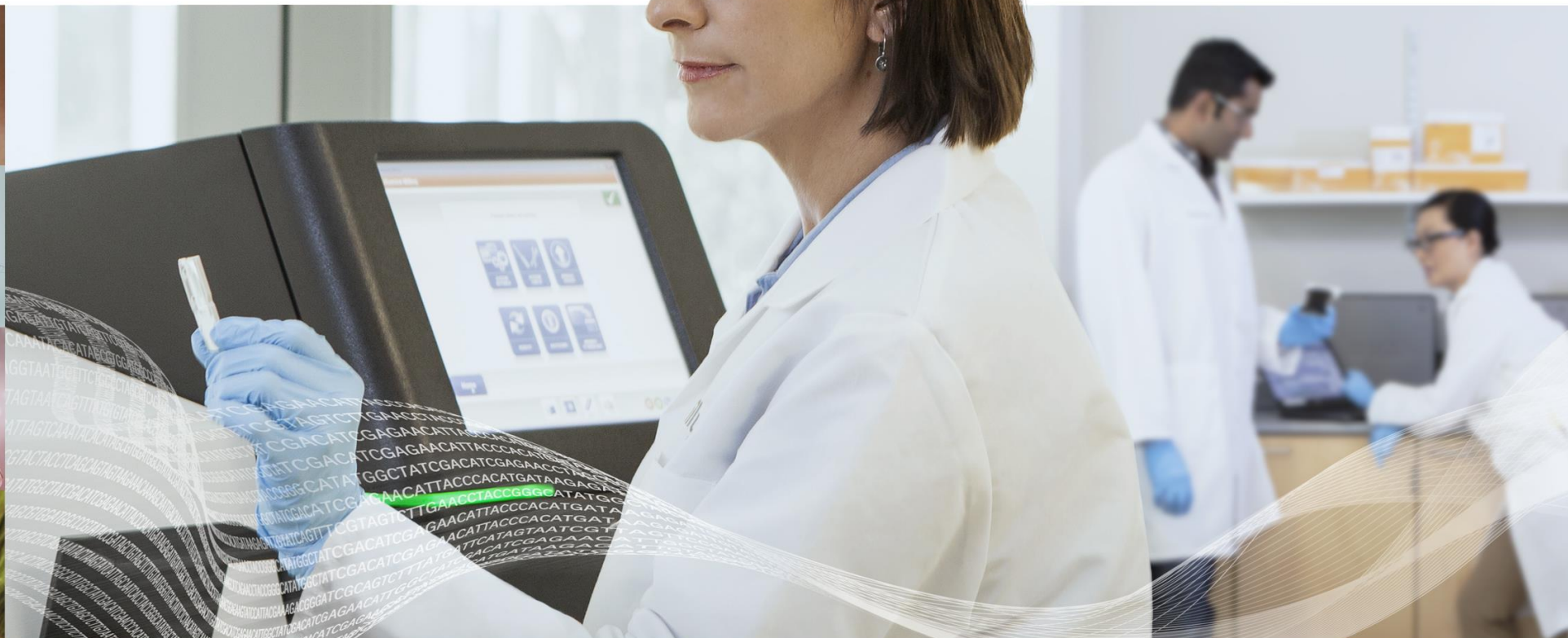


ウイルスゲノムシーケンス入門: ライブラリ調製から解析まで

2020年6月17日



シニアテクニカルアプリケーションサイエンティスト 渡邊 大
テクニカルアプリケーションサイエンティスト 片山 緩子

QB10174

© 2019 Illumina, Inc. All rights reserved.

For Research Use Only. Not for use in diagnostic procedures.

illumina®

はじめに：ウイルスゲノムシーケンスでわかること



ウイルス：ゲノムの種類と特徴

<ウイルスの一般的な特徴>

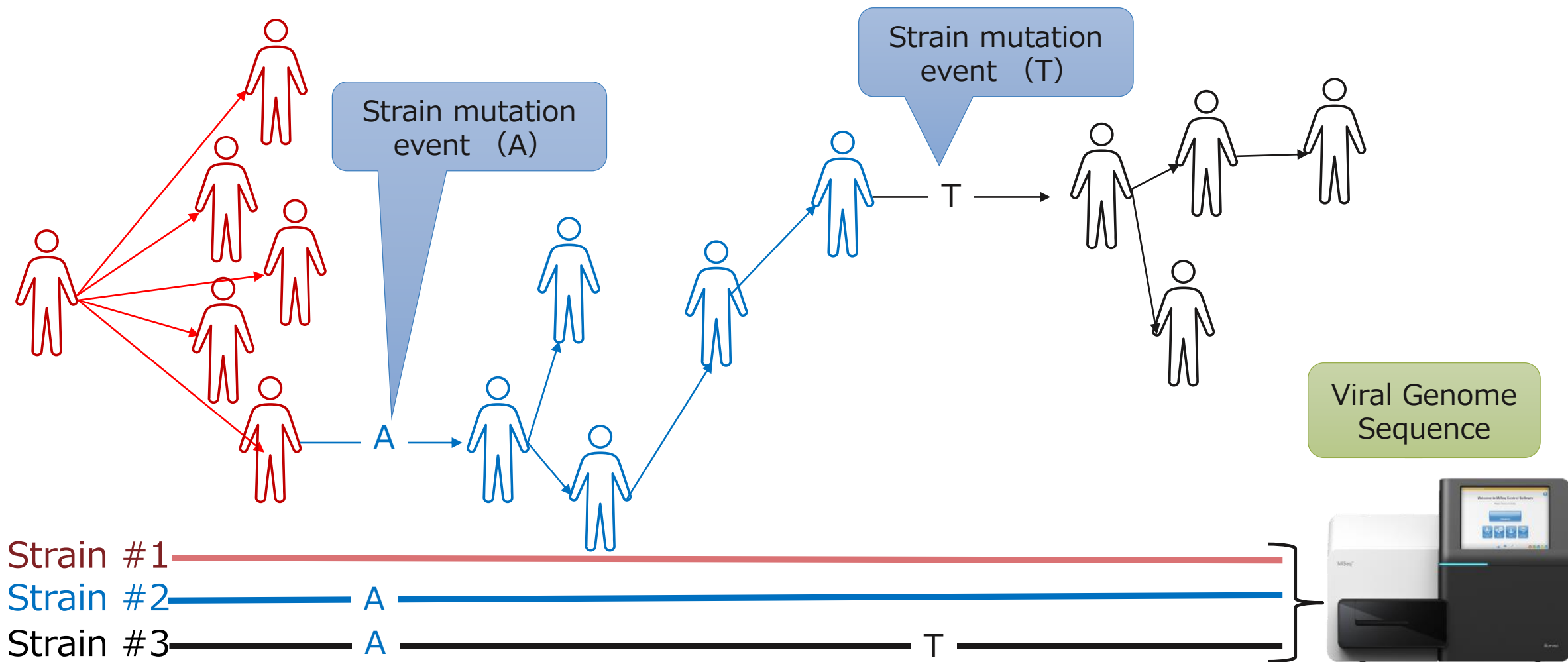
- ds/ssDNA あるいは ds/ssRNAに遺伝情報を持つ
- 自身では自己複製・増殖せず、宿主細胞内で増殖を行う
 - ウイルス培養には時間がかかる
- ウイルス粒子形状、構成はウイルスによって様々
 - ヌクレオカプシド
 - エンベロープ
 - スパイク など
- ゲノムの塩基配列のシーケンスにより、短時間で構成タンパク情報が得ることができる

→ ウイルスゲノム上の変異検出を時系列に沿って調べることで、疫学的な追跡調査が可能になる

ゲノムタイプとウイルスの分類*



ウイルスゲノム解析：変異解析による疫学研究

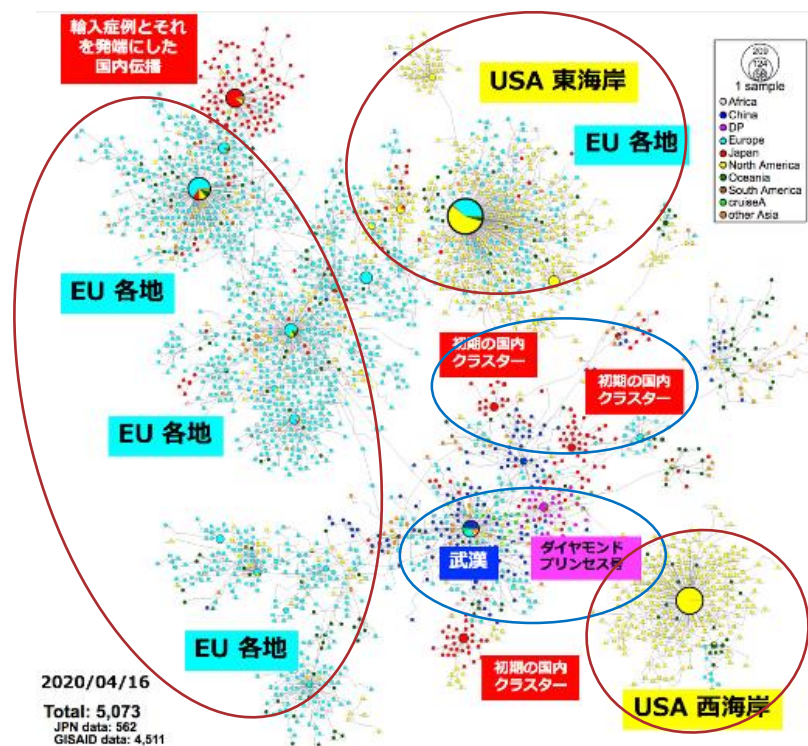


ウイルスゲノム上の変異情報を網羅的に得るためには、シーケンス解析が必須

変異解析から疫学調査へ：SARS-CoV-2研究

● 国立感染症研究所

日本国内の562 患者のSARS-CoV-2ゲノム配列からハプロタイプ・ネットワークを作成



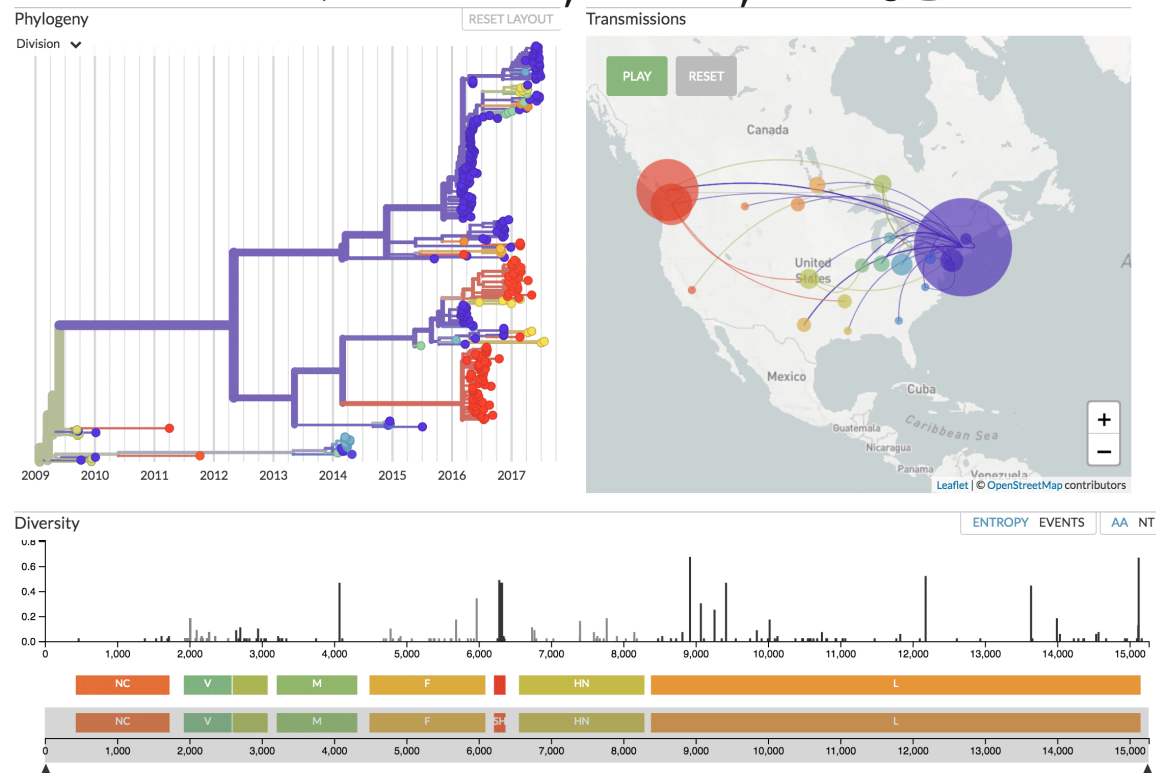
左) 国立感染症研究所ホームページ 2020年4月27日公開記事

<https://www.niid.go.jp/niid/ja/basic-science/467-genome/9586-genome-2020-1.html>

右) Nextstrain <https://nextstrain.org/>

● Nextstrain

GISAIDに登録されている世界中のウイルスゲノムデータについて系統樹解析などに特化したオープンソースプロジェクト。SARS-CoV-2、Influenza, Ebola, Zikaなど



NGS（次世代シーケンサー）によるウイルスゲノムシーケンスでわかること

➤ ウイルスの系統学的分類：

→ 特定ウイルスの検出、新規ウイルスの同定

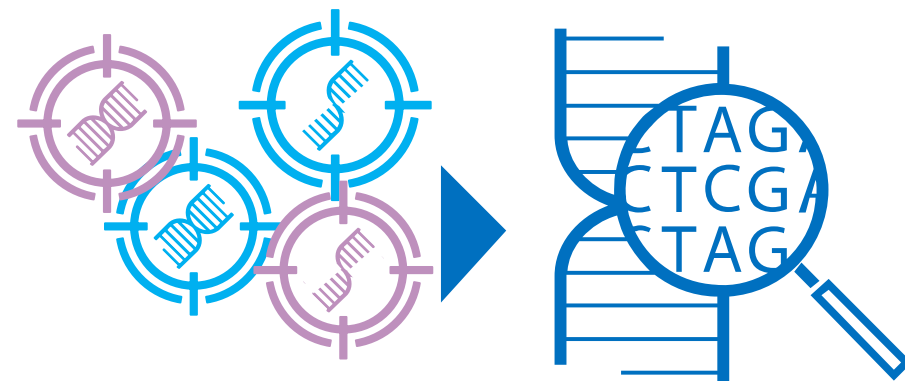
➤ ウイルスを構成するタンパクの情報：

→ ウイルス性状、病原性に関わる情報、治療薬のターゲット探索

➤ 変異情報：

→ 変異株の分類、薬剤耐性株の検出、
時系列に沿った系統樹解析により拡散過程など防疫に関わる情報

→ **NGSによるウイルスゲノムシーケンスにより、臨床的に重要な知見を得ることができる**



本日のAgenda

1. シーケンス法の種類
2. ライブラリ調製とシーケンス
3. BaseSpace Sequence Hub (BSSH) を使った情報解析
4. 本日のまとめ

