp 程度のインサート配列全長決定が可能です。16S rRNA等アンプリコンを対象とした解析に適しています。



### illumina HiSeq4000・HiSeqX

高密度のクラスターを形成するフローセルで、より高いスループットの出力が可能です。多くのデータを必要とする、高等動物の全ゲノム解析や、多検体のエクソーム解析に使用されます。  
※ HiSeqX はホールゲノム解析に限定したシステムです。



## 1 レーンあたりのデータ量

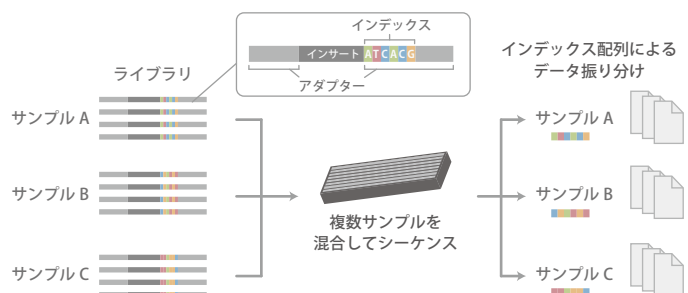
シーケンサー	HiSeq2500				MiSeq						
	High Output Mode		Rapid Run Mode		V2		V3				
試薬バージョン※1	V4		V2		V2		V3				
リード数※2,4	1.8 億		1 億		1000 万		1600 万				
リード長	50bp	100bp※5	125bp	50bp	100bp※5	150bp	250bp	150bp	250bp	75bp	300bp
データ量※3,4	18Gb	36Gb	45Gb	10Gb	20Gb	30Gb	50Gb	3Gb	5Gb	2.5Gb	10Gb

シーケンサー	HiSeq4000	HiSeqX
リード数	～ 3 億	～ 3.7 億
リード長	150bp	150bp
データ量	90Gb	100Gb

- ※1 上記は、弊社にて現在主に使用している試薬バージョンとなります。市販されている他の試薬バージョンの取扱いについては、お問い合わせください。
- ※2 「リード数」は、ペアエンドシーケンス解析のリードペア数となります。
- ※3 「データ量」は、ペアエンドシーケンス解析時の塩基数を示します。
- ※4 取得データ量（リード数）は弊社実測に基づく参考値であり、イルミナ社カタログスペックとは異なる場合があります。また、保証値ではございません。
- ※5 HiSeq 100bp ペアエンド解析については、サンプルあたり Gb 単位でのご依頼を承っております。他のリード長の解析プランについては、お問い合わせください。

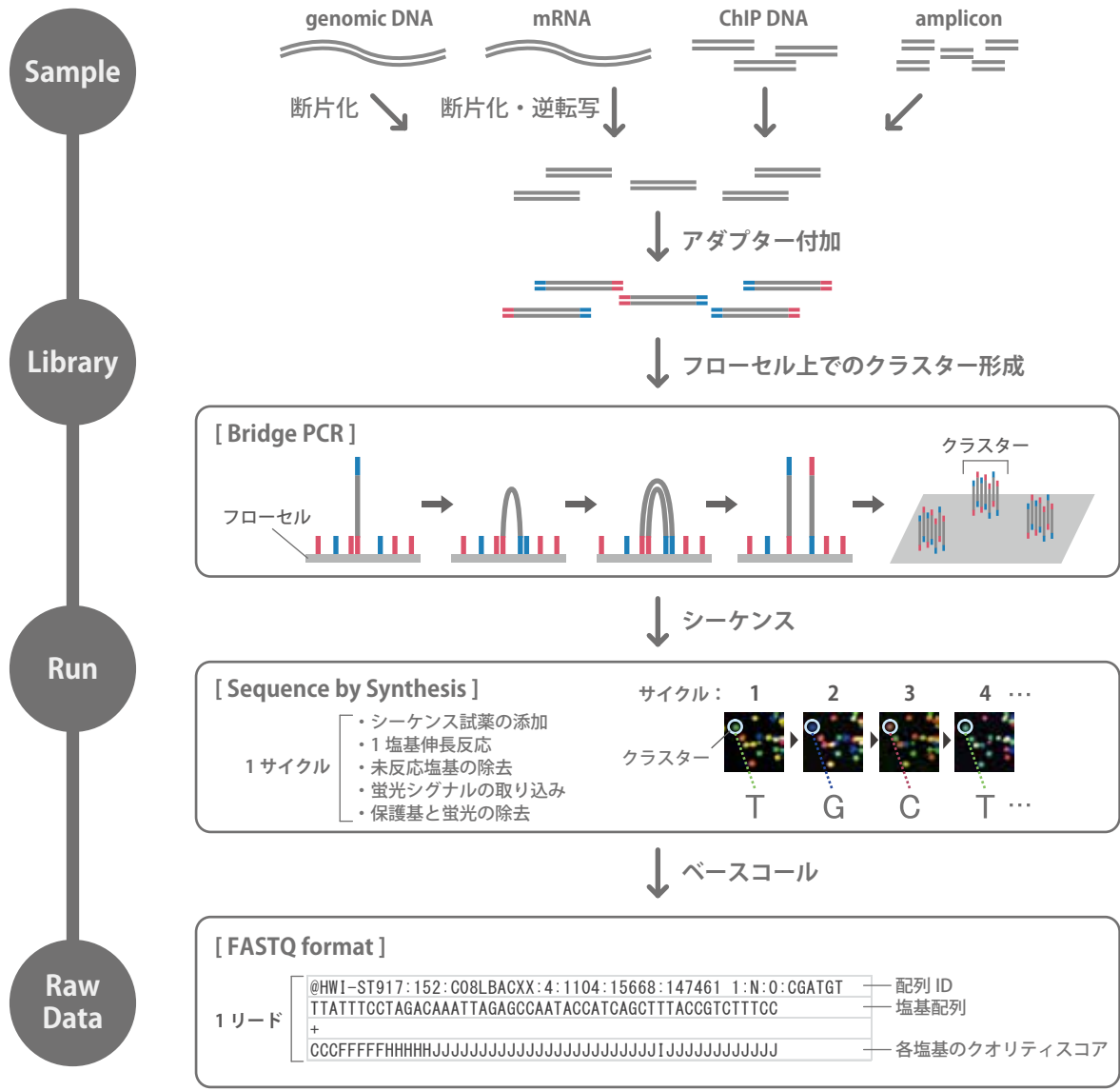
## マルチプレックス解析

アダプター配列には、サンプル識別用のインデックス配列が含まれており、1つのレーンで複数のサンプルを混合してシーケンスすることができます（マルチプレックス解析）。出力されたリードデータは、インデックス配列に従い、サンプル毎に振り分けられます。



# シーケンスの原理

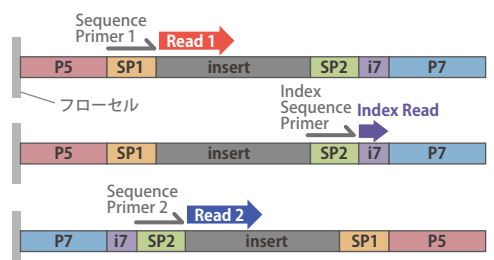
ゲノム DNA、RNA、アンプリコン等に由来する DNA 断片を作製し、アダプターを付加します。アダプター配列を介して、ライブラリをフローセル上に結合し、クラスター（同一ライブラリの集合）を形成します。各クラスターにて、1塩基伸長と蛍光読み取りのサイクルを繰り返す事で、フローセル上で並列的な大量シーケンスが行われます。



## リードの読み取り

リード1、リード2、インデックスの配列は、それぞれ異なるシーケンスプライマーにより読み取られます。同一クラスター由来のリード配列とインデックス配列が、データ上で関連付けられて出力されます。