ウイルスゲノムシーケンス法の種類

・通常得られるサンプルには、複数種のウイルスゲノムや、宿主のDNA/RNAが含まれている

◆ Shotgun metagenomics

・微生物群のゲノムを包括的にシーケンスすることで、複数の生物種の同定、検出を行う

◆ ターゲットシーケンス

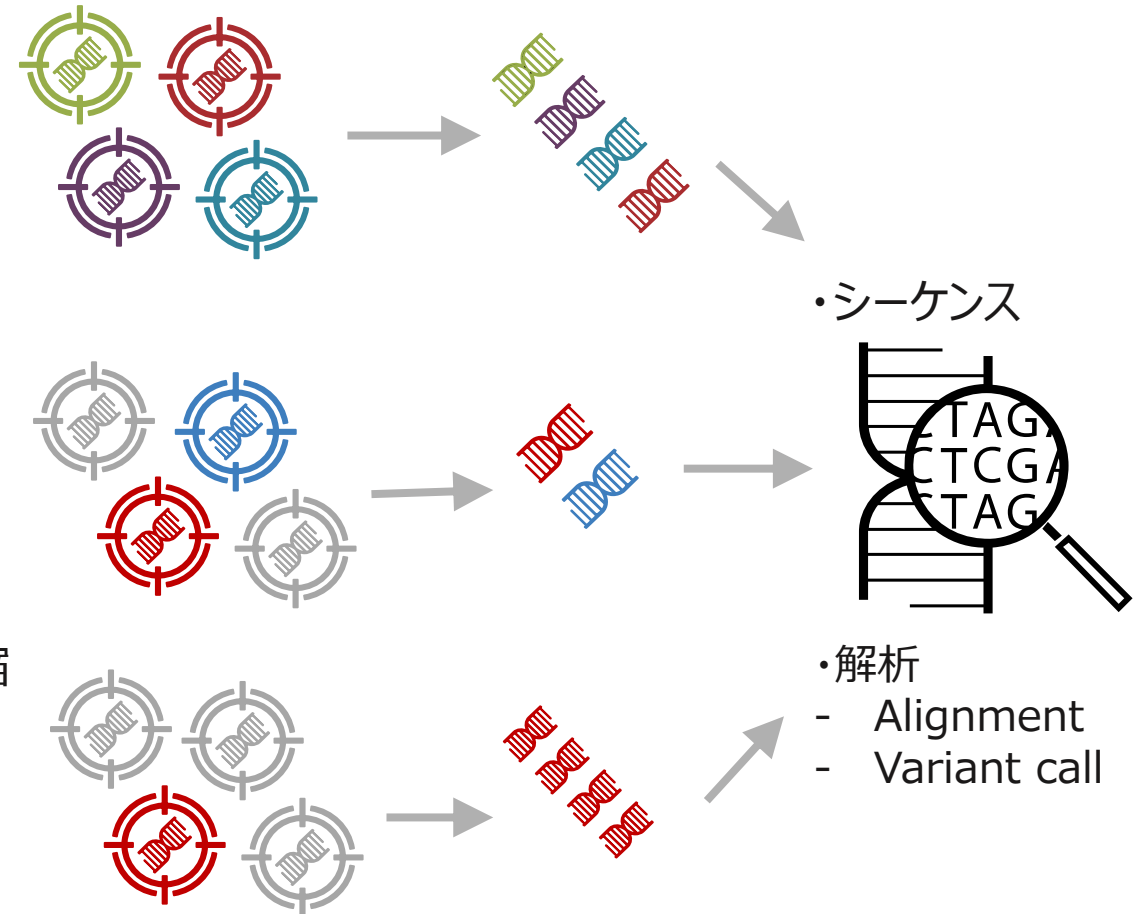
・サンプル中の特定の領域を取り出してシーケンスを行う

➤ Enrichment

・特定の領域に相補的なプローブを用いて目的の領域のみを濃縮

➤ Amplicon

・PCR法でターゲット領域を増幅する



各シーケンス法の特徴

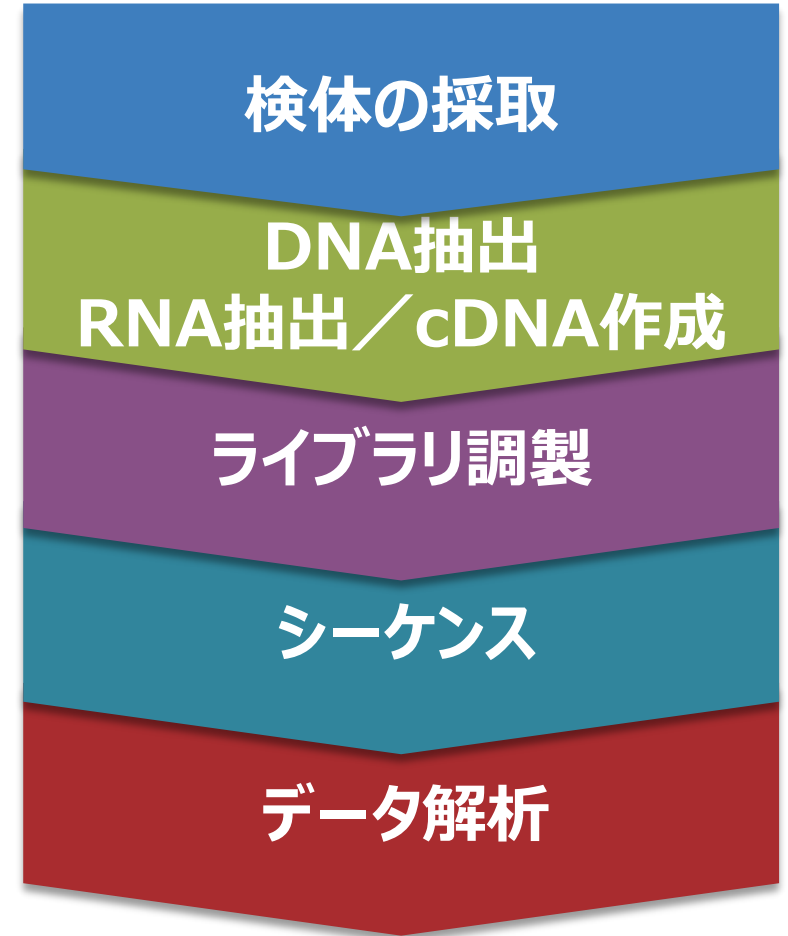
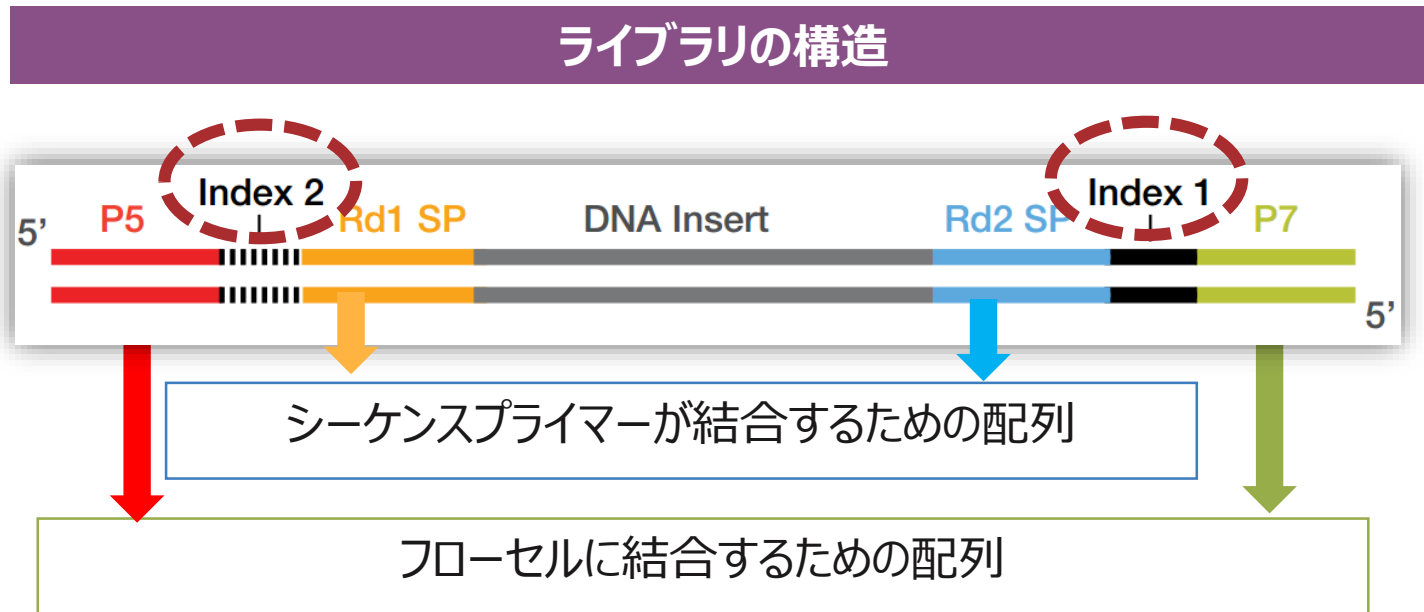


	ウイルス培養法	Shotgun Metagenomics	Enrichment	Amplicon
サンプル・ライブラリ調製法	Culture	None	Hybridization	PCR
新規種の検出	No	✓	No	No
複数の種の検出	No	✓	Possible	Possible
感度	High	Variable	Moderate	Moderate
1 サンプルあたりのリード数	0.5 M	10 M	0.1-1 M	0.5 M
結果が得られるまでの時間	++++	+	++	++
バイオインフォマティクス解析	Easy	Advanced	Easy	Easy
イルミナ社Application Note	✓	✓	✓	-












→ サンプルの状態、解析目的によってシーケンス法を選択し、それぞれに合わせたサンプルの前処理・ライブラリ調製を行う

シーケンス解析の流れ

- ・ライブラリ調製：
シーケンサーにかける前に、サンプルDNAの両端に
シーケンスに必要な配列のついた構造に加工すること



Illuminaの提供するWorkflow (SARS-CoV-2 For Research)

Workflow	Library Prep	Sequencer Fit	Analysis	Data Sharing
Shotgun Metagenomics (RNA)	 <p>Requires RNA extraction followed by TruSeq stranded total RNA</p>	 <p>NextSeq, NovaSeq</p>	 <ul style="list-style-type: none"> • DRAGEN Metagenomics 	
Enrichment	 <p>Requires cDNA followed by Nextera DNA Flex for Enrichment + Respiratory Virus Oligo Panel</p>	 <p>iSeq, MiniSeq, MiSeq</p>	 <ul style="list-style-type: none"> • DRAGEN Enrichment • DRAGEN RNA Pathogen Detection 	 <p>GISAID Submission</p>
Amplicon	 <p>AmpliSeq for Illumina SARS-CoV-2 Research Panel</p>	 <p>iSeq, MiniSeq, MiSeq</p>	 <ul style="list-style-type: none"> • DNA Amplicon 	

2. ライブラリ調製とシーケンス



TruSeq Stranded Total RNAによる Shotgun Metagenomics workflow例

- 229E 200ng
- 229E 10ng (+190ng UHR)
- OC43 200ng
- OC43 10ng (+190ng UHR)

- ・市販のコロナウイルス2株（OC43、229E）で TruSeq Stranded Total RNAによるライブラリ調製を実施
- ・各ウイルス株単独200ng、あるいは各ウイルス株10ngに ヒトレファレンスRNA（UHR）190ngを混合したサンプルからスタート
- ・MiSeq: 2x76 bpシーケンス

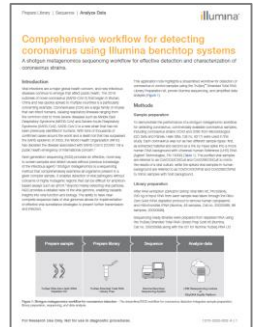
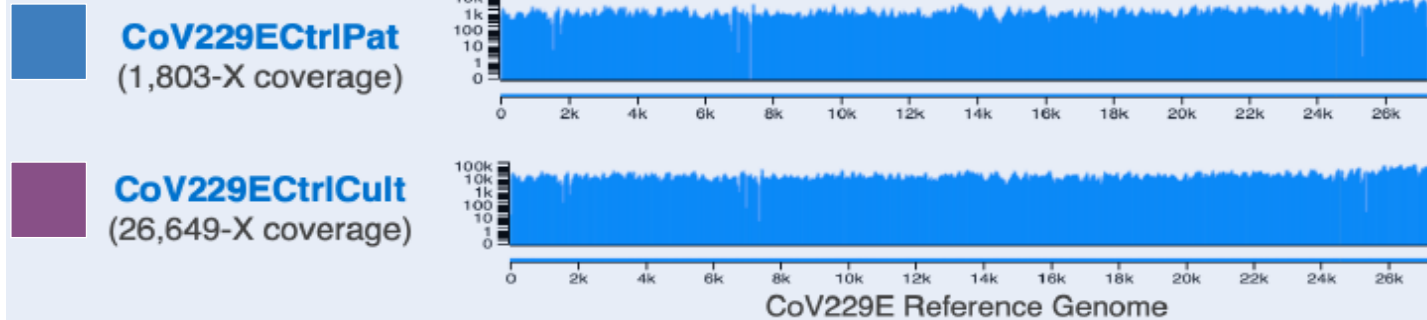
Table 2: Basic metrics using LRM Resequencing Module

Sample	Aligned reads	Mean coverage	Small variants
■ CoVOC43CtrlPat	5.5%	1786.8	47
■ CoV229EctrlPat	6.1%	2266.5	24
■ CoVOC43CtrlCult	78.8%	25,971.5	47
■ CoV229EctrlCult	71.5%	33,833.3	24

Alignment-Based Coverage Maps, Viral Genome Assembly

5% CoV
+95%UHR

100%
CoV



Comprehensive workflow for detecting coronavirus using Illumina benchtop systems

<https://jp.illumina.com/content/dam/illumina-marketing/documents/products/appnotes/ngs-coronavirus-app-note-1270-2020-001.pdf>

(1) Shotgun Metagenomics workflow

- TruSeq Stranded Total RNAのワークフロー
- TruSeq Stranded Total RNAのプロトコルと注意点
- シーケンス条件
- コスト試算



(1) Shotgun Metagenomics workflow

- TruSeq Stranded Total RNAのワークフロー
- TruSeq Stranded Total RNAのプロトコルと注意点
- シーケンス条件
- コスト試算



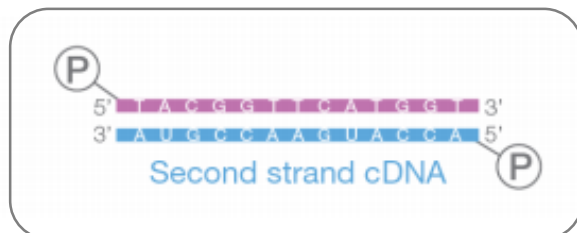
TruSeq Stranded Total RNAワークフロー



① リボソームRNAの除去とRNAの断片化



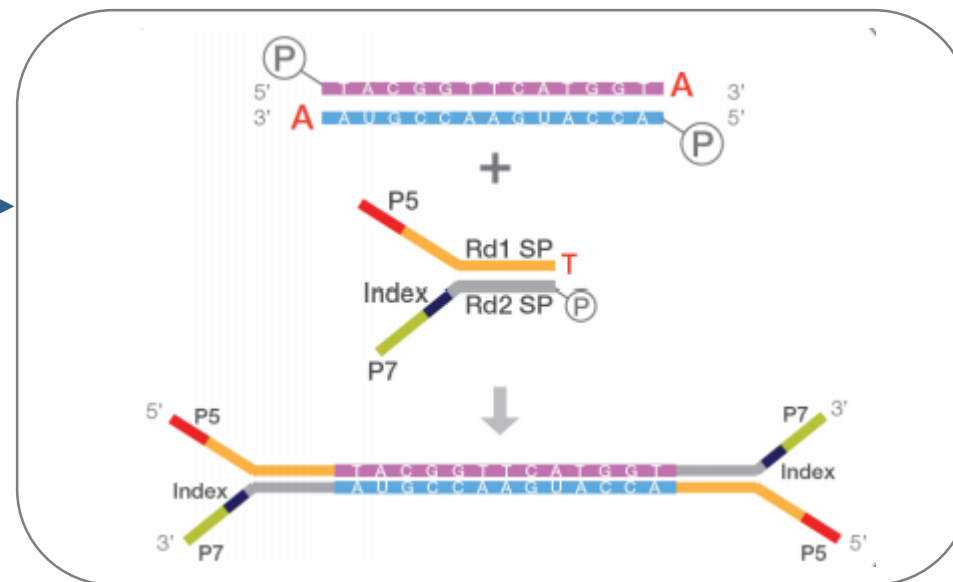
② ランダムプライマーを用いてcDNA合成



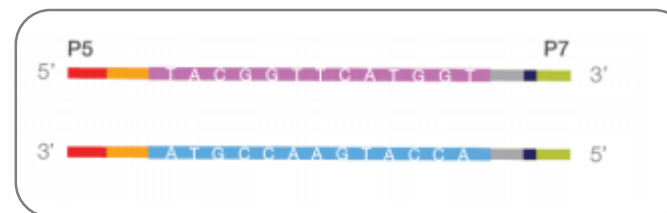
③ 2nd Strandの合成とリン酸化



④ A-Tailの付加



⑤ インデックス付アダプターのライゲーション



⑥ PCR増幅

⑦ ライブラリのクリーンアップ、定量、ノーマライゼーション

(1) Shotgun Metagenomics workflow

- TruSeq Stranded Total RNAのワークフロー
- TruSeq Stranded Total RNAのプロトコルと注意点
- シーケンス条件
- コスト試算



TruSeq Stranded Total RNAの Protokol

・TruSeq Stranded Total RNAの Protokol 詳細については、
下記日本語ウェビナー資料、Reference Guideをご参照ください

【日本語ウェビナー】

RNA-Seqをはじめよう ライブラリー調製編：
絶対に失敗しないライブラリー調製【イルミナiSchool 初級】

<https://jp.illumina.com/events/webinar/2018/webinar-180228-j.html>

- 必要準備品、実験消耗品、機器情報 (p8~)
- 実験の Protokol (p22~)
- トラブルシュート(p37)

【Reference Guide】

TruSeq Stranded Total RNA Reference Guide

<https://jp.support.illumina.com/downloads/truseq-stranded-total-rna-reference-guide-1000000040499.html>

RNA-Seqをはじめようライブラリー調製編
絶対に失敗しないライブラリー調製

フィールドアプリケーションサイエンティスト
仲 健太
2018/2/28

AMPure XP / RNAClean XP

AMPure XPにはポリエチレングリコール (PEG)、NaCl、2価陽イオン、磁性ビーズ (SPRIビーズ) 等が含まれています。

<核酸精製>
ポリエチレングリコール沈殿と同様な原理でDNAを精製。

Agencourt AMPure XP Binding Separation Ethanol Wash Elution Buffer Transfer

AMPure XP / RNAClean XP

mRNA Total RNA
H/M/R Gold
Globin Plant

TruSeq Stranded mRNA / Total RNA sample Prep kit Reference Guide
AMPure XPワークフロー

1. DNA溶液にReference Guide記載の量のAMPure XPを加え、ピペッティング。静置、15分。
2. マグネットスタンドに置き、上清除去。
3. 80% エタノールを加え、30秒、静置し、上清を除去。
4. ステップ3を再度実施。
5. 風乾、15分
6. プロトコル記載のResuspension Bufferを加え、マグネットスタンドから外し、2分、静置。
7. マグネットスタンドに置き、5分 (上清が透明になるまで) 静置。
8. プロトコル記載の量の上清を回収。