- (1) Sequence Primer 1 を用いて、リード1 配列をシーケンス
- (2) Index Sequence Primer を用いて、インデックス配列をシーケンス
- (3) テンプレートの相補鎖を合成し、Sequence Primer 2 を用いて、リード2 配列をシーケンス



- P5, P7 : フローセル上での結合・増幅に使用される配列
- SP1, SP2 : シーケンスプライマーが結合する配列
- i7 : マルチプレックス解析用のインデックス配列

# ゲノムシーケンス解析

## DNA シーケンス解析

### 解析概要

ゲノム DNA を超音波により断片化し、得られた DNA 断片の両端にアダプターをライゲーションします。アセンブル解析や、リシーケンス変異検出に用いるシーケンスデータを取得します。

#### PCR Free のライブラリ調製

TruSeq DNA ライブラリ調製キットには、アダプター付与後に PCR 増幅を行う方法と、PCR 増幅を行わず最終ライブラリとする方法 (PCR-Free) があります。PCR 工程を含むキットでは、PCR-Free の調製に比べて少ない DNA サンプルからライブラリを作製する事ができます (下記、サンプル量参照)。

一方、PCR を行うと、ライブラリ中で PCR 増幅効率のバイアスが生じ、ゲノム上で GC 含量の高い領域等、配列構造によってはカバレッジが少なくなる可能性があります。また、リシーケンスで変異頻度を重視する解析の場合、PCR duplicate により生じたリードが存在すると、正確な変異頻度を算出する事が難しくなるため、PCR-Free の使用を推奨しております。

### サンプル必要量

#### ● DNA-Seq (PCR-Plus)

・精製済みゲノム DNA : 1  $\mu$ g (40ng/ $\mu$ l) 以上

#### ● DNA-Seq (PCR-Free)

・精製済みゲノム DNA : 5  $\mu$ g (40ng/ $\mu$ l) 以上

※ Nuclease Free Water に溶解してください。

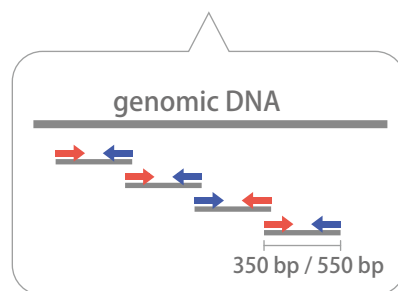
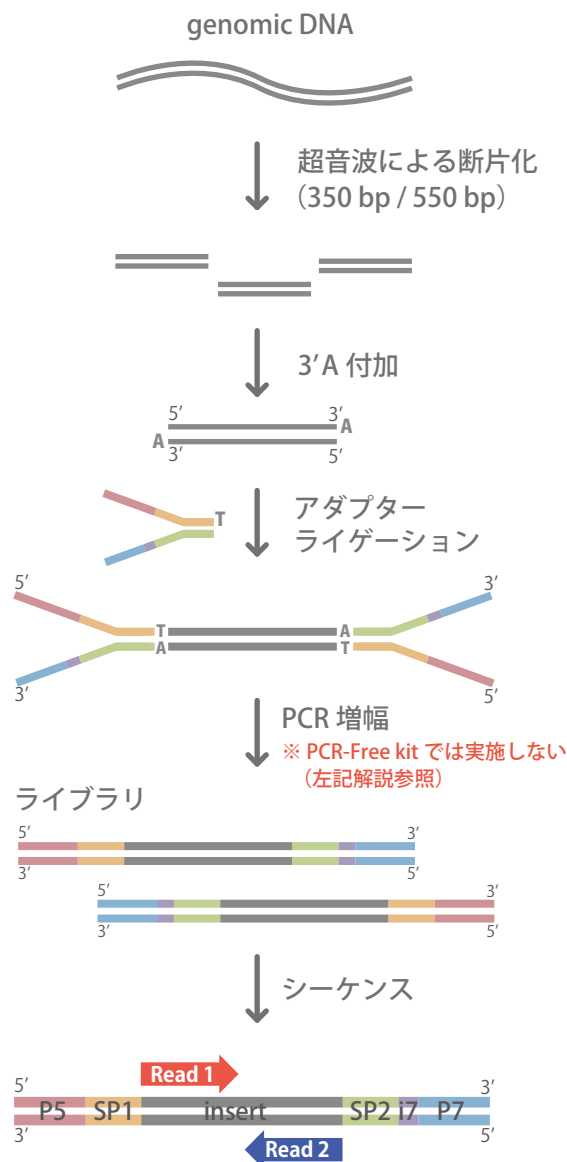
※ 分光光度計での濃度測定は、遊離のヌクレオチド等の夾雑物を検出し、正確な値が出ない可能性がありますので、アガロースゲル電気泳動での既知濃度の DNA とのバンドの蛍光強度比較に基づく測定を推奨しております。

### 取得データ量の目安

De novo アセンブル解析では、ゲノムサイズの 200 $\times$  以上のデータ取得を推奨しております。

リシーケンス変異解析の内、germline mutation の検出にはゲノムサイズの 30 $\times$  以上のデータ取得を目安としております。somatic mutation 等、低頻度変異の解析では、対象の変異頻度を想定した上で、その変異がシーケンスエラーと区別できる程度のカバレッジが得られるデータ量を取得する必要があります。

#### [ TruSeq DNA ライブラリ ]



# メイトペアシーケンス解析

## 解析概要

数 Kb の長いゲノム断片 (インサート) の両末端に相当する配列を含むシーケンスライブラリを作製します。得られたメイトペアリードは、通常のペアエンドリードと合わせてアセンブル解析に使用されます。

## サンプル必要量

- ・精製済みゲノム DNA : 10 $\mu$ g (200ng/ $\mu$ l) 以上
- ※ Nuclease Free Water に溶解してください。

## 🗨️ ゲノムアセンブル解析

De novo アセンブル解析では、インプットのデータ量が多い程、アセンブル効率が上がりますが、一種類のインサートサイズのペアエンドリードデータのみではカバレッジを増加しても繋がらないゲノム配列構造の領域もあります。インサートサイズの異なるメイトペアリードのデータを加える事で、そのような領域のスキファールディングが行われ、アセンブル効率が向上します。

# ChIP シーケンス解析

## 解析概要

ChIP 処理を行ったゲノム DNA 断片にアダプターを付与し、ライブラリを作製します。

得られたシーケンスリードをゲノムにマッピングし、対象因子が結合していたゲノム領域を検出することができます。

## 🗨️ コントロールサンプルについて

IP を行わない DNA サンプル (input) のシーケンスデータを取得した場合、input サンプルのマッピング結果をバックグラウンドとして、IP サンプルのマッピング結果を補正する事で、ピーク検出の特異性が向上します。