AMPure XPは使用する30分以上前に室温に出しておき、使用する直前にはボルテックスでよく撪拌してください。AMPure XPは正確な量を加えて下さい

80% EtOHは用事調製

過度な風乾は溶出効率が低下します。Reference Guideには風乾15分と記載していますが、多検体処理を前提としております。ビーズ表面にヒビが入った時点で過度となりますのでご注意ください。最大5分

illumina

TruSeq Stranded Total RNA : RNA抽出に関する注意点

- 抽出したRNAは吸光度測定し、精製度を確認してください
 - 260/280 Ration Value : ~2.0、260/230 : 2.0-2.2
- Agilent 2100 Bioanalyzer、Tape Station、もしくは同等品を使用し、RIN値 (RNA Integrity Number) でRNAの分解度を確認してください
- Input RNA量は、**10-100 ng/ μ LのTotal RNA 10 μ L、(0.1-1 μ g Total RNA)** です。定量法の指定はありませんが、蛍光色素を用いた方法をご利用ください
- RNA抽出時には、DNase処理を実施してください。コンタミしたDNAはライブラリ化されますので、解析の精度が低下します

Corona Virus Protocol

- ウイルスサンプルは各施設のバイオセーフティ基準に従って取り扱ってください
- イルミナから推奨する特定のRNA抽出キットはありませんが、Corona Virusアプリケーションノートでは下記キットでRNA抽出を行っています

製品名	製造販売元	型番
QIAmp Viral RNA Mini Kit (50 samples)	QIAGEN	52904
QIAmp Viral RNA Mini Kit (250 samples)	QIAGEN	52906

キットに含まれるリボソームRNAの除去試薬（Ribo-Zero）の種類

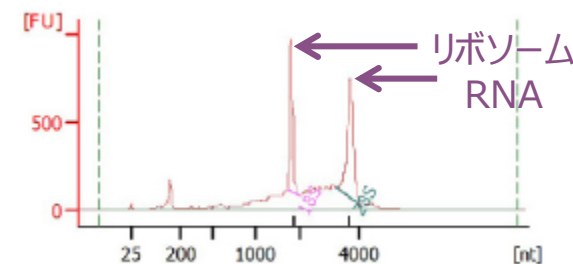
・Ribo-Zero試薬により、抽出RNAに多く含まれる目的外のrRNAなどを除去することで、不要なrRNAにアライメントされるリードが大きく減少します。

表) TruSeq Stranded Total RNAキットに含まれるRibo-Zero試薬の除去ターゲットrRNA

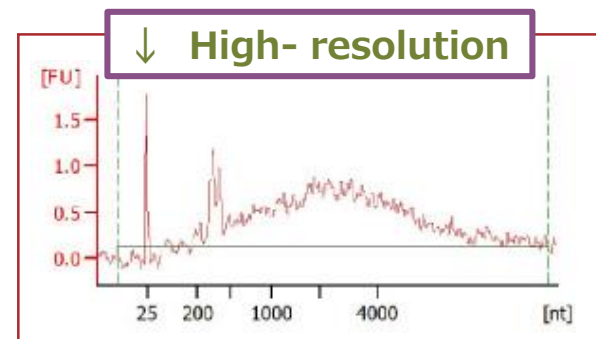
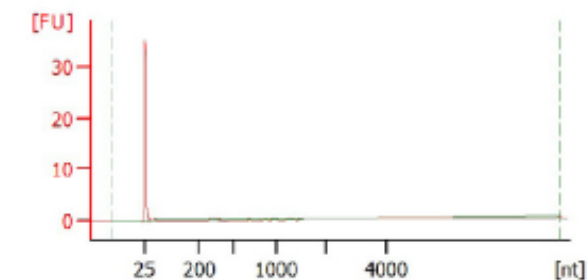
TruSeq Stranded Total RNA kits	ヒト		マウス・ラット			植物			
	細胞質	ミトコンドリア	グロビン*	細胞質	ミトコンドリア	グロビン*	細胞質	ミトコンドリア	葉緑体
Library prep (Human/Mouse/Rat)	○			○					
Library prep Gold	○	○		○	○				
Library prep Globin	○	○	○	○	○	○			
Library Prep Plant							○	○	○

*グロビンmRNAが除去ターゲット

Universal Human Reference Total RNA



▼ Ribo-Zero rRNA Removal Kit (Human/Mouse/Rat).



(1) Shotgun Metagenomics workflow

- TruSeq Stranded Total RNAのワークフロー
- TruSeq Stranded Total RNAのプロトコルと注意点
- シーケンス条件
- コスト試算



TruSeq Stranded Total RNA : シーケンス条件

- RNA-Seqライブラリはインサート長が150-220 bp
- 各シーケンサーのインプットライブラリ量は、各Support Bulletinではおよそ500 bpのライブラリを想定している
 - 短いライブラリは、クラスターを形成しやすいため、初めてRNA-Seqをする場合、推奨濃度よりも少し控えめ（x0.7-x0.8）をお勧めします

Platform	インプット濃度	至適クラスター密度	RNA-Seqライブラリインプット濃度
HiSeq 2000/2500 High OutPut v3	12.0 pM	750 - 850K clusters/mm2	8.5 - 9.5 pM
HiSeq2500 High OutPut v4	18.0 pM	950 - 1050K clusters/mm2	12.5 - 14.5 pM
HiSeq 2500 Rapid Run v2	12.0 pM	850 - 1000K clusters/mm2	8.5 - 9.5 pM
MiniSeq	1.8 pM	170 - 220K clusters/mm2	1.2 - 1.5 pM
MiSeq v2 Reagents	12.5 pM	1000 - 1200K clusters/mm2	8.8 - 10.0 pM
MiSeq v3 Reagents	15.0 pM	1200 - 1400K clusters/mm2	10.0 - 12.0 pM
NextSeq 500/550	1.8 pM	170 - 220K clusters/mm2	1.2 - 1.5 pM

【日本語ウェビナー】 RNA-Seqをはじめよう ライブラリー調製編：絶対に失敗しないライブラリー調製【イルミナiSchool 初級】
<https://jp.illumina.com/events/webinar/2018/webinar-180228-j.html>

(1) Shotgun Metagenomics workflow

- TruSeq Stranded Total RNAのワークフロー
- TruSeq Stranded Total RNAのプロトコルと注意点
- シーケンス条件
- コスト試算



シーケンスコスト試算：TruSeq Stranded Total RNAによる Shotgun Metagenomics workflow

●ライブラリ調製コスト：96 sampleを調製する場合

* サンプル採取およびRNA抽出のコストは含みません。

キット名	1キットあたりのサンプル数	カタログ番号	キットのコスト (円)	1サンプルあたりのコスト (円)
TruSeq Stranded Total RNA Library prep Gold	96 samples	20020599	¥1,640,000	¥17,083
IDT for Illumina TruSeq RNA UD Indexes	96 indexes, 96 samples	20022371	¥107,600	¥1,121
合計			¥18,204	

●シーケンスコスト：2x75bp、1サンプルあたり10Mリードの場合



シーケンサー	NextSeq550		NextSeq 2000	NovaSeq			
	MidOutput	HighOutput	P2	SP	S1	S2	S4
キット (サイクル数)	150 cycle	150 cycle	200 cycle	200 cycle	200 cycle	200 cycle	200 cycle
カタログ番号	20024904	20024907	20040557	20040326	20012864	20012861	20027466
キットのリード数 (CPF**)	130M	400M	400M	800M	1.6G	4.1G	10G
1ラン当たりサンプル数	13	40	40	80	160	410	1000
キットのコスト (円)	¥188,900	¥496,500	¥352,400	¥489,000	¥880,900	¥2,082,100	¥4,404,400
1サンプルあたりのコスト (円)	¥14,531	¥12,413	¥8,810	¥6,113	¥5,506	¥5,078	¥4,404

* 2020年6月現在の希望納入価格 (税抜) に基づきます。

** Cluster Pass Filterを通過したリード

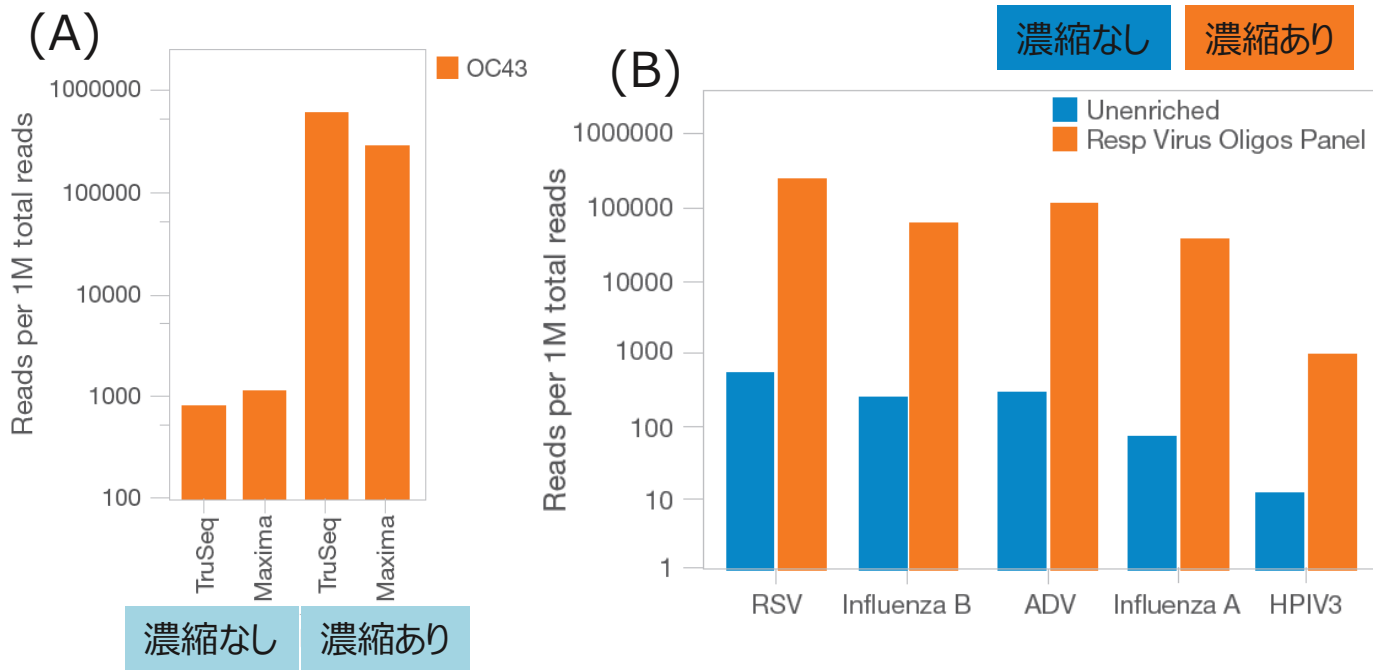
2-2. Enrichment



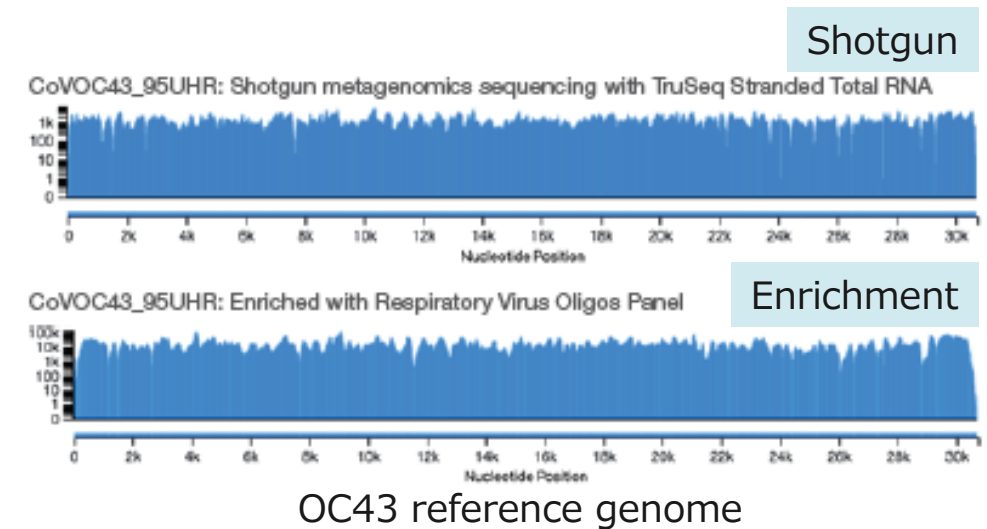
Nextera DNA Flex for Enrichmentによる Enrichment workflow例



- (A) UHRに5%市販のコロナウイルス株 (OC43) RNAを混合したサンプル
- (B) 呼吸器疾患関連DNA/RNAウイルス5種のみ、あるいはUHRに5%5種ウイルスの核酸を混合したサンプル
- Nextera DNA Flex for Enrichment と、Respiratory Virus Oligo Panelでライブラリ調製を実施
- MiSeqで 2x151 bpのシーケンス



(A) Shotgun metagenomicsとの比較



Enrichment workflow for detecting coronavirus using Illumina NGS systems

<https://sapac.illumina.com/content/dam/illumina-marketing/documents/products/appnotes/ngs-enrichment-coronavirus-app-note-1270-2020-002.pdf>

For Research Use Only. Not for use in diagnostic procedures.



(2) Enrichment workflow

- Nextera DNA Flex for Enrichmentとは
- Nextera DNA Flex for Enrichmentのワークフロー
- Nextera DNA Flex for Enrichmentのプロトコルと注意点
- シーケンス条件
- コスト試算

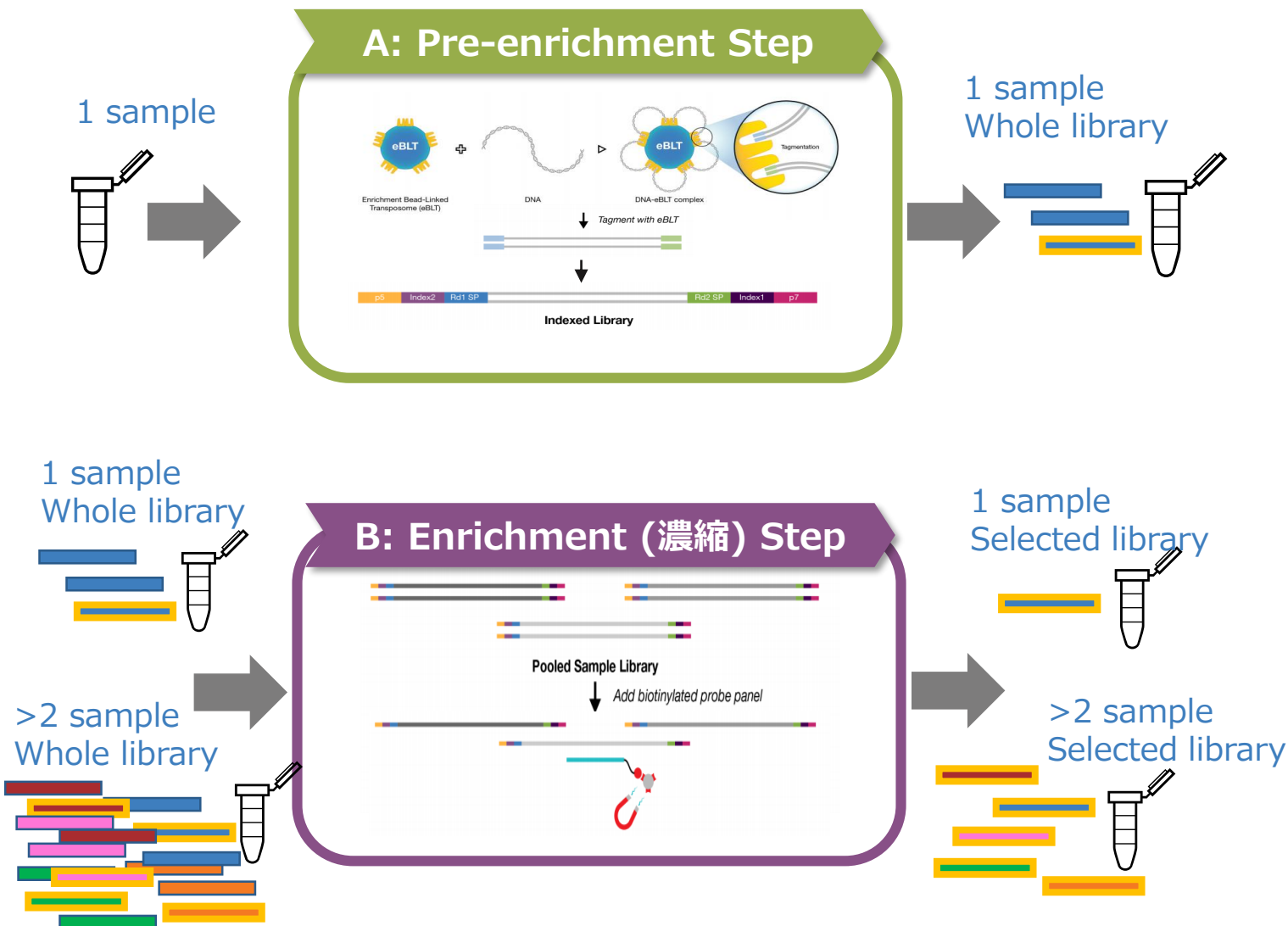


(2) Enrichment workflow

- Nextera DNA Flex for Enrichmentとは
- Nextera DNA Flex for Enrichmentのワークフロー
- Nextera DNA Flex for Enrichmentのプロトコルと注意点
- シーケンス条件
- コスト試算



Nextera DNA Flex for Enrichmentの特徴



- Pre-enrichmentとEnrichment（濃縮）の2段階のワークフロー
- Enrichment（濃縮）ステップは、2サンプル以上同時に処理可能（プレ・プール方式）
- 最短6.5hour*1に短縮されたプロトコル
- 10-1000 ngのインプットゲノムDNAに対応
- FFPE（ホルマリン固定パラフィン標本）由来DNA,末梢血、唾液サンプルからのプロトコルも用意*2
- 各研究分野に最適な、ヒト用濃縮用オリゴプローブパネルを展開
- サードパーティ製濃縮プローブパネルにも対応

*1 12サンプル、12Plexの場合
 *2 サンプルによってプロトコル、オプション必要試薬が異なりますのでご注意ください。

Nextera DNA Flex for Enrichment : 製品-1

【注意事項】

- Enrichment(濃縮ステップ)を、何サンプルプールしてから行うか（12サンプルプールして1回で濃縮するなら12 plex）によって、必要になるコンポーネントキット数が異なりますのでご注意ください
- Nextera DNA Flex for Enrichment には、①～④の**4つのコンポーネントが必要**です

①ライブラリ調製試薬 : サンプル数分

②インデックス試薬 : サンプル数分 あるいは 一度にランするサンプル数分

③濃縮試薬 : Enrichmentを行う回数分

④オリゴプローブパネル : Enrichmentを行う回数分

Pre-enrichment Stepで使用

Enrichment (濃縮) Stepで使用

例

全部で96個のサンプルを
12サンプルずつプールして濃縮
→濃縮は8回分

- 96サンプル
- 12-plex
- 8 Enrichment

①と②は96サンプル分
③と④は8 Enrichment分必要

下記のページで必要になる各キット数の確認ができます

Nextera DNA Flex for Enrichment選択ツール

<https://jp.illumina.com/destination/s/library-prep-checker/selector-enrichment.html>

注意) Nextera DNA Flex for Enrichmentは、1-plexまたは12-plexでの濃縮反応に最適化されています。その他のプレックス数の濃縮も可能ですが、プロトコルの最適化が必要になります

Nextera DNA Flex for Enrichment : 製品-2

① ライブラリ調製試薬 および ③ 濃縮試薬（下記の中からご選択ください）

カタログ番号	製品名	①ライブラリ調製試薬 (サンプル数)	③濃縮試薬 (Enrichment反応量)
20025523	旧) Nextera DNA Flex Pre-Enrichment Library Prep and Enrichment Reagents – 16 samples	16 sample	16 Enrichment
	新) Illumina DNA Prep with Enrichment (S) Tagmentation 16 Samples		
20025524	旧) Nextera DNA Flex Pre-Enrichment Library Prep and Enrichment Reagents – 96 samples	96 sample	8 Enrichment
	新) Illumina DNA Prep with Enrichment (S) Tagmentation 96 Samples		
20025519	旧) Nextera DNA Flex Pre-Enrichment Library Prep Reagents (16 samples)	16 sample	(なし)
	新) Illumina DNA Prep (S) Tagmentation 16 Samples		
20025520	旧) Nextera DNA Flex Pre-Enrichment Library Prep Reagents (96 samples)	96 sample	(なし)
	新) Illumina DNA Prep (S) Tagmentation 96 Samples		

注意) 2020年5月、製品名の変更がありました。
製品選択をより分かりやすくするための変更で、品番、製品内容に変更はございません。

Nextera DNA Flex for Enrichment : 製品-3

②インデックス製品 (下記の中からお選びください)

【旧】

カタログ番号	製品名	②インデックス製品
20027213	IDT® for Illumina Nextera DNA Unique Dual Indexes Set A	96 Indexes, 96 Samples
20027214	IDT® for Illumina Nextera DNA Unique Dual Indexes Set B	96 Indexes, 96 Samples
20027215	IDT® for Illumina Nextera DNA Unique Dual Indexes Set C	96 Indexes, 96 Samples
20027216	IDT® for Illumina Nextera DNA Unique Dual Indexes Set D	96 Indexes, 96 Samples



【新】

カタログ番号	製品名	②インデックス製品
20027213	IDT® for Illumina® - DNA/RNA UD Indexes Set A, Tagmentation	96 Indexes, 96 Samples
20027214	IDT® for Illumina® - DNA/RNA UD Indexes Set B, Tagmentation	96 Indexes, 96 Samples
20042666	IDT® for Illumina® - DNA/RNA UD Indexes Set C, Tagmentation	96 Indexes, 96 Samples
20742667	IDT® for Illumina® - DNA/RNA UD Indexes Set D, Tagmentation	96 Indexes, 96 Samples

注意) 2020年5月、製品名、品番、内容の変更がありました。製品使用方法に変更はございません。

Nextera DNA Flex for Enrichment : 製品-4

④ 濃縮用オリゴプローブパネル製品 (下記の中からご選択ください)

カタログ番号	濃縮用オリゴプローブパネル製品名	プローブのターゲット	④濃縮用オリゴ Enrichment反応量
20020183	Illumina Exome Panel (CEX)	ヒトコーディング領域	8 enrichments
FC-121-0202	TruSight Cancer	がん関連94遺伝子	8 enrichments
20029227	TruSight One	疾患関連4,813遺伝子	6 enrichments
20029226	TruSight One Expanded	遺伝性がん素因関連113遺伝子	6 enrichments
20029551	TruSight Hereditary Cancer	疾患関連6,704遺伝子	8 enrichments
20029229	TruSight Cardio	心疾患関連174遺伝子	8 enrichments
20029550	TruSeq Neurodegeneration	神経変性疾患関118連遺伝子	8 enrichments
20042472	Respiratory Virus Oligos Panel NEW	呼吸器疾患関連ウイルス株41種	8 enrichments
20025371	Illumina® Custom Enrichment Panel via Design Studio	(カスタム)	8 enrichments

• サードパーティ製濃縮プローブパネルをご使用の場合は、以下の基準に沿っているかご確認ください

- プローブの長さ : 80 あるいは 120 bp
- プローブ数 : 500-675,000 probes
- SingleあるいはDouble stranded
- プローブ量は3 pmol以上 (1 – 12 plex/enrichmentあたり)

• RNAプローブをご使用の場合は、“Nextera Flex for Enrichment with RNA Probes”をご参照ください

https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_nextera/nextera-flex-enrichment/nextera-flex-for-enrichment-rna-probes-demonstrated-protocol-1000000070581-01.pdf

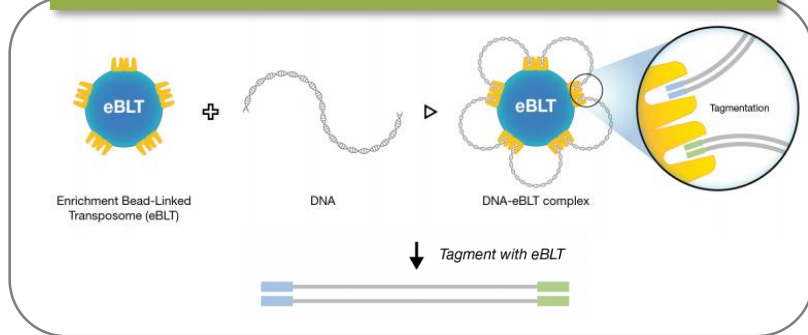
(2) Enrichment workflow

- Nextera DNA Flex for Enrichmentとは
- Nextera DNA Flex for Enrichmentのワークフロー
- Nextera DNA Flex for Enrichmentのプロトコルと注意点
- シーケンス条件
- コスト試算

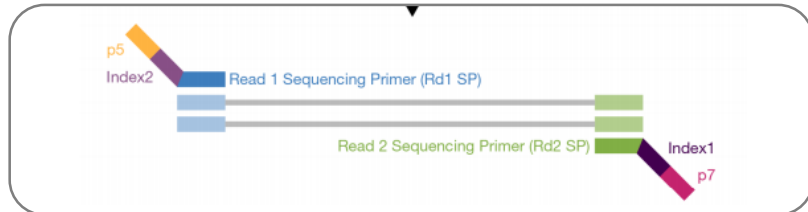


Nextera DNA Flex for Enrichmentワークフロー

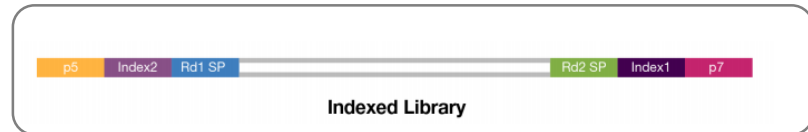
Pre-enrichment library prep



① cDNAの断片化とタグメンテーション



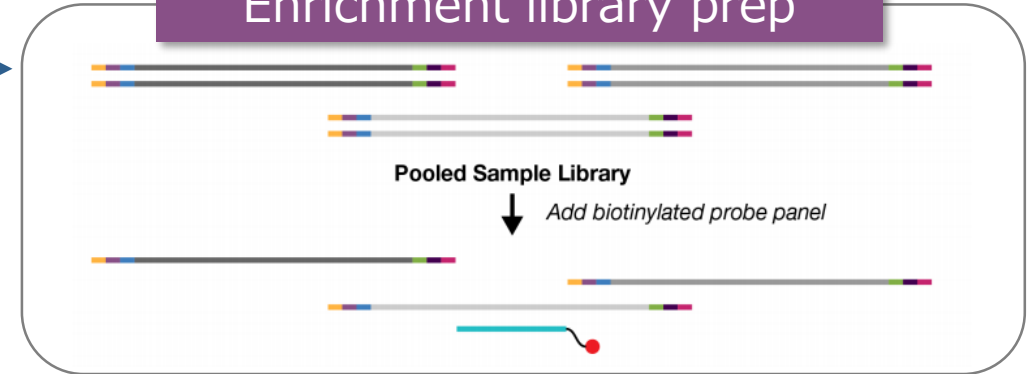
② PCRによるインデックスアダプターの付加



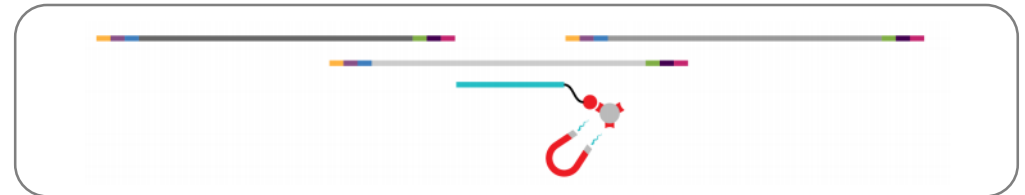
③ ライブラリの溶出、クリーンアップ

④ ライブラリの定性、定量とプーリング

Enrichment library prep



⑤ パネルプローブのハイブリダイゼーション



⑥ ビーズによるプローブ結合ライブラリの濃縮と溶出

Enriched and Indexed Library Ready for Sequencing

⑦ PCRによる選択されたライブラリプールの増幅

⑧ プールしたライブラリの定性、定量

(2) Enrichment workflow

- Nextera DNA Flex for Enrichmentとは
- Nextera DNA Flex for Enrichmentのワークフロー
- Nextera DNA Flex for Enrichmentの**プロトコルと注意点**
- シーケンス条件
- コスト試算

*このセクションではNextera DNA Flex for Enrichment の標準プロトコルをご紹介しますが、Respiratory Virus Oligos Panelと組み合わせてコロナウイルス株をサンプルとした社内実験の際の情報を、Corona Virus protocolとして添え書きしています。

Corona Virus Protocol

★Corona Virus protocol



Nextera DNA Flex for Enrichment : 準備品(抜粋)

製品名	製造販売元	型番	容量
Agencourt AMPure XP kit	Beckman Coulter	A63880	5 ml
		A63881	60 ml
Qubit dsDNA BR Assay Kit	Thermo Fisher	Q32850	100 assay
		Q32853	500 assay
Agilent DNA 1000 Kitなど	Agilent	5067-1504	25 Chip

← 精製ステップで使用

← ライブラリQCで使用

製品名	製造販売元	型番
Agilent 2100 Bioanalyzer (もしくは同等品)	Agilent	G2940CA
Qubit Fluorometer 3.0 (もしくは同等品)	Thermo Fisher	Q32866
C100 Touch Thermal cycler with 96-Deep well Reaction Module	BioRad	1851197
Magnetic Stand-96*	Thermo Fisher	AM10027
(*あるいは MagneSphere Technology Magnetic Separation Stands)	Promega	Z5342
96ウェル深底プレート (MIDIプレート)	Thermo Fisher	AB-0859
96ウェルPCRプレート	Eppendorf/Bio-Rad	0030129512/HSP-9601
96 well プレートインキュベーター	SciGene	1057-30-O
MIDI Head Blocker for SciGene Hybex system	Illumina	BD-60-601
プレートシェーカー	BioShake iQ/XP Hi-Speed Thermal Mixer	1808-0506/1808-505

その他、一般的な実験用の消耗品が必要になります。詳しくは、Reference Guideをご参照下さい。

公式プロトコル : Nextera Flex for Enrichment (=Illumina DNA Prep with Enrichment) Reference Guide

<https://jp.support.illumina.com/downloads/illumina-dna-prep-with-enrichment-reference-guide-1000000048041.html>

Nextera DNA Flex for Enrichment : Input DNAに関する注意点 (genomic DNAサンプルの場合)

- Genomic DNAの場合：推奨input量は **50-1000 ng**です
(30 μ l inputの場合、1.67 ~ 33.3 ng/ μ l)
10-49 ngの場合は、蛍光色素を用いた方法で定量を行ってください
- イルミナから推奨する特定のDNA抽出キットはありません
30 μ l input DNAからスタートですので、それ以下の容量に溶出してください
- 抽出、作成したDNAは吸光度測定し、精製度を確認しましょう
 - 260/280 Ration Value : 1.8-2.0、260/230 : 2.0-2.2
- FFPE由来のDNAの場合は、必ずQCを行って、 Δ Cqが5以下であることを確認してください
 - Infinium HD FFPE QC Assay Protocol (15020981)
- ヒト末梢血、唾液サンプルに関しては、個別のプロトコル、キットをご用意しています。Reference Guideをご参照ください

【重要】

Input DNAに以下のものが含まれないこと

- 1mM以上のEDTA
- 有機化合物 (フェノール、エタノールなど)
→ 10mM Tris-HCl(pH7.5-8.5)に Suspendすることをお勧めします

【重要】 ウェビナー資料はプロトコル概要のご紹介ですので、
詳細は必ず公式Reference Guide (英語版) をご参照ください。

公式プロトコル : Nextera Flex for Enrichment (=Illumina DNA Prep with Enrichment) Reference Guide
<https://jp.support.illumina.com/downloads/illumina-dna-prep-with-enrichment-reference-guide-1000000048041.html>

Nextera DNA Flex for Enrichment Corona virus検出時のInput DNAとプロトコル

Corona Virus
Protocol
(Respiratory Virus
Oligos Panel)

- cDNA作成へのRNA持ち込み量は、最少で10 ngを推奨します
- 推奨する特定のRNA抽出・cDNA合成キットはありませんが、社内テストでは下記キットを使用しています

製品名	製造販売元	型番
QIAmp Viral RNA Mini Kit (50 samples)	QIAGEN	52904
QIAmp Viral RNA Mini Kit (250 samples)	QIAGEN	52906
QIAGEN AllPrep PowerViral DNA/RNA Kit (50 samples)	QIAGEN	28000-50
Maxima H Minus Double-Stranded cDNA Synthesis Kit (10 reactions)	Thermo Scientific	K2561

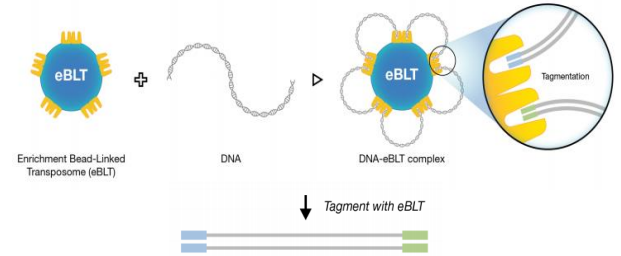
- cDNA合成後のRNA degradationは必要ありません。
もしご使用のキットにRNA degradationステップがありましたら、incubationは1秒以内にとどめることを推奨します
- もっともよい結果を得るためには、抽出後すぐのRNAを使用されることをお勧めします
- 推奨インプット量は**50-100 ng cDNA**です。
(最低でも 10 ng) サンプルは30 µl以下に溶出してください
- 通常のgenome DNAインプットの場合と異なるプロトコルになっていますので、ご注意ください

Table 4: Recommended Nextera Flex for Enrichment workflow parameters

Step	Parameter
Amplify tagmented DNA	12 cycles
Hybridize probes	ramp down to 58°C
Capture hybridized probes	58°C
Amplify enriched library	12 cycles at 58°C

Detecting coronavirus with the Respiratory Virus Oligo Panel <https://www.illumina.com/content/dam/illumina-marketing/documents/products/appnotes/coronavirus-enrichment-product-list-1270-2020-004.pdf>

Nextera DNA Flex for Enrichment プロトコル① : cDNAの断片化とタグメンテーション



- (1) サンプル (50-1000ng) を30 μ LずつPCR plateに用意する
- (2) tagmentation master mix★を20 μ L添加してよくピペッティング
- (3) プログラム名 : TAG program
55°C, 5分
10°C, hold
- (4) 室温に2分置く
- (5) ST2 10 μ Lを添加して反応を止める → 室温5分
- (6) マグネットスタンドにセットし、TWB bufferで3回Washする

TWBは泡立ちやすいので
ピペッティングは慎重に

→OverDryを防ぐために、最後のWashで上清を除かずに次のステップへ

★(2)で使用するtagmentation master mix

Reagent	1 rxn
eBLT	11.5 μ L
TB1	11.5 μ L
Total Volume	23 μ L

重要 : eBLTは完全にSuspendすること

実験のポイント : eBLT

- ビーズが乾かないように容器を立てて保管する (2-8°C)
- 凍結、加熱しないこと
- 乾燥を防ぐために、サンプル数が多い場合は、wash/clean upを数サンプルずつ分けて行う
- 推奨のマグネットを使用すること