## 🗨️ 断片化について

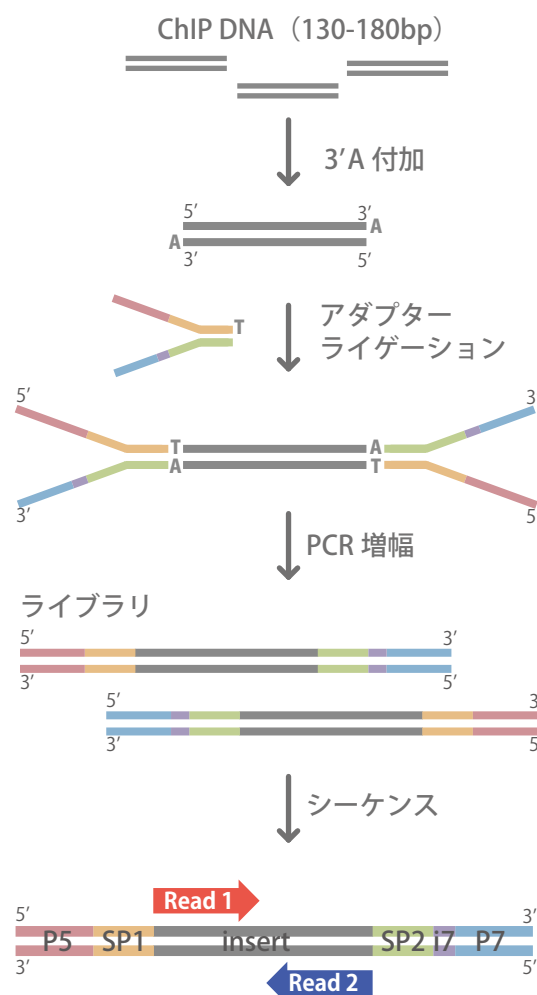
ChIP-DNA サンプルの断片配列長は、130 ~ 180 bp となるようご用意ください。DNA 断片のサイズが大きい場合には、更に断片化処理を行ってからライブラリ作製を行いますが、IP 処理時に該当サイズまで断片化が行われた場合に比べ、検出されるピーク領域が broad になる事が考えられます。

## サンプル必要量

- ・精製済み ChIP-DNA : 50ng (10ng/ $\mu$ l) 以上
- ※ サイズレンジ : 130 ~ 180 bp

- ※ Nuclease Free Water に溶解してください。
- ※ 分光光度計では正確な測定ができない可能性があるため、市販の 2 本鎖 DNA 定量キット (Qubit fluorometer、Invitrogen 社 Quant-iT dsDNA kit 等) での測定を推奨しております。

## [ TruSeq ChIP ライブラリ ]



# ターゲットシーケンス解析

## ● ターゲットシーケンス

ゲノムサイズの大きい高等生物でも、解析対象のゲノム領域を限定する事により、比較的少ないデータ量で変異解析を行う事ができます。

ターゲット解析の手法としては、(1) 対象領域を PCR で増幅する方法、(2) ゲノム DNA を断片化した後、キャプチャプローブで対象領域を濃縮する方法があります。

## エクソーム解析 (Agilent SureSelect)

### 解析概要

ゲノム DNA を断片化し、アダプターを付加した後、エクソン領域をキャプチャするオリゴプローブを用い、エクソン領域の DNA 断片を濃縮します。

Agilent SureSelect では、[Human・Mouse・Bovine・Zebrafish] のエクソーム解析に対応した既製のデザインがあります。その他の生物種についてもカスタムキットを使用して設計する事が可能です。

### サンプル必要量

・精製済みゲノム DNA : 5  $\mu$ g (40ng/ $\mu$ l) 以上

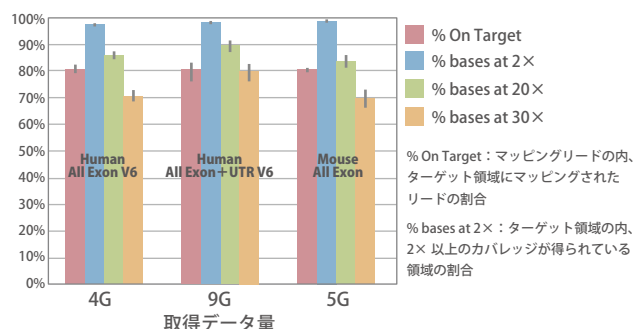
※ Nuclease Free Water に溶解してください。

※ 分光光度計での濃度測定は、遊離のヌクレオチド等の夾雑物を検出し、正確な値が出ない可能性がありますので、アガロースゲル電気泳動での既知濃度の DNA とのバンドの蛍光強度比較に基づく測定を推奨しております。

### 取得データ量の目安

SureSelect を使用したエクソーム解析では、ターゲットの 8 割以上の領域で 20 $\times$  以上のカバレッジが見込まれる程度のデータ取得を目安としてご案内しております。低頻度変異 (somatic mutation 等) の解析では、取得データ量を更に増加する事をお勧めいたします。

### ◆ SureSelect 取得データ量とカバレッジの関係



## Agilent SureSelect エクソームキャプチャキット

	ターゲットサイズ	対象領域
SureSelect Human All Exon V6	60 Mb	ヒトのコーディングエクソン領域
SureSelect Human All Exon V6 + COSMIC	66 Mb	上記 Human All Exon V6 に、Pseudogene 等の COSMIC 領域を追加
SureSelect Human All Exon V6 + UTR	91 Mb	上記 Human All Exon V6 に、非コード UTR 領域を追加
SureSelect Human Mouse All Exon	52 Mb	マウスのコーディングエクソン領域

## カスタムターゲットシーケンス

下記のキットを使用して、任意のゲノム領域を対象としたカスタムパネルを作製する事ができます。

	ターゲットサイズ	設計について
Roche SeqCap EZ Choice / Choice XL Library	- 7 Mb / -200 Mb	デザインツール Nimble Design を使用して設計
Roche SeqCap EZ Developer Library	-200 Mb	
Agilent SureSelect Custom	-0.5 / -2.9 / -5.9 / -11.9 / -24 Mb	デザインツール SureDesign を使用して設計
Agilent HaloPlex Custom	-0.5 / -2.5 / -5 Mb	

※ デザインに関しては、キット販売元のメーカーにお問い合わせください。

# ターゲットシーケンス 既製パネル

ヒトのがんや疾患に関連する遺伝子をターゲットとした既製のパネルです。

	ターゲットサイズ	対象領域
<b>Agilent ClearSeq SS</b>		
ClearSeq SS Comprehensive Cancer	0.79 Mb	がんに関連する 151 遺伝子 ※ 下記遺伝子リスト参照
ClearSeq SS 遺伝性疾患	10.5 Mb	約 4000 種の遺伝性疾患に関連する 2742 遺伝子
ClearSeq SS DNA Kinome	3.2 Mb	キナーゼ遺伝子とがん関連遺伝子
<b>Agilent ClearSeq Halo</b>		
ClearSeq Halo Cancer	10 kb	がんや薬剤標的との関連性が知られる 47 遺伝子のホットスポット
ClearSeq Halo AML	23.7 kb	骨髄性悪性腫瘍で変異が見られる 20 遺伝子
ClearSeq Halo 心筋症	208.4 kb	心筋症に関連する 34 遺伝子
ClearSeq Halo 遺伝性結合組織病	143 kb	結合組織病に関連する 40 遺伝子
ClearSeq Halo 遺伝性不整脈関連疾患	83 kb	4 種の遺伝性不整脈心疾患に関連する 19 遺伝子
ClearSeq Halo ヌーナン症候群	41 kb	RAS/MAPK 症候群に関連する 13 遺伝子
ClearSeq Halo ICCG	463 kb	ICCG が定義する約 180 遺伝子 ※ ICCG: International Collaboration for Clinical Genomics
ClearSeq Halo X 染色体	3.2 Mb	X 染色体上にある遺伝子のコーディングエクソン、UTR 領域
<b>illumina TruSight</b>		
TruSight Tumor 15	44 kb	固形がんに変異の見られる 15 遺伝子
TruSight One Sequencing Panel	12 Mb	臨床的表現型に関連する約 4800 遺伝子
TruSight Myeloid Sequencing Panel	141 kb	骨髄性悪性腫瘍で体細胞変異の見られる 54 遺伝子
TruSight Cancer	255 kb	がんの素因となる生殖細胞変異の見られる 94 遺伝子
TruSight Cardio Sequencing Kit	572 kb	17 種の遺伝性心臓疾患に関連する 174 遺伝子
TruSight Inherited Disease	2.25 Mb	小児発症疾患に関連する 552 遺伝子