Nextera DNA Flex for Enrichment

② PCRによるインデックスアダプターの付加

- (1) PCR master Mixを調製する
- (2) ①のTWB Bufferでのwashを完了させる
(透明になった上清を除く) →マグネットスタンドから外す
- (3) すぐにPCR master Mix 40 μ lを加える
→ すぐにピペティングしてビーズを完全にsuspendする
- (4) Index Adapters 10 μ lを加えてピペティング

(5)プログラム名 : eBLT program

72°C, 3分

98°C, 3分

98°C, 20秒

60°C, 30秒

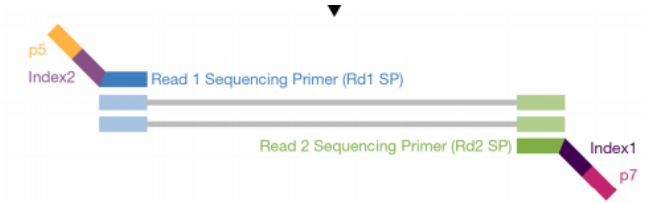
72°C, 1分

72°C, 3分

10°C, hold

何サイクル行うかは
Input DNAによって異なります
(右表参照)

★ビーズを乾燥させないこと



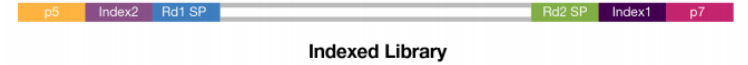
★PCR master Mix

Reagent	1 rxn
EPM	23 μ l
Nuclease-free water	23 μ l
Total Volume	46 μ l

Sample Input Type	Number of PCR Cycles (X)
10–49 ng genomic DNA	12
50–1000 ng genomic DNA	9
50–1000 ng extracted FFPE	12
Saliva	9
Blood	9

★Corona Virus protocol
12 cycleを推奨します

Nextera DNA Flex for Enrichment



③ ライブラリの溶出、クリーンアップ

(1) マグネットスタンドにセットして45 μ l の上清を回収

(2) **AMPure XP Beads**を使用したクリーンアップ

* High quality gDNA, blood and saliva samplesの場合

- 77 μ l のnuclease free waterと88 μ l のAMPure XP Beadsを添加
- マグネットに静置して上清200 μ l を回収
- 準備しておいた20 μ l のAMPure XP Beadsと混ぜる

* FFPE由来DNAサンプルの場合

- 81 μ l のAMPure XP Beadsを添加

(3) 室温に5分置く

(4) マグネットスタンドにセットし、上清を捨てる

(5) 80%エタノールで2回Wash→Air Dryでエタノールを除く

(6) 17 μ l のRSBを添加、ボルテックス、室温2分インキュベーション
→ マグネットにセットして15 μ l の上清を回収

実験のポイント：クリーンアップ

- 2-Step クリーンアップなので間違えないこと
- 1st Clean up- Large fragmentを除く
→ **上清を回収**
- 2nd Clean up-Small fragmentを除く
→ **上清を捨てる**
→最後にRSBにElute

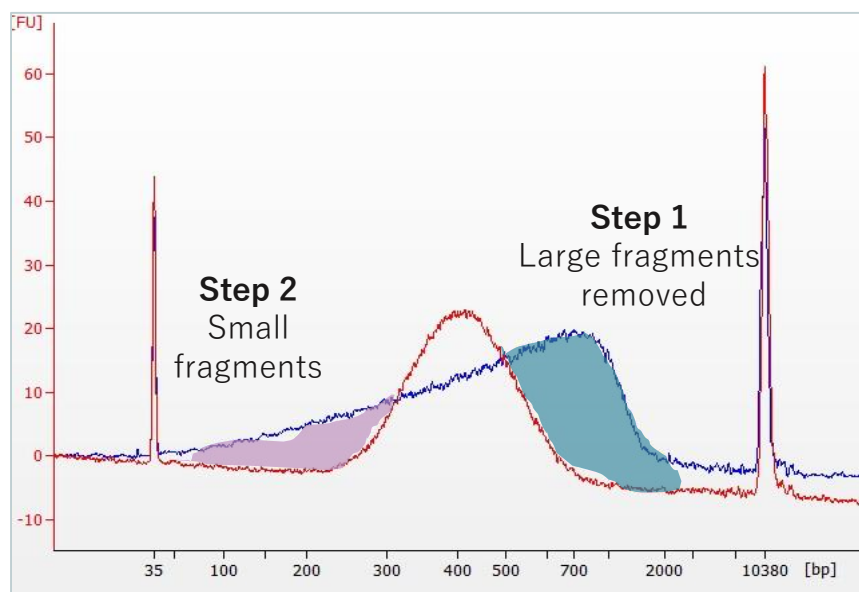
- ★重要：FFPE由来DNAサンプルの場合は、1-Stepでのクリーンアップ（2nd Clean upのみ）を行う
- * Small fragmentを除く
→新しくビーズを加えて上清を捨てる
→最後にRSBにElute

★Corona Virus protocol
もしcDNA量が少ない場合は、FFPE用の
81 μ lのプロトコル（1-Step）でClean upしてください

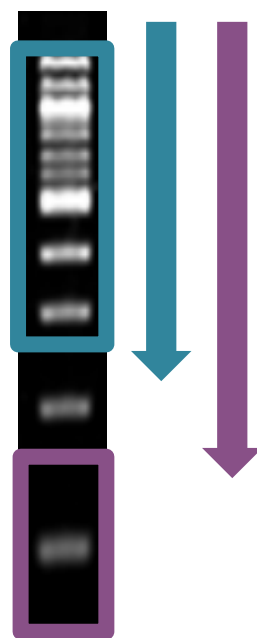
参考 : AMPure XP BeadsによるClean up

核酸溶液とAMPure XP Beadsの体積比によって
ビーズ上に結合して回収される核酸のサイズが変わりますので、
ビーズに対するDNA溶液の体積比が非常に重要です。

* AMPure XP Beadsに
結合するライブラリサイズ



* こちらはクリーンアップ結果を分かりやすくするための図で、
Nextera DNA Flex for Enrichmentのライブラリではありません



実験のポイント : AMPure XP Beads

- 実験前に室温に戻しておくこと
- Vortexをして完全にSuspendしておくこと
- 乾燥しすぎないように、Reference Guideの乾燥時間は参考程度に
- 正確なピペティングが重要

<AMPure XP Beadsの取り扱い情報を含む日本語資料>

AMPure XP / RNAClean XP
AMPure XPにはポリエチレングリコール (PEG)、NaCl、2価陽イオン、磁性ビーズ (SPRIビーズ) 等が含まれています。

<核酸精製>
ポリエチレングリコール沈殿と同様な原理でDNAを精製。

1. 核酸を含む反応溶液にAMPure XPを加える。
2. ポリエチレングリコールによってDNAが沈殿し、ビーズに結合。マグネットスタンドに置き、上清除去。
3. EtOHでウォッシュ。
4. Elution Buffer (Resuspension Buffer, milliQなど) を加え、溶出。
5. マグネットスタンドに置き、上清回収。

illumina

Step4-ii ライブラリーの精製 ~AMPure ビーズでの2段階精製

<準備>
- AMPureXPビーズを室温に戻す (30分以上)
- 70% EtOHを作成。1サンプル当たり300µlを使用 (用事調製)

<精製の原理>
AMPureXPビーズ溶液と核酸溶液の体積比によって、ビーズに結合する核酸のサイズが変わる

illumina

【日本語ウェビナー】RNA-Seqをはじめよう ライブラリー調製編 :
絶対に失敗しないライブラリー調製【イルミナiSchool 初級】
<https://jp.illumina.com/events/webinar/2018/webinar-180228-j.html>
【日本語ウェビナー】イルミナで AmpliSeq for Illumina パネル解析をはじめよう
-製品紹介とライブラリー調製編-【イルミナiSchool 初級】
<https://jp.illumina.com/events/webinar/2018/webinar-180425-j.html>

Nextera DNA Flex for Enrichment

④ ライブラリの定性、定量とプーリング

- (1) 15 μ l のサンプル中、1 μ lをBioanalyzerでの定性に使用
- (2) 体積比によるプール方法か、質量比によるプール方法のどちらかで、**Total 30 μ lのライブラリプール**を作成

Table 2 Recommended Pooling Methods

Sample Input	Pooling Method
10-49 ng gDNA	Mass or volume*
50-1000 ng gDNA	Volume
50-1000 ng extracted FFPE	Mass or volume*
Saliva	Volume
Blood	Volume

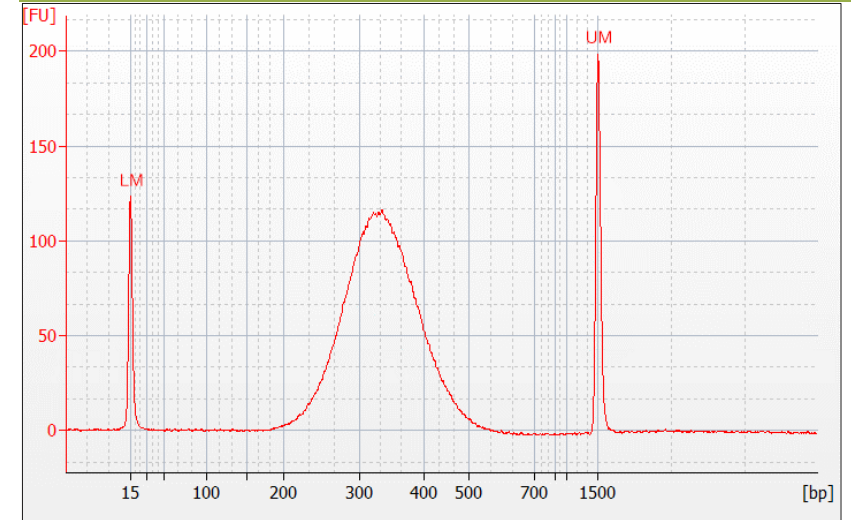
★定量は、1 μ lをQubit dsDNA BR Array Kitで定量してください。

<Input DNAと想定収量>

- 10-49 ng genomic DNA : ≥ 100 ng
- 50-1000 ng genomic DNA : ≥ 250 ng
- blood*, saliva* : ≥ 250 ng

*イルミナ既定のキット、プロトコルで抽出した場合

Nextera DNA Flex enrichment ライブラリ : Pre-enrichment終了時のトレース例



★想定ライブラリサイズはおよそ300-400bp
FFPEサンプルの場合は少し小さくなり、250bp

★Corona Virus protocol

- 12-plex、定量して質量比でのプールを強く推奨します。
- 500 ng/libraryが理想ですが、200 ngしかないライブラリがある場合では、等質量比 (200ng/library) でのプールをお勧めします

Nextera DNA Flex for Enrichment

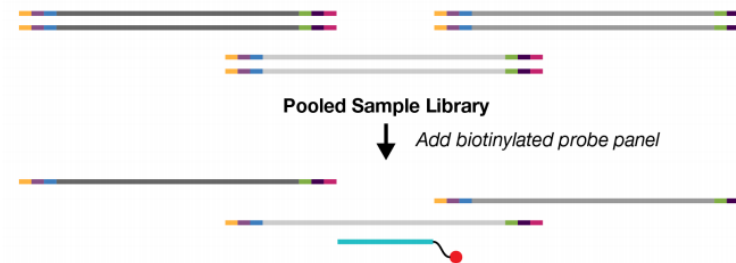
⑤ パネルプローブのハイブリダイゼーション

(1) プールサンプル30 μ l に、NHB2、プローブパネル、EHB2を、この順番に添加してピペティング

Reagent	1 tube
Pooled/1-plex library	30 μ l
NHB2	50 μ l
Oligo probe panel	10 μ l
EHB2	10 μ l
Total Volume	100 μ l

実験のポイント：NHB2

- -20°C保存においては沈殿が発生
→完全な溶解が必須
- 室温で解凍後、50°C5分加熱
- Vortexとピペティングを繰り返して、沈殿を完全に溶かすこと



(2) Program名：NF-HYB

95°C, 5分
 94°C, 1分
 92°C, 1分
 ...
 64°C, 1分
 62°C, hold(90分)

1分に
2°Cずつ
下げる
16 cycle

* FFPEサンプル、CEXパネル、Somatic Variant callingの場合

95°C, 5分
 94°C, 1分
 92°C, 1分
 ...
 60°C, 1分
 58°C, hold(90分)

1分に
2°Cずつ
下げる
18 cycle

★Corona Virus protocol
18 cycle, 58°Cを推奨します

Nextera DNA Flex for Enrichment

⑥ ビーズによるプローブ結合ライブラリの濃縮と溶出

(1) ハイブリダイズ後のサンプル100 μ l にSMB 250 μ lを添加し混合

(2) MIDI Heat Blockerに移し、62°C(★) 15分インキュベーション

→ この間にEEWを62°C(★)に温めておく

(* Washが終わるまでEEWの温度キープ)

(3) マグネットスタンドに移す。上清が透明になったら上清を捨てる

(4) マグネットスタンドから外し、温めておいたEEW 200 μ lを添加

→ 62°C(★)で5分インキュベーション

→マグネットスタンドにセット、上清が透明になったら上清を捨てる

(5) (4)を合計4回繰り返す

★4回目のWashでは

新しいチューブ/MIDI plateを使用

(6) 完全に上清を取り除いた後、23 μ l Elution Mix (*) を添加

(7) 室温2分インキュベーション,マグネットスタンドにセット、21 μ lの上清を回収

(8) 4 μ l のET2を添加、静かにピペッティング

★FFPEサンプル、CEXパネル、Somatic Variant callingの場合は 全て58°C

* (6)で使用するElution Mix

Reagent	1 rxn
EE1	28.5 μ l
HP3	1.5 μ l
Total Volume	30 μ l

★Corona Virus protocol
58°Cを
選択してください

実験のポイント：濃縮ステップ

- 温度コントロールが非常に重要
- SMBビーズは完全にSuspendする
- SMBは乾かし過ぎない
- EEWでのWashの際、ピペッティングはしない (ShakerかVortexで懸濁)

Nextera DNA Flex for Enrichment

⑦ PCRによる選択されたライブラリの増幅

(1) 25 μ l のenrichedライブラリに5 μ l PPC、20 μ l EPMをこの順に加える

(2) Program名 : AMP

98°C, 30秒

98°C, 10秒

60°C, 30秒

72°C, 30秒

72°C, 5分

10°C, hold

12 cycle (★)

★CEXパネルの場合は10 cycleを推奨。
サードパーティ製オリゴプローブを使用の場合は、
Reference Guideをご確認ください

★Corona Virus protocol
58°C 12 cycleを推奨します

★最終ライブラリQCは、1 μ lをもちいて
Qubit dsDNA BR Assay kitでの定量、
1 μ lを用いてBioanalyzerのHigh sensitivity
DNA kitでの定性を推奨します

【ライブラリのクリーンアップ】

(1) 50 μ lのenrichedライブラリに45 μ lのAMPure XP Beadsを添加して混和

(2) マグネットスタンドにセットして95 μ lの上清を捨てる

(3) 80%エタノールでのWash2回 → Air Dryでエタノールを除く

(4) 32 μ l RSBを加えて5分室温インキュベーション

→ マグネットにセットして30 μ lの上清を回収

Nextera DNA Flex enrichment ライブラリ :
Enrichment終了時のトレース例



★想定ライブラリサイズはおよそ350bp
(サイズ範囲は200-1000bp)

(2) Enrichment workflow

- Nextera DNA Flex for Enrichmentとは
- Nextera DNA Flex for Enrichmentのワークフロー
- Nextera DNA Flex for Enrichmentのプロトコルと注意点
- シーケンス条件
- コスト試算



Nextera DNA Flex for Enrichment : シーケンス条件

- ライブラリの調製後、下記のスタート濃度から、各装置の推奨希釈プロトコルを開始してください。
- 最終ライブラリ濃度の最適化が必要な場合があります。

Platform	ライブラリ希釈スタート濃度	インプット時最終ライブラリ濃度
HiSeq 2000/2500 (high output modes)	2 nM	16 – 18 pM
HiSeq 2500 System (rapid run mode)	2 nM	7 – 8 pM
HiSeq 3000/4000	2-3 nM	150 – 200 pM
iSeq 100	2 nM	100 pM
MiniSeq	2 nM	1.7 – 1.8 pM
MiSeq (v3 reagents)	4 nM	10 – 12 pM
NextSeq 500/550	2 nM	1.4 – 1.5 pM
NovaSeq 6000 System (standard workflow)	2 nM	175 – 185 pM

- Reference Guide (公式プロトコル) では、2x 101 bpを想定していますが、長く読むことも可能です。

★Corona Virus protocol

•2x 76 bpを推奨

•一般的には、1サンプルあたりのリード数は0.5 M-1 Mを推奨
(初めてのシーケンスの場合は 1 M/Sample)

(2) Enrichment workflow

- Nextera DNA Flex for Enrichmentとは
- Nextera DNA Flex for Enrichmentのワークフロー
- Nextera DNA Flex for Enrichmentのプロトコルと注意点
- シーケンス条件
- コスト試算



シーケンスコスト試算：Nextera DNA Flex for EnrichmentとRespiratory Virus Oligo PanelによるEnrichment workflow