## ClearSeq SS Comprehensive Cancer ターゲット遺伝子リスト

ABCB1	ASXL1	CDKN2A	CYP19A1	DNMT3A	EZH2	FSTL5	IDH2	KDM6A	MAP2K4	MTOR	PDGFRB	PTCH1	ROS1	SLC22A2	SNCAIP	TRRAP
ABCC2	ATM	CDKN2B	CYP2A6	DPYD	FBXW7	GNA11	IKZF1	KDR	MAP3K1	MYC	PHF6	PTEN	RPS6KB1	SLC31A1	SOS1	TYK2
ABL1	ATRX	CEBPA	CYP2B6	EGFR	FGFR1	GNAQ	IL2RA	KIT	MAPK1	MYD88	PIK3CA	PTPN11	RUNX1	SLC34A2	SPRED1	UGT1A1
ABL2	BRAF	CHD7	CYP2C19	ERBB2	FGFR2	GNAS	IL2RB	KRAS	MED13	NELL2	PIK3R1	RAF1	RXRA	SLC45A3	SRC	VHL
AKT1	BRCA1	CHIC2	CYP2C9	ERBB3	FGFR3	GSTP1	IL2RG	LAMA2	MET	NF1	PSMB1	RARA	RXRB	SLC01B1	STK11	WT1
AKT2	BRCA2	CREBBP	CYP2D6	ERBB4	FGFR4	H3F3A	INPP4B	LCK	MLH1	NOTCH1	PSMB2	RARB	RXRG	SMAD4	SUFU	YES1
AKT3	CBL	CRLF2	DDR1	ERG	FLT1	HNF1A	JAK1	LTK	MLL	NPM1	PSMB5	RARG	SHH	SMARCA4	TAS2R38	ZMYM3
ALK	CDA	CSF1R	DDR2	ESR1	FLT3	HRAS	JAK2	MAP2K1	MPL	NRAS	PSMD1	RB1	SHOC2	SMARCB1	TET2	
APC	CDH1	CTNNB1	DDX3X	ESR2	FLT4	IDH1	JAK3	MAP2K2	MST1R	PDGFRA	PSMD2	RET	SLC22A1	SMO	TP53	

## アンプリコンを使用した変異解析

### ● 対象領域が 10kb 程度の場合

ターゲット領域の全長が 10 kb 程度である場合、該当領域をカバーする数 Kb ずつのアンプリコンをサンプルとして、ホールゲノム解析と同様のライブラリ調製（アンプリコンの断片化、アダプター付加）を行い、ターゲット領域のシーケンスデータを得る事ができます。データ解析では、対象領域のリファレンス配列に対してリードをマッピングし、変異検出を行います。

#### ◇ ターゲットの例

- ・ 数個の遺伝子の cDNA 全長配列
- ・ 1 つの遺伝子のエクソン・イントロンを含むゲノム配列

### ● 対象領域が ~ 500bp 程度の場合

特定のゲノム領域について、サンプルに含まれる多型配列の種類や存在割合を調べる事を目的とする場合には、該当領域 ~ 500bp 程度のアンプリコンをサンプルとして、配列全長のシーケンス解析を行う事が可能です。

※ P10「アンプリコンシーケンス解析」をご参照ください。

#### ◇ ターゲットの例

- ・ 遺伝子の一部領域の体細胞変異解析
- ・ TCR 遺伝子, BCR 遺伝子の多型解析
- ・ ゲノム編集領域
- ・ ウイルスの多型領域

# トランスクリプトーム解析

## RNA シーケンス解析

### 解析概要

mRNA を断片化、逆転写して得られた cDNA 断片の両端にアダプターを付加します。  
発現解析や、De novo トランスクリプトーム解析に用いるシーケンスデータを取得します。

#### 💬 ストランド情報を維持したライブラリ

TruSeq stranded RNA ライブラリ調製キットでは、元の RNA の転写方向性を維持したリードデータが得られます。mRNA の逆転写により 1st 鎖 cDNA を合成した後、2nd 鎖 cDNA 合成時に、dTTP ではなく dUTP を使用する事で、最終工程の PCR において 2nd 鎖 cDNA の増幅が抑制されます（右図参照）。その結果、リードデータのほとんどが 1st 鎖 cDNA 由来となります。リードのストランド情報を利用して、マッピングやアセンブル解析の精度を高める事ができるため、stranded RNA ライブラリ調製キットの使用をお勧めしておりますが、過去に行われたシーケンス解析と条件を揃えたい場合等、ストランド情報の無いライブラリ調製キットを使用する事も可能です。

#### 💬 rRNA 除去方法について

##### ● バクテリア

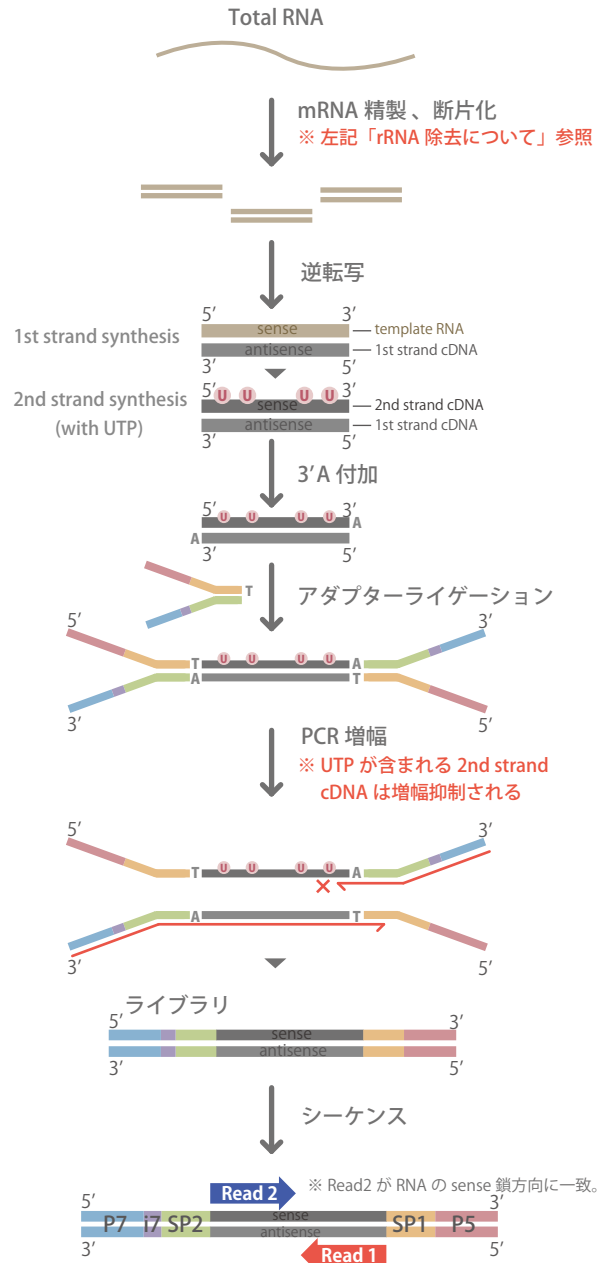
バクテリアの RNA-Seq 解析では、Total RNA サンプルに対し、rRNA キャプチャ除去キット「Ribo-Zero」を使用した rRNA 除去を行います。

##### ● 真核生物

真核生物の RNA-Seq ライブラリ調製では、一般的に oligo-dT ビーズを用いた poly(A)-RNA の精製を行います。poly(A) tail を持たない一部の long non coding RNA を解析対象とする場合には、poly(A)-RNA 精製ではなく、rRNA キャプチャ除去キット「Ribo-Zero」を使用した rRNA 除去を行う事も可能です（※対応可能な生物種については、下記 URL をご参照ください）。また、RNA の分解が進んだサンプルでは、poly(A)-RNA 精製を行う事で、mRNA の 3' 側に偏ったデータとなる可能性があるため、Ribo-Zero を用いた rRNA 除去に変更する事も可能です。（ヒト由来サンプルの場合、P8「ヒト FFPE サンプルを用いた解析」もご参照ください。）

※ Ribo-Zero rRNA Removal Kits 対応生物種について  
<http://www.illumina.com/products/rrna-removal-kit-species-compatibility.html>

### [ TruSeq stranded RNA ライブラリ ]



### サンプル必要量

#### ● mRNA-Seq

- 精製済み Total RNA : 5  $\mu$ g (40ng/ $\mu$ l) 以上
- 精製済み mRNA : 200ng (10ng/ $\mu$ l) 以上

#### ● Ribo-Zero + mRNA-Seq

- 精製済み Total RNA : 10  $\mu$ g (40ng/ $\mu$ l) 以上

※ Nuclease Free Water に溶解してください。



### トランスクリプト配列決定

ゲノムの解読されていない生物等で、網羅的なトランスクリプト配列決定を目的とした場合（発現定量を行わない場合）は、複数組織由来の RNA を混合したサンプルを用いる事で、各組織で特異的な発現プロフィールが均一化されたデータが得られます。

### 微量 RNA サンプルを用いた解析

数細胞由来の微量 Total RNA を用いた RNA-Seq 解析では、cDNA の増幅を行う必要があります。quartz-Seq 法 (Sasagawa et al., 2013) は、① 逆転写時のアニール温度の最適化、② cDNA 増幅に使用する PCR 酵素の選定、③ PCR 時にプライマーダイマー等の短い非特異産物の増幅を抑制するプライマーを用いる事で、増幅バイアスが少なく検出感度の高い RNA-Seq を行う手法とされています。

### 解析対象の RNA サイズ

RNA-Seq 用ライブラリ調製では、ビーズ精製工程にて低分子が除去されるため、100 bp 以下の短い RNA は解析対象から除外されます。small RNA を対象とした解析については、別途、small RNA-Seq 用ライブラリを調製する必要があります（下記参照）。

### ヒト FFPE サンプルを用いた解析

TruSeq RNA access キットでは、ヒトのコード領域をキャプチャするプローブを用い、poly(A) に依存しない mRNA 精製を行い、分解の進んだ少量の Total RNA から RNA-Seq 用ライブラリを作製する事が可能です。なお、分解の度合いによっては、品質の高い RNA サンプルを使用した解析に比べ、マッピング結果や、その後の発現解析、融合遺伝子解析の精度が劣る場合がございます。

## small RNA シーケンス解析

### 解析概要

TruSeq small RNA ライブラリ調製キットでは、Total RNA をサンプルとして、small RNA に対するアダプター付加と PAGE・切り出し精製によるサイズセレクションを行い、small RNA 由来配列のシーケンスライブラリを作製します。リード 1 本が small RNA 1 本に相当し、同一配列の集計を行う事により、各 small RNA の相対的な存在量を算出できます。

### 解析対象の small RNA

TruSeq small RNA ライブラリ調製キットで用いられる 3' アダプターは、生体内で RNA プロセッシングを受けた small RNA (Dicer により切断された miRNA 等) の末端に見られる 3'-OH に結合します。そのため、機能的 small RNA 以外の RNA 分解産物等はライブラリ化されず、Total RNA サンプルから効率良く miRNA をライブラリ化する事ができます。

### サンプル必要量

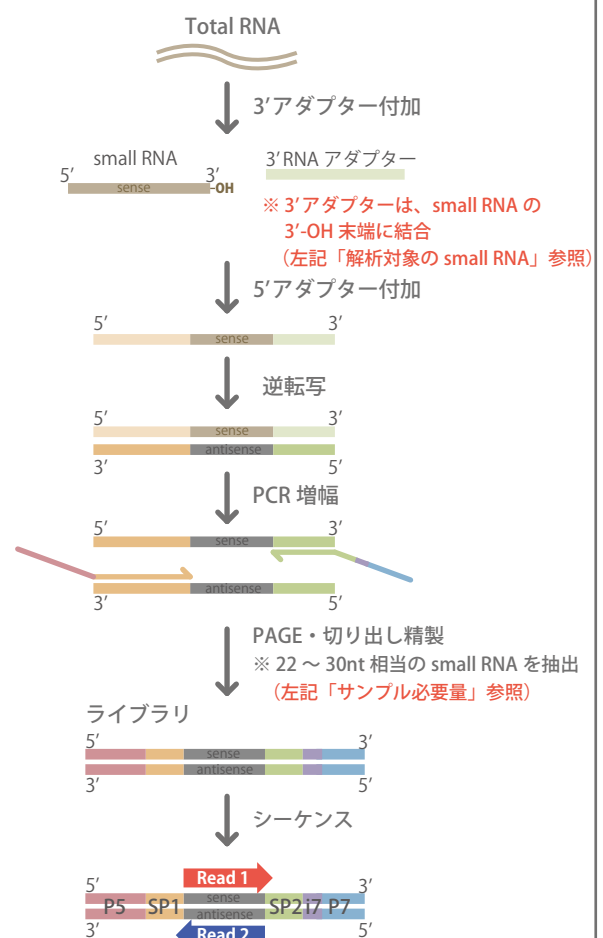
#### ● small RNA-Seq

- ・精製済み Total RNA : 5 µg (200ng/µl) 以上
- ・精製済み small RNA : 200ng (10ng/µl) 以上

※ Nuclease Free Water に溶解してください。

※ PAGE・切り出し精製では、通常 22-30 nt 相当の small RNA 配列を抽出しますが、解析対象とされる small RNA のサイズにより、切り出しサイズを変更する事も可能です。

### [ TruSeq small RNA ライブラリ ]



# 16S rRNA 細菌叢解析

## 解析概要

土壌や糞便等の環境サンプルから抽出したメタゲノム DNA をテンプレートとして、バクテリア 16S rRNA 領域のアンプリコン (約 450bp) を作製します。アダプター配列が付加されたプライマーを用い、PCR によるライブラリ調製を行います。

MiSeq 300bp ペアエンドシーケンスで、アンプリコン配列全長を決定し、菌叢解析に使用するデータを取得します。

## サンプル必要量

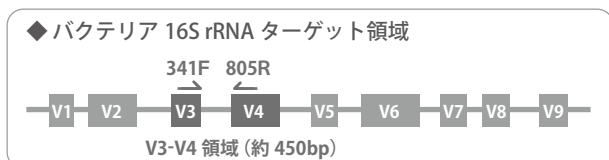
- ・精製済み 1st PCR 産物 : 5ng (0.2ng/μl) 以上 (オーバーハング配列を含むアンプリコンサイズ: 約 550 bp)
- ・精製済み 2nd PCR 産物 : 150ng (5ng/μl) 以上 (アダプター配列を含むアンプリコンサイズ: 約 630 bp)

※ 10mM Tris pH8.5 Buffer、又は Nuclease Free Water に溶解してください。

※ 分光光度計での濃度測定は、遊離のヌクレオチド等の夾雑物を検出し、正確な値が出ない可能性がありますので、アガロースゲル電気泳動での既知濃度の DNA とのバンドの蛍光強度比較に基づく測定を推奨しております。

## 16S rRNA ターゲット領域

バクテリア 16S rRNA V3-V4 領域を対象としたプライマーセットをご提供いたします。



※ プライマー配列の詳細については、イルミナ社公開資料「16S Metagenomic Sequencing Library Preparation」にてご確認ください。