●ライブラリ調製コスト：96 sampleを調製する場合

* サンプル採取およびRNA抽出、cDNA作成のコストは含みません。

キット名	1キットあたりのサンプル数	カタログ番号	キットのコスト (円)	1サンプルあたりのコスト (円)
Illumina DNA Prep with Enrichment (S) Tagmentation 96 Samples	96 samples (8 enrichment)	20025524	¥1,306,900	¥13,614
IDT® for Illumina® - DNA/RNA UD Indexes Set A, Tagmentation	96 Indexes, 96 Samples	20027213	¥107,600	¥1,121
Respiratory Virus Oligo Panel	8 enrichment	20042472	¥92,300	¥961
合計			¥15,696	

●シーケンスコスト： 2x76bp、 1サンプルあたり 0.5Mリードの場合



シーケンサー	iSeq100		MiniSeq		MiSeq			
	i1 reagent	MidOutput	HighOutput	Nano	Micro	v2	v3	
キット (サイクル数)	300 Cycle	300 cycle	150 cycle	300 cycle	300 cycle	300 cycle	150 cycle	
カタログ番号	20021533	FC-420-1004	FC-420-1003	MS-103-1001	MS-103-1002	MS-102-2002	MS-102-3001	
キットのリード数 (CPF**)	4M	8M	25M	1M	4M	15M	25M	
1ラン当たりサンプル数	8	16	50	2	8	30	50	
キットのコスト (円)	¥100,100	¥96,600	¥169,200	¥51,000	¥76,700	¥183,000	¥158,400	
1サンプルあたりのコスト (円)	¥12,513	¥6,038	¥3,384	¥25,500	¥9,588	¥6,100	¥3,168	

* 2020年6月現在の希望納入価格（税抜）に基づきます。

** Cluster Pass Filterを通過したリード

2-3. Amplicon

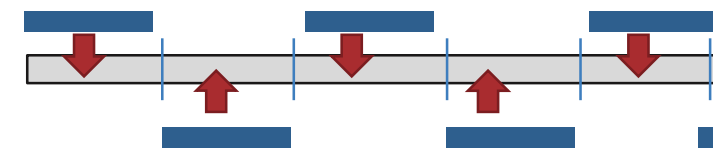
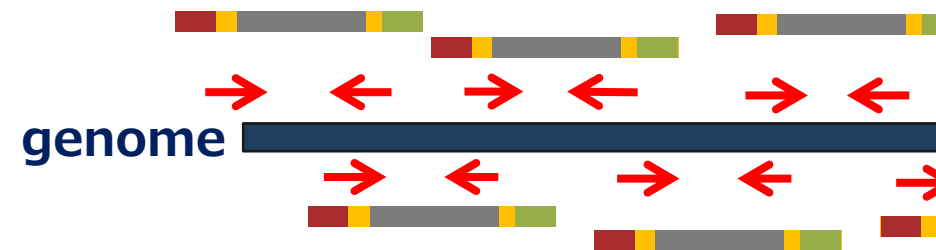


Amplicon workflowの種類について

- PCR Ampliconを複数設計することで、ゲノム上の複数の領域、あるいはゲノムの全長をライブラリ化することが可能

● SARS-CoV-2の場合

- ・複数のアンプリコンでゲノム全長をカバーするWGS tiling法が主に用いられている
- ・一般的に用いられている3rd Partyのtiling法
 - ARTIC法：98 PCR ampliconsのプライマーがV3までWEBで公開されている
<https://artic.network/>
 - ARTIC法をベースとした改変版プロトコル（日本国立感染研究所）
<https://www.protocols.io/view/ncov-2019-sequencing-protocol-for-illumina-betejeje>
- ・イルミナが提供しているSARS-CoV-2用プライマーセット
 - **AmpliSeq for Illumina SARS-CoV-2 Research Panel**



illumina®

(3) Amplicon workflow

- AmpliSeq for Illumina SARS-CoV-2 Research Panel
- AmpliSeq for Illuminaのワークフロー
- AmpliSeq for Illuminaのプロトコルと注意点
- シーケンス条件
- コスト試算



(3) Amplicon workflow

- AmpliSeq for Illumina SARS-CoV-2 Research Panel
- AmpliSeq for Illuminaのワークフロー
- AmpliSeq for Illuminaのプロトコルと注意点
- シーケンス条件
- コスト試算

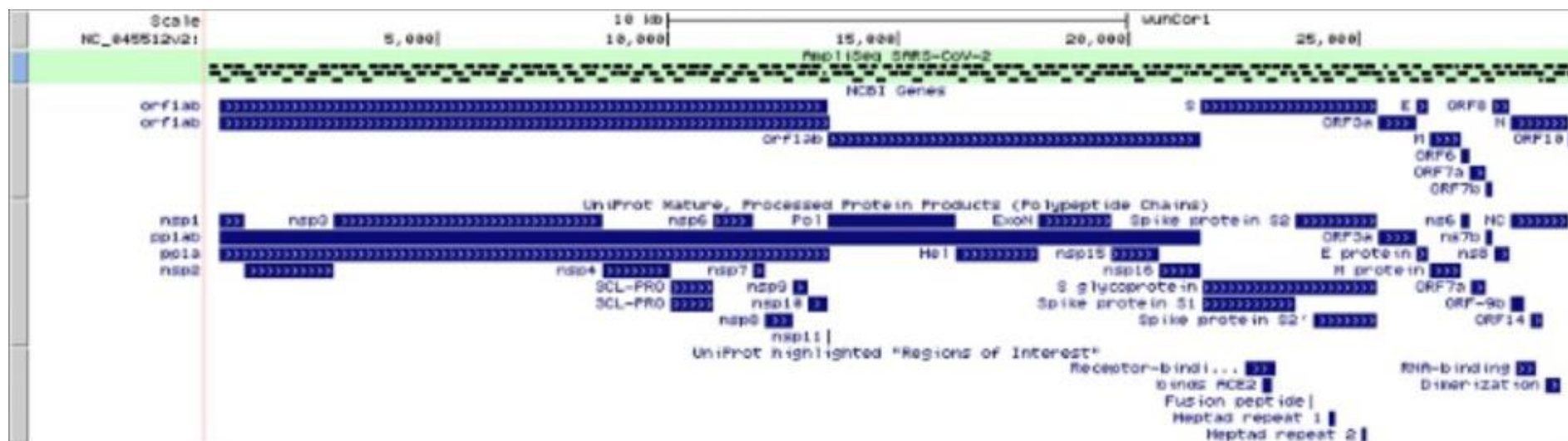


AmpliSeq for Illumina SARS-CoV-2 Research Panel

- 247 Amplicon (125-275 bp、2 pool) でSARS-CoV-2 ゲノムの全長をカバー (増幅コントロールとしてHuman 5 geneをターゲットとしたAmpliconも含まれる)
- 網羅的かつ高感度で低コストな変異解析を実現
- イルミナコンシェルジュサービスを使用することで、カスタムでアンプリコンを追加することが可能



AmpliSeq SARS-CoV-2 panelのAmpliconをUCSC Genome Browserで展開

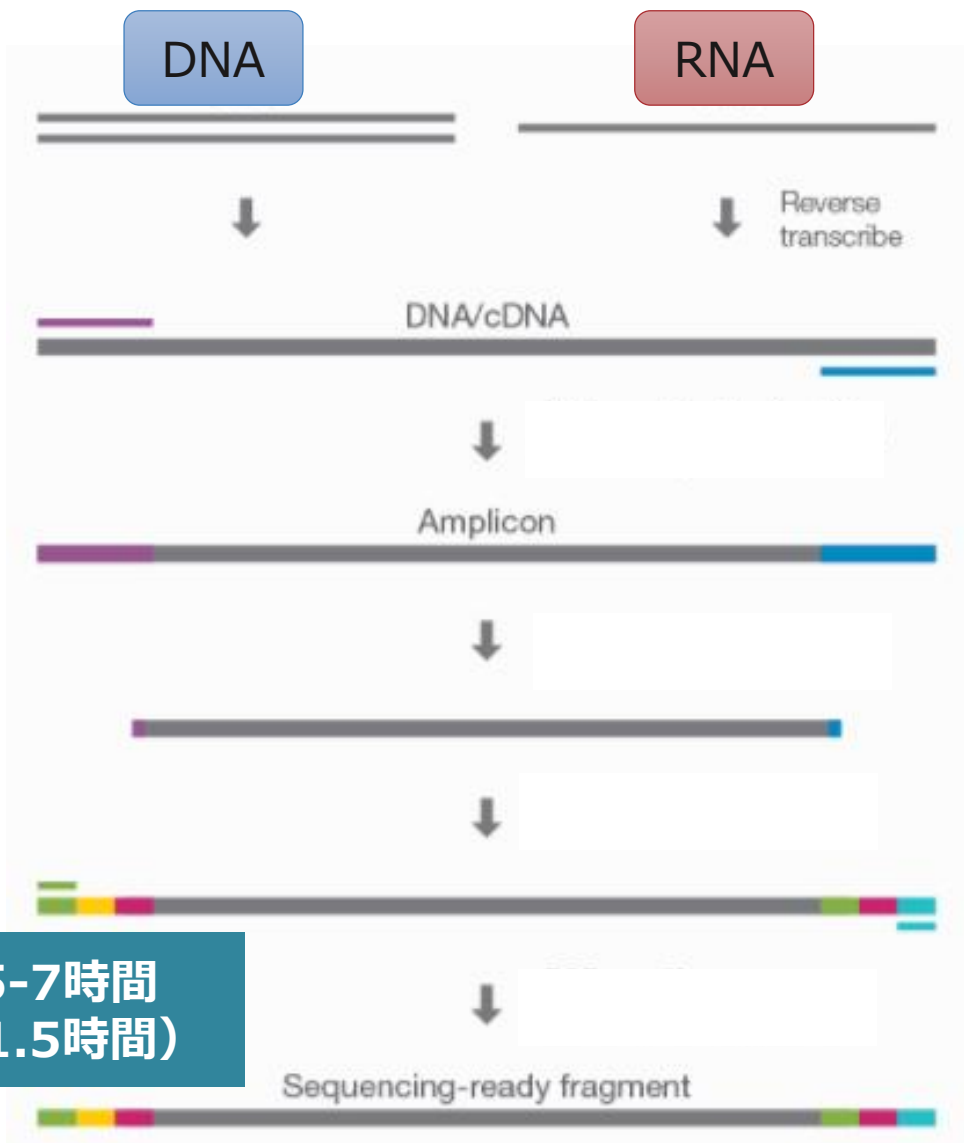


(3) Amplicon workflow

- AmpliSeq for Illumina SARS-CoV-2 Research Panel
- AmpliSeq for Illuminaのワークフロー
- AmpliSeq for Illuminaのプロトコルと注意点
- シーケンス条件
- コスト試算



AmpliSeq for Illumina (DNA/RNA)のワークフロー



所用時間5-7時間
(ハンズオン1.5時間)

① 逆転写cDNA合成

② ターゲット領域の増幅 (PCR)

③ プライマー部分を分解 (FuPa処理)

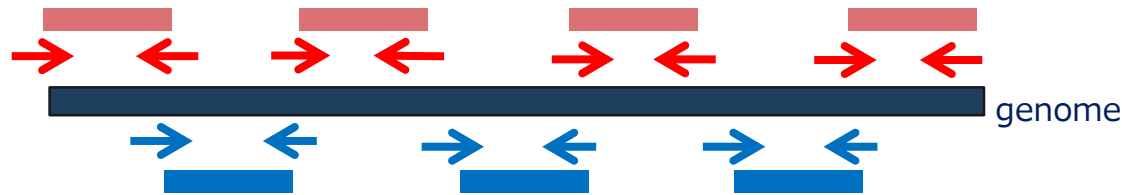
④ インデックス-アダプターのライゲーション およびライブラリーの精製 (AMPureビーズ精製)

⑤ ライブラリーの増幅 (PCR) および精製 (AMPureビーズ精製×2)

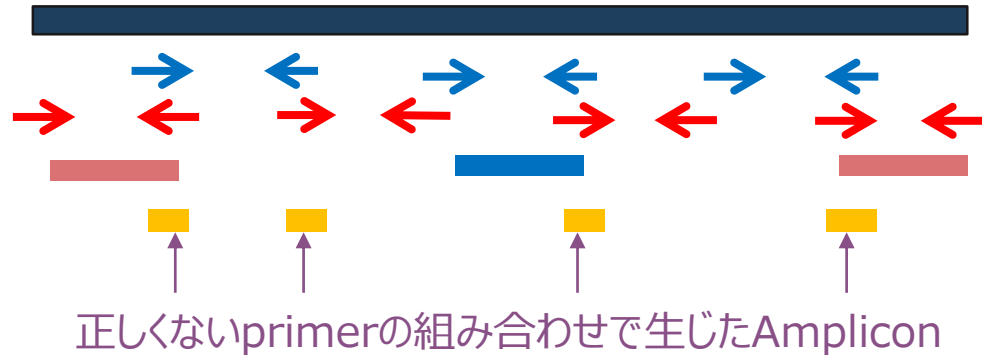
AmpliSeq for Illumina の注意点: プール数と反応数

- Amplicon数が多い場合、最初のPCR反応を分けて行う必要があります
- プール数 = 1サンプル当たりのPCR反応チューブ数
- 反応数 = Library PLUS 試薬の消費単位

PCR反応を2つのチューブに分けて行った場合



PCR反応を1つのチューブで行った場合



Primer set A
Amplicon

Primer set B
Amplicons

① ターゲット領域の増幅
→ 2 pool

② プライマー部分を分解 以降
→ 1チューブにまとめて1反応

【注意点】


- 1プールは最大3,072アンプリコンまで
- **Library PLUS の1反応分の試薬は2プールまで対応**
- パネルが4 プールの場合、Library PLUS は2反応分必要
- 1 プールの時も、Library PLUSは1反応必要

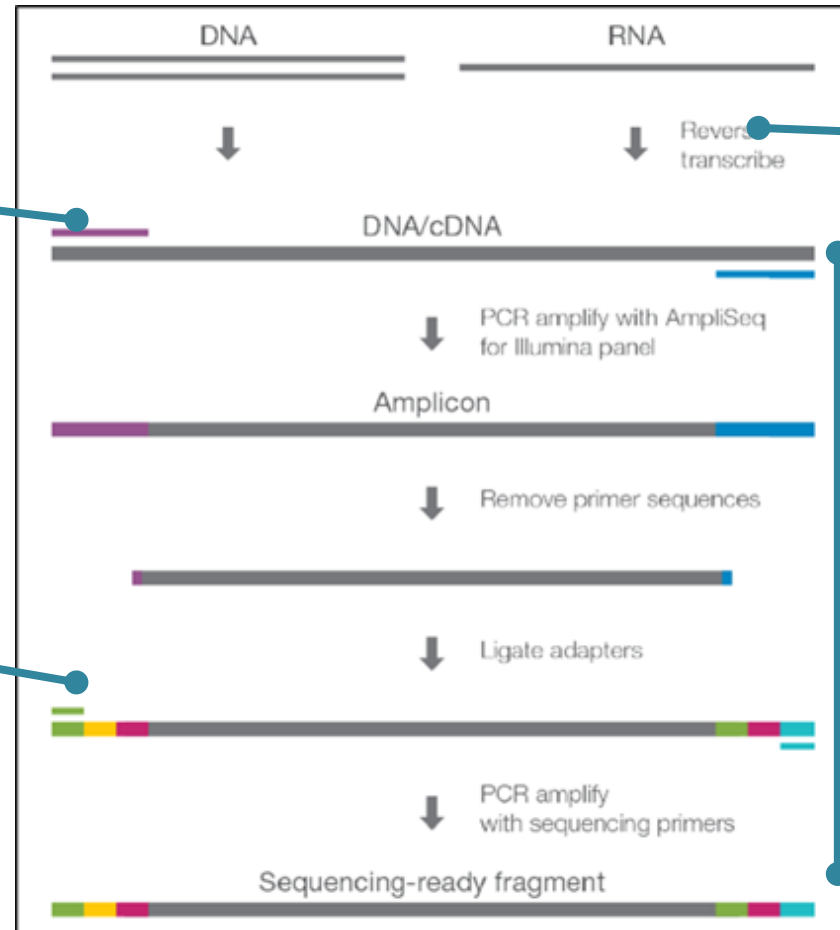
AmpliSeq SARS-CoV-2 Panel実験に必要な製品

Corona Virus Protocol
(AmpliSeq for illumina SARS-CoV-2 Research Panel)

- ①～④の4つの製品を個別にご購入いただく形になります


AmpliSeq for illumina SARS-CoV-2 Research Panel
(2,250反応分、247Amplicon, 2 pool)


Index plate
(AmpliSeq CD or UD Indexes for Illumina)




AmpliSeq cDNA Synthesis
(cDNA変換試薬)
(100反応分)


AmpliSeq Library PLUS試薬
(24 反応、96 反応、384 反応)

(3) Amplicon workflow

- AmpliSeq for Illumina SARS-CoV-2 Research Panel
- AmpliSeq for Illuminaのワークフロー
- AmpliSeq for Illuminaのプロトコルと注意点
- シーケンス条件
- コスト試算



AmpliSeq for Illuminaの Protokol

AmpliSeqのライブラリ調製の詳細に関しては、下記の日本語ウェビナーおよびReference Guideをご参照ください

【日本語ウェビナー】

AmpliSeq for Illumina パネル解析をはじめよう -製品紹介とライブラリー調製編-【イルミナiSchool 初級】

<https://jp.illumina.com/events/webinar/2018/webinar-180425-j.html>

- **必要準備品 (p36~)**
- **Input DNA、RNAの条件 (p42~)**
- **Protokol (p46~)**
- **トラブルシュート (p71~)**

【Reference Guide】

AmpliSeq for Illumina Custom and Community Panels Reference Guide

<https://support.illumina.com/downloads/ampliseq-for-illumina-custom-and-community-panels-reference-guide-100000036408.html>

- **Chapter 3: Protokol for RNA panels (p24~)**

【注意】AmpliSeq for IlluminaのProtokolは、インプットがDNAかRNAか、またパネルによって異なります。必ずお使いになるパネルのReference Guideをご参照ください



AmpliSeq SARS-CoV-2 Panel プロトコル① Input RNAと逆転写cDNA合成

● Input RNAの条件

- 20～50 ng/μlの濃度で保管
- 標準使用量は**10 ng/pool** (high quality RNA 1-100 ng)
- 7 μl以下に溶出する必要があります
- Qubit RNA HS Assay kitなど、蛍光色素を用いた方法で定量 (NanoDropなど、UVスペクトルを用いた定量方法は**不可**)

実験のポイント : Input RNA

- DNase処理をしておく
 - RNA溶解バッファーは高濃度のEDTAなど、PCR反応を阻害する試薬が入っていないバッファーを選ぶ。
- LowTEを推奨

★Corona Virus protocol
4 ng/μlのRNA希釈溶液を 5 μl使用 (計 20 ng)

● 逆転写cDNA合成

(1) PCRチューブに反応液を準備 (★)

(2) Program名 : RT
42°C, 30 分
85°C, 5分
10°C, hold

★逆転写反応液

Reagent	1 tube
5x AmpliSeq cDNA Reaction Mix	2 μl
10x AmpliSeq RT Enzyme mix	1 μl
Total RNA(1-100 ng per pool)	<7 μl
Nuclease-free water	(to 10 μl)
Total Volume	10 μl

AmpliSeq SARS-CoV-2 Panel プロトコル①：ターゲット領域の増幅（PCR）

(1) 5xAmpliSeq HiFi MixとStep①で作成したcDNA、および Nuclease-free water を混合し、Master mix(★)を作成

(2) 2 wellに Master mixを 8 μlずつ分注する

(3) 片方のwellに5xAmpliSeq SARS-CoV-2 panel Pool 1 を2 μl, もう片方のwell に5xAmpliSeq SARS-CoV-2 panel Pool 2 を2 μl加える (total 10 μl/Well)

(4) Program名: AMP_RNA

99°C, 2 分

| 99°C, 15秒

| 60°C, 4* 分

10°C, hold



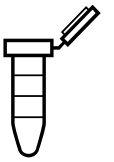
18*サイクル

*プール数、プライマーセット数、RNAのクオリティなどによって、該当サイクル数、温度が異なります

Primer set
Pool 1



Primer Set
Pool 2



★Master Mix

Reagent	per 1 sample
5xAmpliSeq HiFi Mix	4.5 μl
cDNA	10 μl
Nuclease-free water	3.5 μl
Total Volume	18 μl

Table 7 X Cycles and X Minutes

Primer Pairs per Pool	X Cycles for 10 ng High-Quality RNA per Pool	X Minutes
12-24	21	4
25-48	20	4
49-96	19	4
97-192	18	4
193-384	17	4
385-768	16	4
769-1200	15	8

AmpliSeq SARS-CoV-2 Panel プロトコル②：プライマー部分を分解（FuPa処理）

(1) Pool 1 (10 μ l)とPool 2 (10 μ l) を混合 (total 20 μ l)

(2) FuPa 試薬 2 μ lを加えて、混合 (total 22 μ l)

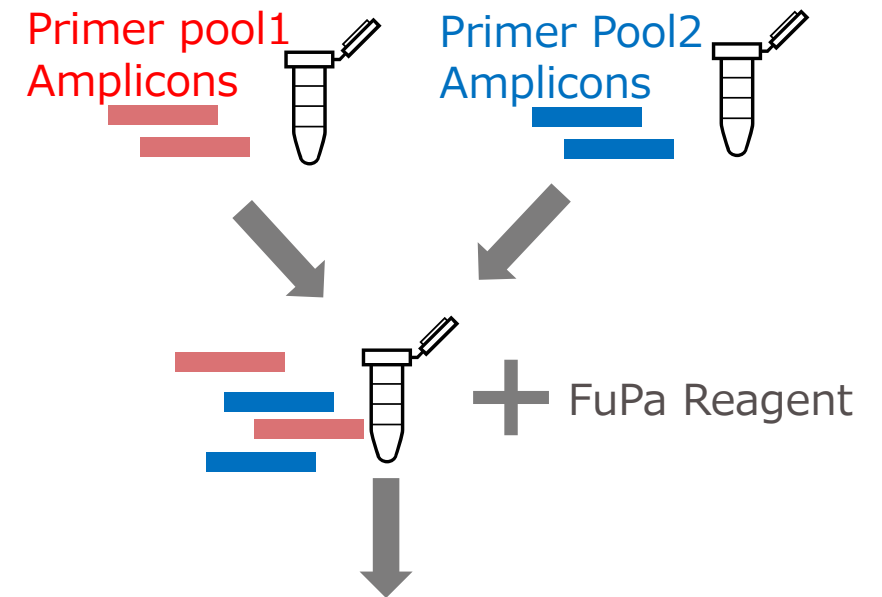
(3) Program名: FUPA

50°C, 10 分

55°C, 10 分

62°C, 20 分

10°C, hold



★プロトコル③以降は、下記資料をご参照ください。

【日本語ウェビナー】AmpliSeq for Illumina パネル解析をはじめよう -製品紹介とライブラリー調製編-【イルミナiSchool 初級】

<https://jp.illumina.com/events/webinar/2018/webinar-180425-j.html>

【Reference Guide】AmpliSeq for Illumina Custom and Community Panels Reference Guide

<https://support.illumina.com/downloads/ampliseq-for-illumina-custom-and-community-panels-reference-guide-1000000036408.html?langsel=/us/>

(3) Amplicon workflow

- AmpliSeq for Illumina SARS-CoV-2 Research Panel
- AmpliSeq for Illuminaのワークフロー
- AmpliSeq for Illuminaのプロトコルと注意点
- シーケンス条件
- コスト試算



AmpliSeq for Illumina Custom Panel : シーケンス条件

- ・ライブラリの調製後、下記のスタート濃度から、各装置の推奨希釈プロトコルを開始してください。
- ・最終ライブラリ濃度の最適化が必要な場合があります。
- ・2 x 151 bpでのランを推奨します。

Platform	ライブラリ希釈スタート濃度	入力時最終ライブラリ濃度
iSeq 100	2 nM	50 pM
MiniSeq	2 nM	1.1 – 1.9 pM
MiSeq (v3 reagents)	2 nM	7 - 9 pM
NextSeq 500/550	2 nM	1.1 – 1.9 pM



★ Corona virus protocol
・1サンプルあたりのリード数は0.25Mを推奨

(3) Amplicon workflow

- AmpliSeq for Illumina SARS-CoV-2 Research Panel
- AmpliSeq for Illuminaのワークフロー
- AmpliSeq for Illuminaのプロトコルと注意点
- シーケンス条件
- コスト試算



シーケンスコスト試算：AmpliSeq SARS-CoV-2 Panelによる Amplicon workflow

●ライブラリ調製コスト：96 sampleを調製する場合

* サンプル採取およびRNA抽出のコストは含みません。

キット名	1キットあたりのサンプル数	カタログ番号	キットのコスト（円）	1サンプルあたりのコスト（円）
Ampliseq™ cDNA Synthesis for Illumina	100 reactions	20022654	¥59,000	¥590
AmpliSeq™ SARS-CoV-2 Research Panel for Illumina	2,250 reactions	20020496	¥244,500	¥109
AmpliSeq for Illumina Library PLUS	96 reactions	20019102	¥1,122,000	¥11,688
AmpliSeq for Illumina CD Indexes Set A	96 index, 96 sample	20019105	¥73,900	¥770
合計				¥13,156

●シーケンスコスト： 2x151bp、 1サンプルあたり 0.25Mリードの場合

シーケンサー	iSeq100		MiniSeq		MiSeq			
	キット (サイクル数)	300 Cycle	MidOutput 300 cycle	HighOutput 300 cycle	Nano 300 cycle	Micro 300 cycle	v2 300 cycle	v3 600 cycle
カタログ番号	20021533	FC-420-1004	FC-420-1003	MS-103-1001	MS-103-1002	MS-102-2002	MS-102-3003	
キットのリード数 (CPF**)	4M	8M	25M	1M	4M	15M	25M	
1ラン当たりサンプル数	16	32	100	4	16	60	100	
キットのコスト（円）	¥100,100	¥96,600	¥270,500	¥51,000	¥76,700	¥183,000	¥268,100	
1サンプルあたりのコスト（円）	¥6,256	¥3,019	¥2,705	¥12,750	¥4,794	¥3,050	¥2,681	

* 2020年6月現在の希望納入価格（税抜）に基づきます。

** Cluster Pass Filterを通過したリード



3. BaseSpace Sequence Hub (BSSH) を使った情報解析





BaseSpaceで利用可能なSARS-CoV-2解析ツール

	Tool/Pipeline	Overview	Price
検出・同定	 DRAGEN Metagenomics	<ul style="list-style-type: none"> • ホストゲノムの除去、Kraken 2を使用した分類同定 • Krona plotなど視覚的にわかりやすい解析レポートの生成 	10月までアプリ 使用料無料
	 DRAGEN RNA Pathogen Detection	<ul style="list-style-type: none"> • Respiratory virus panel for Nextera Flexが使用可能 • DRAGEN k-mer matcherを使用したウイルス検出 	
データ共有	 GISAID Submission App	<ul style="list-style-type: none"> • GISAIDとのデータ共有を簡便化する 	
	 SRA Import App	<ul style="list-style-type: none"> • SRAデータベースから簡単にデータをインポート 	












引き続きアップデートや新規リリースが見込まれます

BaseSpaceで利用可能なSARS-CoV-2解析ツール

Tool/Pipeline	Overview	Price
 DRAGEN Enrichment	<ul style="list-style-type: none">• Nextera Flex + Respiratory Virus Panelに対応	<ul style="list-style-type: none">• 5 iCredit /node hr
 DNA Amplicon	<ul style="list-style-type: none">• AmpliSeq for Illumina + AmpliSeq Coronavirus Panelに対応	<ul style="list-style-type: none">• 3 iCredit /node hr

引き続きアップデートや新規リリースが見込まれます

Illumina Workflows

Workflow	Library Prep	Sequencer Fit	Analysis	Data Sharing
<p>Shotgun Metagenomics (RNA)</p>	 <p>Requires RNA extraction followed by TruSeq stranded total RNA</p>	 <p>NextSeq, NovaSeq</p>	<p>NEW</p>  <ul style="list-style-type: none"> • DRAGEN Metagenomics 	
<p>Enrichment</p>	 <p>Requires cDNA followed by Nextera DNA Flex for Enrichment + Respiratory Virus Oligo Panel</p>	 <p>iSeq, MiniSeq, MiSeq</p>	<p>NEW</p>  <ul style="list-style-type: none"> • DRAGEN Enrichment • DRAGEN RNA Pathogen Detection 	<p>NEW</p>  <p>GISAIID Submission</p> 
<p>Amplicon</p>	 <p>AmpliSeq for Illumina SARS-CoV-2 Research Panel</p>	 <p>iSeq, MiniSeq, MiSeq</p>	 <ul style="list-style-type: none"> • DNA Amplicon 	

3-1. パブリックデータ



パブリックデータのご用意

データ名	内容
Respiratory Virus Panel reference file	<ul style="list-style-type: none">• Nextera Flex for Enrichment, Respiratory Virus Panelのリファレンスfastaとbedファイル
NFE_CoVOC43_95UHR_Enriched	<ul style="list-style-type: none">• OC43株5% + UHR 95%, NFE + RVSパネルで濃縮• Dragen Metagenomicsアプリの実行例
NFE_CoVOC43_95UHR_Unenriched	<ul style="list-style-type: none">• OC43株5% + UHR 95%, NFEライブラリ• Dragen Metagenomicsアプリの実行例
NFE_CoVOC43_Enriched	<ul style="list-style-type: none">• OC43株, NFE + RVSパネルで濃縮• Dragen Metagenomicsアプリの実行例
NFE_CoVOC43_Unenriched	<ul style="list-style-type: none">• OC43株 + UHR 95%, NFEライブラリ• Dragen Metagenomicsアプリの実行例

引き続きアップデートが見込まれます

サンプルプロジェクト

データ名	入手場所	内容
DRAGEN Metagenomics Sample Project	アプリのトップページから	・ アプリのデモラン (de-host ON/OFFなど)
DRAGEN RNA Pathogen Detection Sample Project	アプリのトップページから	・ アプリのデモラン (異なるリファレンスでの実行結果など) ・ SFS参考リファレンスのダウンロード

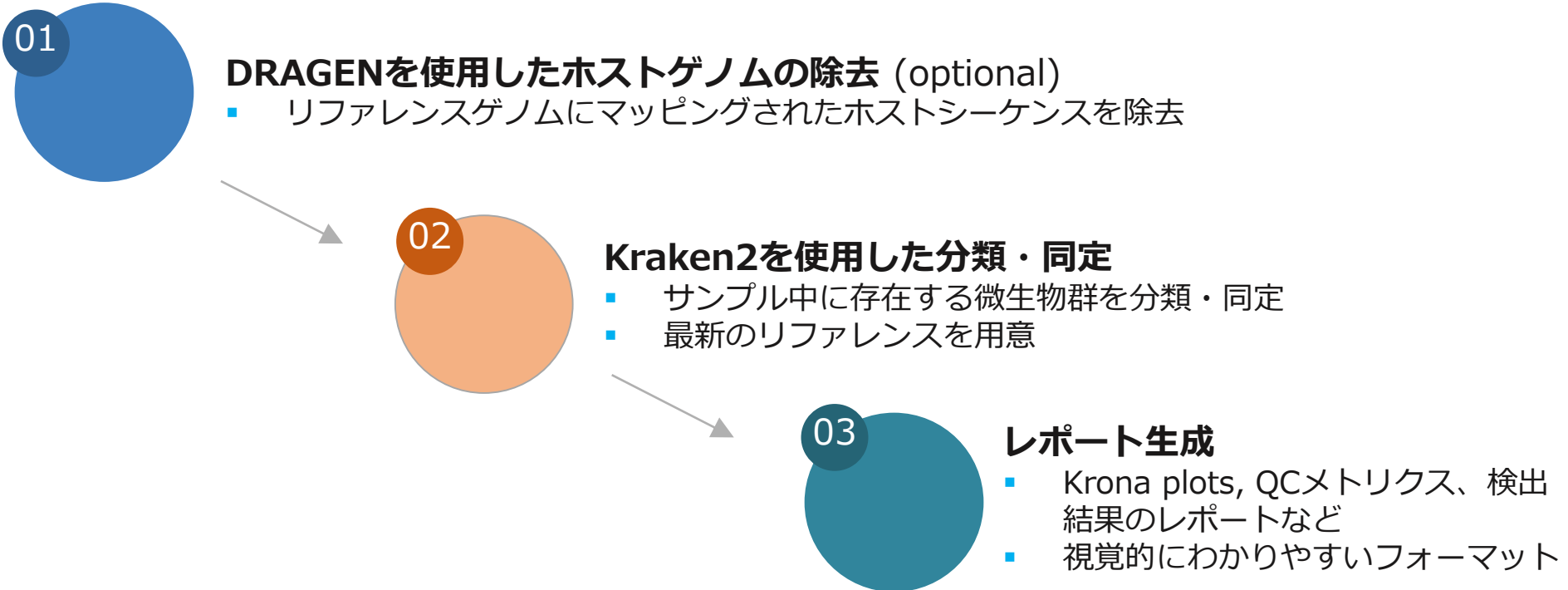
引き続きアップデートが見込まれます

DRAGEN Metagenomics Pipeline

ショットガンメタゲノムワークフロー



DRAGEN Metagenomicsアプリワークフロー



インプットフォーム

The image shows a screenshot of the DRAGEN configuration form. Several sections are highlighted with red dashed boxes. The "Reference Database" section includes a dropdown menu set to "MiniKraken2 (March 2020)". The "Custom Reference Database" section has a "Select Dataset File(s)" button. The "Host Reference" section has a dropdown menu set to "Human (hg38 Alt-Aware, with HLAs)". The "Dehosting" section has a checked checkbox for "Enable Dehosting". The "Organism Detection Report" section has three radio buttons, with "From text input" selected. The "Organism Identifiers" section has a text input field containing "Severe acute respiratory syndrome coronavirus 2", "Influenza A", and "Influenza B".

Reference Database

- MiniKraken2
- Extended Kraken2

Dehosting

Dragenによるホストの除去を実行するか

Organism Detection report, Organism Identifier

実行レポートのサンプル名の指定

結果サマリー CoV2_Enriched

AGGREGATE SUMMARY

Aggregate Organism Detection Report [Export \(CSV\)](#)

Color Palette: 

[Copy](#) [CSV](#) [Excel](#)

サンプル名

分類群名

Search:

Sample ID	Human coronavirus OC43	Influenza A	Influenza B
CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	High Detected (1740576)	ND (0)	ND (0)
CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep2	High Detected (1308243)	ND (0)	ND (0)
CoVOC43_Thermo_cDNA_Synthesis_Enriched_Rep1	High Detected (5235615)	ND (0)	ND (0)
CoVOC43_Thermo_cDNA_Synthesis_Enriched_Rep2	High Detected (5576985)	ND (0)	ND (0)

Showing 1 to 4 of 4 entries

検出の様子

Sample identifierの指定

<https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi>

The "Token set" option returns longer names that include the search terms, e.g., hybrid taxa. See what happens if you query "Bos taurus" using the "Complete match" option versus the "Set of tokens" option. The "Phonetic search" option can be used when you are not sure about the exact spelling of an organism name. It tries to find the phonetically closest match.