※ 本プライマーセットは、バクテリア 16S rRNA 解析用であり、アーキアの 16S rRNA には対応していません。

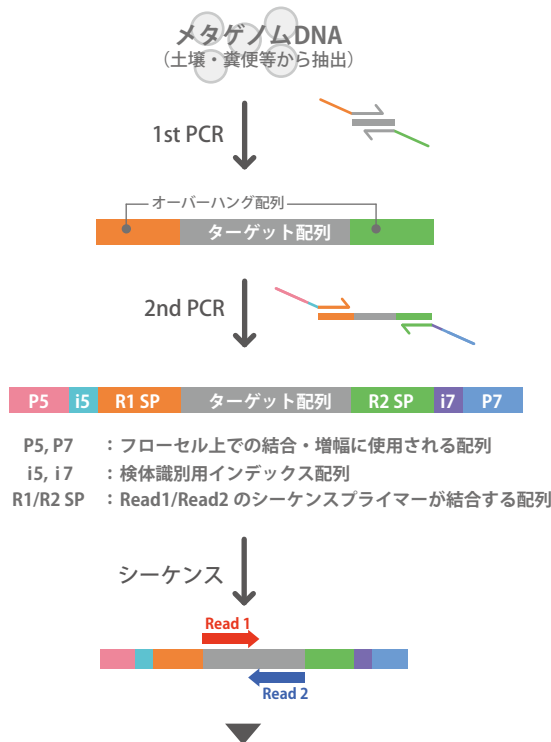
※ 本プライマーで増幅され得る菌種の詳細は、RDP (<https://rdp.me/msu.edu/>) の Probe Match にてご確認ください。

### 植物由来配列の混入について

植物が混入する可能性のあるサンプルでは、植物細胞の葉緑体、ミトコンドリア由来の 16S rRNA 配列が多量にデータに含まれる可能性があります。

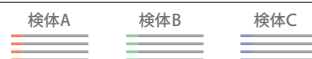
データ解析時に、植物由来と推定された配列を除外する事も可能ですので、事前にご相談ください。

## [ 16S メタゲノムシーケンス ]

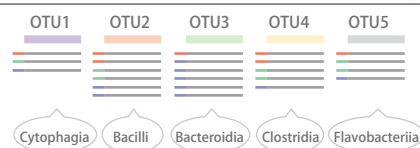


## データ解析：細菌叢解析

### インデックスによるリードデータの検体識別



### 全サンプルのリードから OTU を作成・Taxonomy 情報付与



### 各 OTU に帰属するリード数を基に菌叢グラフ作成



※ 本解析は、イルミナ社公開資料「16S Metagenomic Sequencing Library Preparation」に従い実施しております。

ライブラリ調製の詳細プロトコルに関しては、弊社注文書 (16S 群集解析あいのリプラン) 内、プロトコルシートをご参照ください。

# アンプリコンシーケンス解析

## 解析概要

500 bp 程度のアンプリコンを対象としたディープシーケンスを行います。

ゲノム編集領域や、細胞・ウイルス集団内の多型領域を対象とした解析では、得られたアンプリコン配列を集計し、配列多型の種類と存在割合を算出できます。

また、rRNA や ITS 領域等、マーカー遺伝子を用いた生物叢解析に使用するデータを取得することもできます。

※ 16S rRNA 細菌叢解析については P9 をご参照ください。

## サンプル必要量

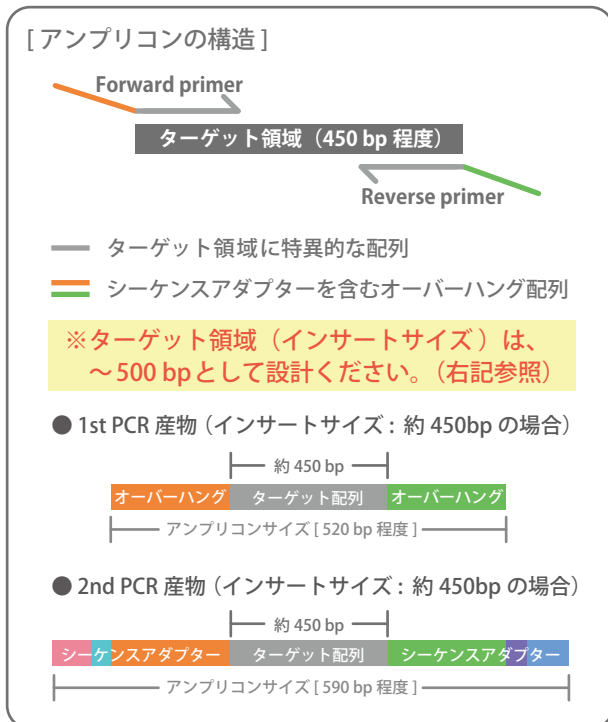
- ・精製済み 1st PCR 産物 : 5ng (0.2ng/μl) 以上
- ・精製済み 2nd PCR 産物 : 150ng (5ng/μl) 以上

※ 10mM Tris pH8.5 Buffer、又は Nuclease Free Water に溶解してください。

※ 分光光度計での濃度測定は、遊離のヌクレオチド等の夾雑物を検出し、正確な値が出ない可能性がありますので、アガロースゲル電気泳動での既知濃度の DNA とのバンドの蛍光強度比較に基づく測定を推奨しております。

## アンプリコンの構造

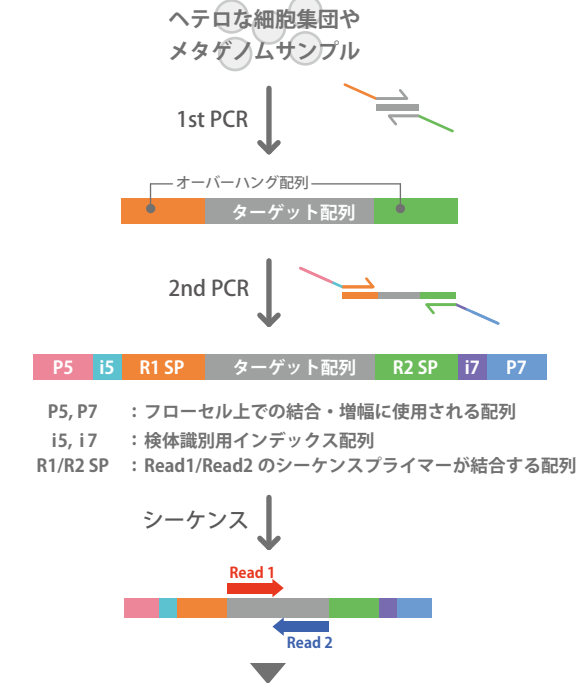
指定のオーバーハング配列を付与したプライマーをご用意いただき、アンプリコンを作製してください。



※ 本解析は、イリミナ社公開資料「16S Metagenomic Sequencing Library Preparation」に従い実施しております。

オーバーハング配列、及び、ライブラリ調製の詳細プロトコルに関しましては、弊社注文書 (Amplicon Deep Sequence 解析注文書) 内、プロトコルシートをご参照ください。

## [アンプリコンシーケンス]



## [データ解析例]

### ● ペアリードの連結



### ● リードカウント

Amplicon Seq_1	Amplicon Seq_2	Amplicon Seq_3
TACCCTGTATAGCTGTA	AGCCTACTTACGGGTAT	CATTITGACACTACGAC
TACCCTGTATAGCTGTA	AGCCTACTTACGGGTAT	
TACCCTGTATAGCTGTA	AGCCTACTTACGGGTAT	
	AGCCTACTTACGGGTAT	
	AGCCTACTTACGGGTAT	

## 💬 インサート配列の塩基長について

MiSeq 300 bp ペアエンドシーケンス解析において、アンプリコンシーケンスの乗り合い解析を行っております。最大 550 bp 程度のインサート配列を読むことができますが、リードの後半にかけてシーケンスクオリティが低下する可能性を考慮し、~500 bp 程度のインサートサイズを推奨しております。また、乗り合い解析では塩基長が極端に異なるライブラリを同時に解析した場合、双方のデータに悪影響 (一方のデータ量が極端に少なくなる等) が出る可能性があるため、アンプリコンサイズは 450 ~ 500 bp 程度となるよう設計をお願いしております。

上記よりも短い、または長いサイズのアンプリコンを対象とした解析に関しては、他のシーケンス解析プランもございますので、ご相談ください。

# PacBio RS II

## PacBio シーケンサー

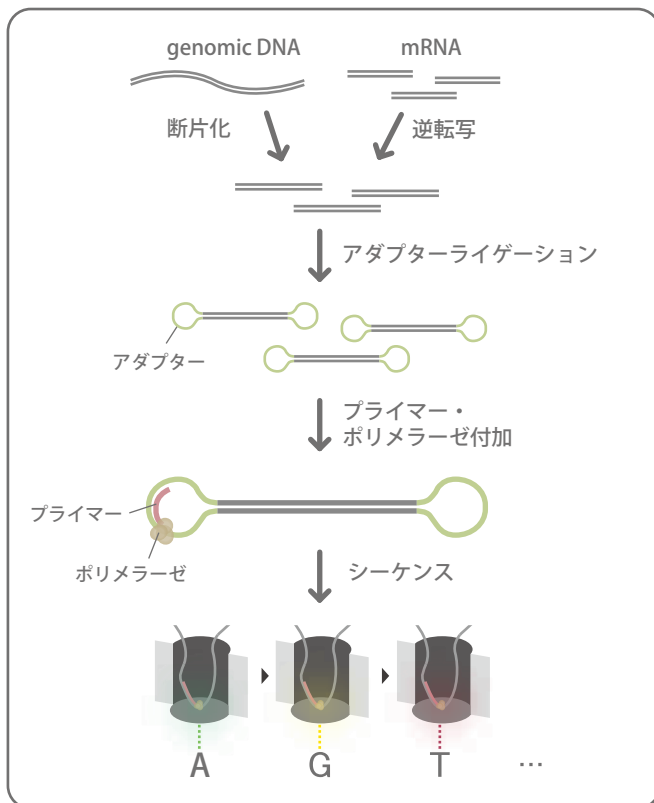
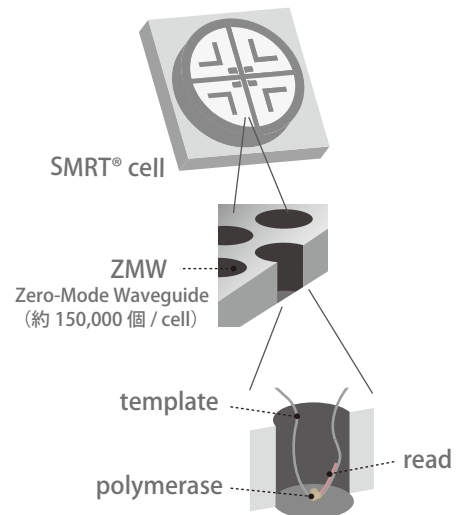
PacBio RS II では、数 kb のロングリードデータを出力します。ライブラリ調製時に PCR を行わないため、GC 含量に関わらず均一なシーケンスデータを取得することができます。バクテリア等ゲノムサイズの小さい生物種のゲノムアセンブル解析や、高等生物の完全長トランスクリプト配列決定に適しています。

## シーケンス解析の原理

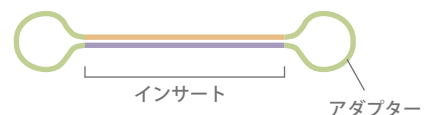
ゲノム DNA 断片、又は、全長 mRNA 由来の cDNA をインサートとして、両末端にヘアピン型のアダプターを付加し、環状のライブラリ (SMRT bell ライブラリ) を作製します。ライブラリのアダプター部分にプライマーとポリメラーゼを結合させ、SMRT cell にアプライします。SMRT cell には ZMW と呼ばれる約 15 万個の孔が開いており、1 個の ZMW の底面に 1 分子の DNA-ポリメラーゼ複合体が固定されます。1 塩基伸長反応の際に塩基毎に標識された蛍光が検出され、環状のライブラリを繰り返し周回する様にシーケンスが行われます。1 個の ZMW から 1 本のポリメラーゼリードデータが得られ、そこからアダプター配列を除いたサブリード配列を以降の解析に使用します。



PacBio RS II



### [ SMRT bell ライブラリ ]



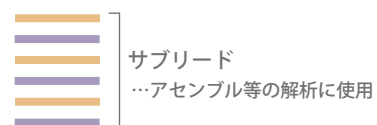
### [ ポリメラーゼリード ]

1 本の環状 SMRT bell ライブラリを繰り返しシーケンスし、アダプターとインサートが交互に出現するリードデータ



### [ サブリード ]

ポリメラーゼリードからアダプター配列を除去したリード配列



## 取得データ量について

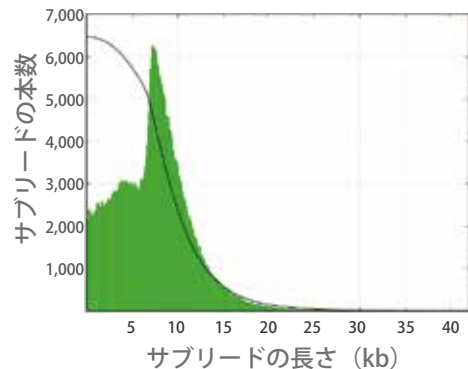
PacBio では、1 SMRT Cell あたり 300Mb ~ 1 Gb のデータが得られます。サンプルの品質や配列構造により、取得データ量、リード長は変動するため、下記の数値は参考値となります。

### 1 SMRT cell あたりのデータ量

サブリード長	約 ~ 10 Kb
サブリード数	約 40,000 ~ 60,000
データ量	約 300Mb ~ 1 Gb

※ 上記は弊社実測に基づく参考値であり、保証値ではありません。

◆ サブリードの長さの分布例 (緑色)



## DNA-Seq

ゲノム DNA をサンプルとしてシーケンスを行い、アセンブル解析に用いるデータを取得します。

### 解析セル数の目安

	ゲノムサイズ	解析セル数	アセンブル結果
Bacteria	~ 10 Mb	1 cell	1 ~ 10 contigs
Fungi	10 Mb ~	6 - 8 cell	100 ~ 500 contigs

※ 上記、アセンブル結果は、目安であり保証値ではありません。  
ゲノムの品質や、配列構造によりアセンブルの効率は異なります。

### サンプル必要量

・精製済みゲノム DNA : 20  $\mu$ g (150ng/ $\mu$ l) 以上

※ Nuclease Free Water に溶解してください。  
※ 分光光度計での濃度測定は、遊離のヌクレオチド等の夾雑物を検出し、正確な値が出ない可能性がありますので、アガロースゲル電気泳動での既知濃度の DNA とのバンドの蛍光強度比較に基づく測定を推奨しております。

### 推奨のゲノム抽出キット

- ・ Qiagen® MagAttract® HMW kit
- ・ Qiagen Genomic-tip kit
- ・ Qiagen Gentra® Puregene® kit

### サンプルの注意事項

- ① 二本鎖 DNA であること。  
一本鎖 DNA でのライブラリ調製はできません。
- ② 複数回の凍結融解を経っていないこと。
- ③ OD260/280 の比が 1.8 から 2.0 であること。
- ④ 不溶物を含まないこと。
- ⑤ RNA 混入が無いこと。
- ⑥ 挿入蛍光色素、又は紫外線放射に曝されていないこと。
- ⑦ キレート剤 (EDTA)、二価金属カチオン (Mg<sup>2+</sup>)、変性剤 (グアニジン塩、フェノール)、または浄化剤 (SDS、Triton-X100) などが含まれていないこと。
- ⑧ 組織 (ヘム、フミン酸、ポリフェノール、多糖類など) のキャリア・オーバーコンタミネーションがないこと。

## Iso-Seq

mRNA をサンプルとしてシーケンスを行い、トランスクリプト配列決定に用いるデータを取得します。

### サンプル必要量

・精製済み Total RNA : 10  $\mu$ g (50ng/ $\mu$ l) 以上  
(RIN 値 8 以上・OD260/280 値 1.8 - 2.0 以上)

※ Nuclease Free Water に溶解してください。  
※ 濃度測定については基本的に分光光度計で問題ございませんが、可能であれば Agilent2100BioAnalyzer 等での事前確認を行っていただく事を推奨しております。

### 解析セル数の目安

多様なインサートサイズのライブラリが混在している場合、SMRT cell には短いライブラリが優先的にロードされます。そのため、Iso-Seq 解析では、mRNA のサイズでライブラリを分割し (1-2 kb, 2-3 kb, 3-6 kb, >6 kb 等)、各サイズ分画 2 セル程度でシーケンスを行います。解析対象の遺伝子が限定されている場合、該当 mRNA を含むサイズ画分のみシーケンスを行う事も可能です。

※ PacBio ではライブラリサイズによりシーケンス量にバイアスが出るため、遺伝子発現量の算出はできません。

# ご依頼の流れ

## 1 事前のご相談

弊社ホームページのお問い合わせフォーム、または、お電話にてご連絡ください。  
必要に応じて弊社担当がご訪問することも可能ですので、お気軽にご相談ください。

HSS ホームページ <http://www.hssnet.co.jp>

「サポート・お問い合わせ」にございます  
お問い合わせフォームをご利用ください。

お電話でのお問い合わせ

 **0120-613-190**

解析目的、ご用意いただくサンプル、弊社作業内容、シーケンスデータ量についてご相談の後、  
お見積りをいたします。

## 2 ご注文書の受け付け

弊社ホームページより注文書をダウンロードいただくか、弊社担当者よりメールでご案内させていただきます。  
エクセルファイルの注文書にご記入いただき、Eメールに添付の上お送りください。

HSS ホームページ <http://www.hssnet.co.jp>

「ご注文（Online 申込）」の「次世代シーケンス」  
より該当の注文書をご選択ください。



お送り先（E-mail）：

**[hss-ngs@hssnet.co.jp](mailto:hss-ngs@hssnet.co.jp)**

※ 記載に関してご不明な点がございましたら、  
お気軽にご相談ください。

## 3 サンプルのお預かり

弊社よりサンプル送付セット（梱包用の箱、緩衝材、送料弊社負担の日通送付状）をお送りいたします。  
ドライアイス（5kg 程度）をご用意いただき同梱の上、下記までお送りください。

サンプル送付先：

〒001-0932 札幌市北区新川西2条1丁目2-1  
北海道システム・サイエンス株式会社  
解析部 次世代シーケンス解析サービス担当  
TEL：011-768-5903

お受け取りは平日（月～金）とさせていただきます。  
土日祝日のお受け取りはできませんので、ご注意ください。

※ 必要サンプル量や濃度測定方法に関しては、本誌の各アプリケーションページ、及び、  
弊社ホームページの「次世代シーケンス解析」…「サンプル必要量」をご確認ください。

## 4 受け入れサンプルの品質検査

お送りいただいたサンプルについて、弊社にて品質検査を行い、結果をご報告します。  
弊社基準に満たなかったサンプルについては、追加、再送をお願いするなど、取扱いについて  
ご相談させていただく可能性があります。

## 5 作業内容確認書のご案内

ご依頼の内容を記載した「作業内容確認書」をお送りいたしますので、ご確認ください。

## ⑥ ライブラリ調製

上記、品質検査を通過したサンプルを用いて、シーケンス用ライブラリを作製いたします。

※ ライブラリ作製の流れについては、本誌の各アプリケーションのページをご確認ください。

※ ライブラリの品質検査にて異常が見られた場合には、お取扱いをご相談させていただきます。

## ⑦ 次世代シーケンスのラン

上記、品質検査を通過したライブラリについて、次世代シーケンサーでのランを行います。

## ⑧ Raw データの出力

シーケンサーより出力された画像データについてベースコール、フィルタリングを行い、リードの塩基配列を取得します（Index 配列による各サンプルデータの振り分けも行います）。

Raw データ取得までのご依頼では、本工程で出力されるリードの塩基配列データを納品いたします。

※ 納品物の詳細は、冊子「次世代シーケンス解析サービスーデータ解析ー」をご参照ください。

## ⑨ データ解析・納品

各種アプリケーションのデータ解析までご依頼いただいた場合、上記 Raw Data を使用

してデータ解析を行います。納品物には、Raw Data、及び、解析結果のデータが含まれます。

データ解析  
のみのご依頼も  
承ります。

# データ解析のご案内

### ◆ 冊子「次世代シーケンス解析サービスーデータ解析ー」

各種アプリケーションのデータ解析をご紹介します。

シーケンス Raw データ / De novo ゲノムシーケンス解析 / ゲノムリシーケンス解析 / mRNA 発現頻度解析 / De novo トランスクリプトーム解析 / small RNA シーケンス解析 / 16S rRNA 細菌叢解析 / アンプリコンシーケンス解析 / CHIP シーケンス解析

### ◆ 次世代シーケンス解析 デモデータ

ご注文の前に、デモデータを確認する事ができます。ご希望がございましたら、お問い合わせください。

※ デモデータのご用意のないアプリケーションもございます。

ゲノム De novo アセンブル解析 / ゲノムマッピング+SNP 候補検索 / RNA De novo アセンブル解析 / RNA マッピング+発現頻度解析 / small RNA-Seq リードカウント+ miRBase BLAST / CHIP-Seq マッピング+ピーク検出

---

## 注意事項

- ※ 製品の規格仕様・サービス内容などにつきまして、予告なしに変更することがあります。
- ※ 受注後、実作業に入ってからキャンセルはサービスの仕様上お受けいたしかねます。  
やむを得ない場合は実作業分の料金をご請求させていただきます。
- ※ ご依頼いただくサンプルは、文部科学省の「遺伝子組換え生物等の使用等の規制による生物の多様性の確保に関する法律」(カルタヘナ法)におけるP1レベルのサンプルに限らせていただきます。
- ※ 感染性のあるサンプル(HCV・HIVなど)の受け入れは弊社では行っておりません。また、ヒト臨床サンプルの場合はインフォームドコンセントを得ていることをご確認ください。
- ※ 本解析サービスは、試験研究を目的にご利用ください。その他の目的(医療品・食品の製造・品質管理や医療診断など)には使用しないでください。
- ※ 本解析サービスにより得られた結果が原因となり生じた損失・損害等について、サービスの仕様上、責任を負いかねます。
- ※ ご提供・ご指示いただくサンプルや手法から生じる工業所有権・安全性などの問題については、一切責任を負いかねます。
- ※ ご提供いただくサンプルの保管および返却は行っておりません。予めご了承ください。

代理店

 **北海道システム・サイエンス株式会社**

〒001-0932 札幌市北区新川西2条1丁目2-1

 0120-613-190

TEL:011-768-5903 FAX:011-768-5951

E-mail:hss-ngs@hssnet.co.jp

URL:<http://www.hssnet.co.jp>