This is the top level of the taxonomy database maintained by NCBI/GenBank. You can explore any of the taxa listed below by clicking on the link.

- [Archaea](#)
- [Bacteria](#)
- [Eukaryota](#)
- **[Viruses](#)**
- [Other](#)
- [Unclassified](#)

These are direct links to some of the organisms commonly used in molecular research projects:

Arabidopsis thaliana	Escherichia coli
Bos taurus	Hepatitis C virus
Caenorhabditis elegans	Homo sapiens
Chlamydomonas reinhardtii	Mus musculus
Danio rerio (zebrafish)	Mycoplasma pneumoniae
Dictyostelium discoideum	Oryza sativa
Drosophila melanogaster	Plasmodium falciparum

Disclaimer: The NCBI taxonomy database is not an authoritative source for nomenclature or classification - please consult the literature.

NCBI Taxonomy Browserで参照できる名称

Sample identifierの指定

Kraken 2 Taxonomic Sequence Classification System



Home Manual Downloads CCB » Software » Kraken2

2020/03 UPDATE

All links are currently being updated.
ftp://ftp.ccb.jhu.edu/pub/data/kraken2

名前	サイズ	更新日
16S_Greengenes13.5_20200326.tgz	72.1 MB	2020/03/26 20:19:00
16S_RDP11.5_20200326.tgz	168 MB	2020/03/26 20:21:00
16S_Silva132_20200326.tgz	117 MB	2020/03/26 18:28:00
16S_Silva138_20200326.tgz	112 MB	2020/03/27 13:09:00
README.md	1.9 kB	2020/03/27 15:29:00
README.txt		
minikraken_8GB_20200326.tgz		
minikraken_8GB_20200312.tgz		
old/		

```
>kraken:taxid|273360|NC_043646.1 Potosi virus strain IL94-1899 segment L, complete sequence
>kraken:taxid|273360|NC_043645.1 Potosi virus strain IL94-1899 segment M, complete sequence
>kraken:taxid|273360|NC_043647.1 Potosi virus strain IL94-1899 segment S, complete sequence
>kraken:taxid|204928|NC_021926.1 Calibrachoa mottle virus isolate California, complete genome
>kraken:taxid|21698|NC_006213.1 Human coronavirus OC43 strain ATCC VR-759, complete genome
>kraken:taxid|31631|NC_006213.1 Human coronavirus OC43 strain ATCC VR-759, complete genome
>kraken:taxid|75734|NC_006213.1 Human coronavirus OC43 strain ATCC VR-759, complete genome
>kraken:taxid|2169971|NC_006213.1 Human coronavirus OC43 strain ATCC VR-759, complete genome
>kraken:taxid|11020|NC_006213.1 Human coronavirus OC43 strain ATCC VR-759, complete genome
>kraken:taxid|1076255|NC_006213.1 Human coronavirus OC43 strain ATCC VR-759, complete genome
>kraken:taxid|10869|NC_006213.1 Human coronavirus OC43 strain ATCC VR-759, complete genome
```

Human coronavirus OC43 strain ATCC VR-759, complete genome
NCBI Reference Sequence: NC_006213.1
[FASTA](#) [Graphics](#)

Go to:

LOCUS NC_006213 30741 bp RNA linear VRL 21-FEB-2019
DEFINITION Human coronavirus OC43 strain ATCC VR-759, complete genome.
ACCESSION NC_006213
VERSION NC_006213.1
DBLINK BioProject: [PRJNA485481](#)
KEYWORDS RefSeq.
SOURCE Human coronavirus OC43 (HCoV-OC43)
ORGANISM [Human coronavirus OC43](#)
Viruses; Riboviria; Nidovirales; Coronidovirineae; Coronaviridae;
Orthocoronavirinae; Betacoronavirus; Embecovirus.

アプリが認識できるフォーマット



ORGANISM [Human coronavirus OC43](#)

レポート

DRAGEN METAGENOMICS RESULTS FOR SAMPLE [COVOC43_ILLUMINA_CDNA_SYNTHESIS_ENRICHED_REP1](#)

[Kraken2 Classification Report](#)
 [Organism Detection Report](#)

エクセル形式のレポート

SAMPLE INFORMATION

Fragments Total	Fragments After De-hosting	Fragments Classified	Fragments Unclassified
6948512		3200309	373947

サンプル情報

レポート

DRAGEN METAGENOMICS RESULTS FOR SAMPLE [COVOC43_ILLUMINA_CDNA_SYNTHESIS_ENRICHED_REP1](#)

- Kraken2 Classification Report
- Organism Detection Report

エクセル形式のレポート

SAMPLE INFORMATION

Fragments Total	Fragments After De-hosting	Fragments
6948512		3200309

サンプル情報

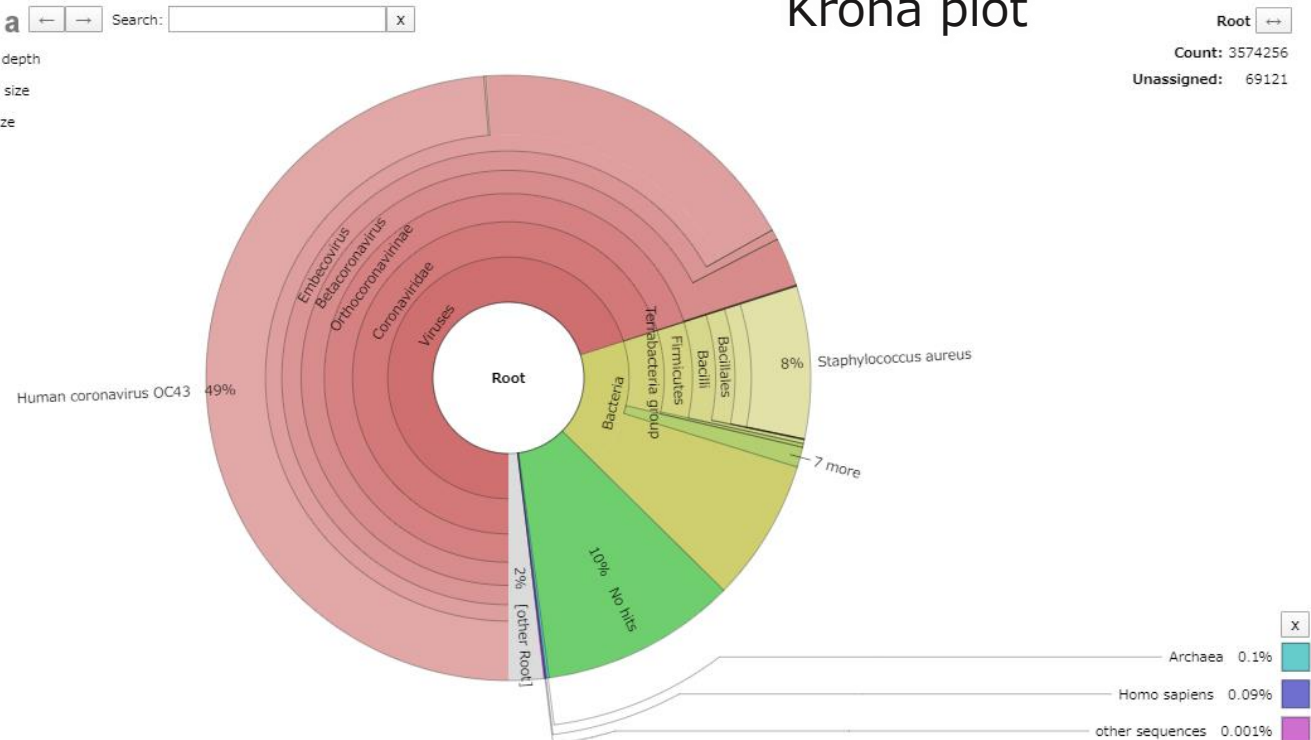
KRONA CLASSIFICATION CHART

K r o n a ← → Search:

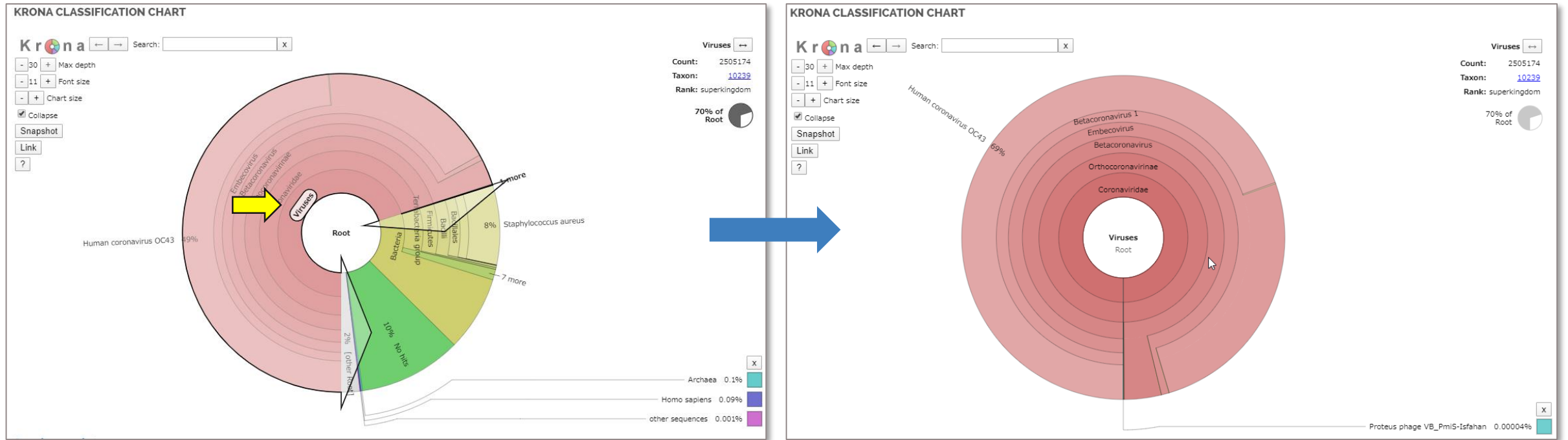
- 30 + Max depth
- 11 + Font size
- + Chart size

Collapse
Snapshot
Link
?

Krona plot



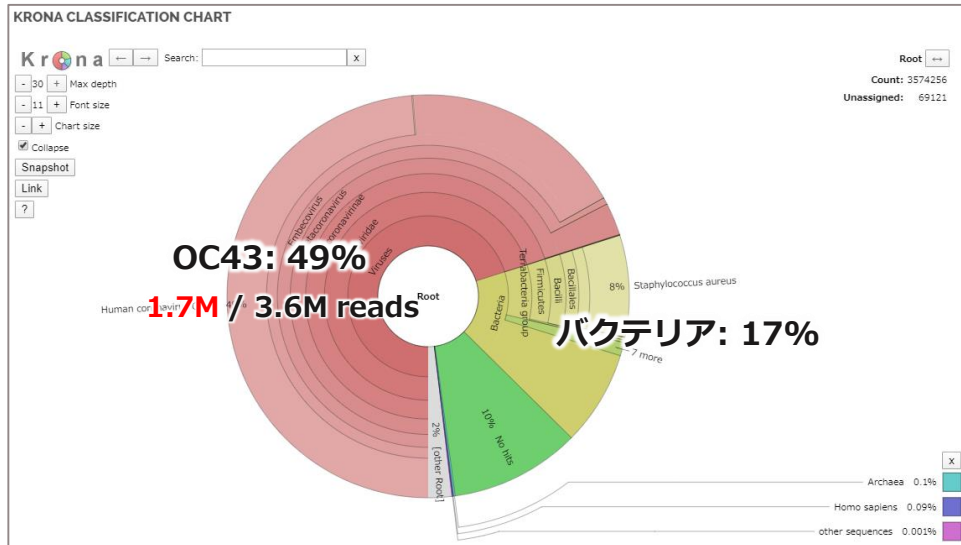
レポート



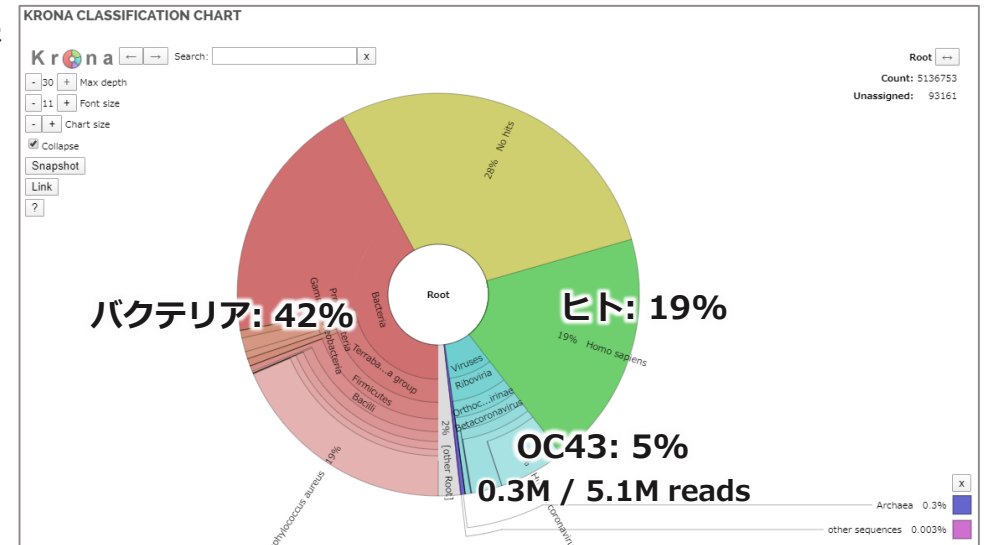
ルートを変更し、対象を絞り込んでいくことができる

アプリケーションによる検出量差異 De-hosting OFF

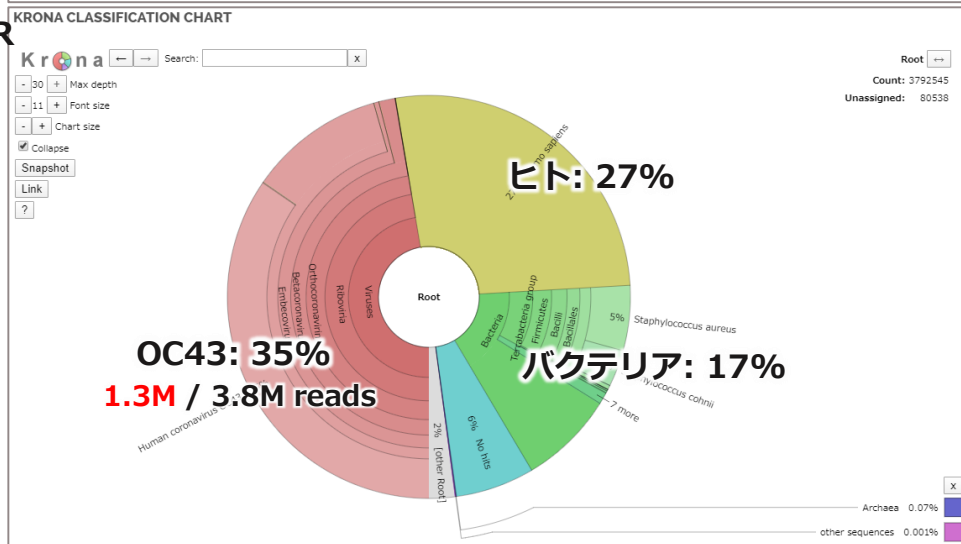
ウイルス株
RVS Enriched



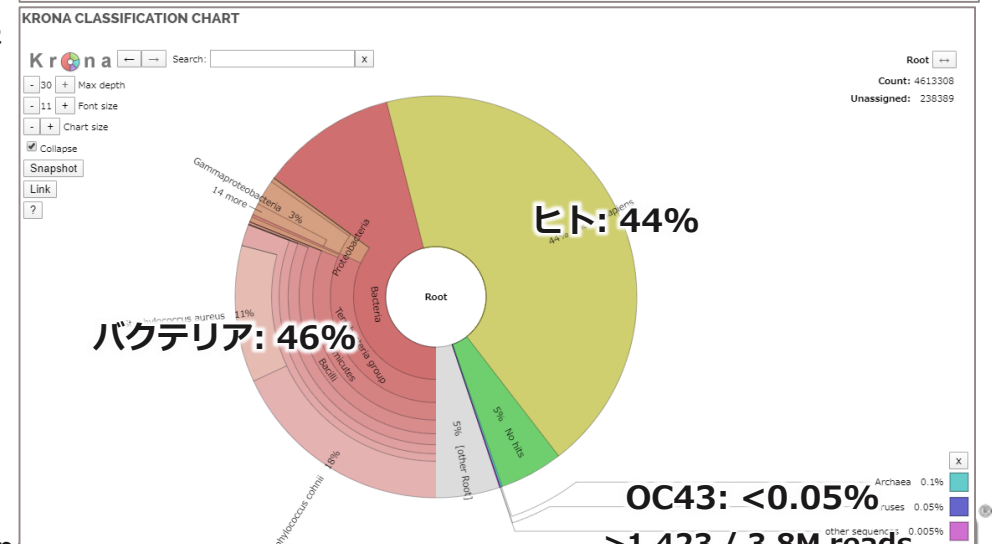
ウイルス株



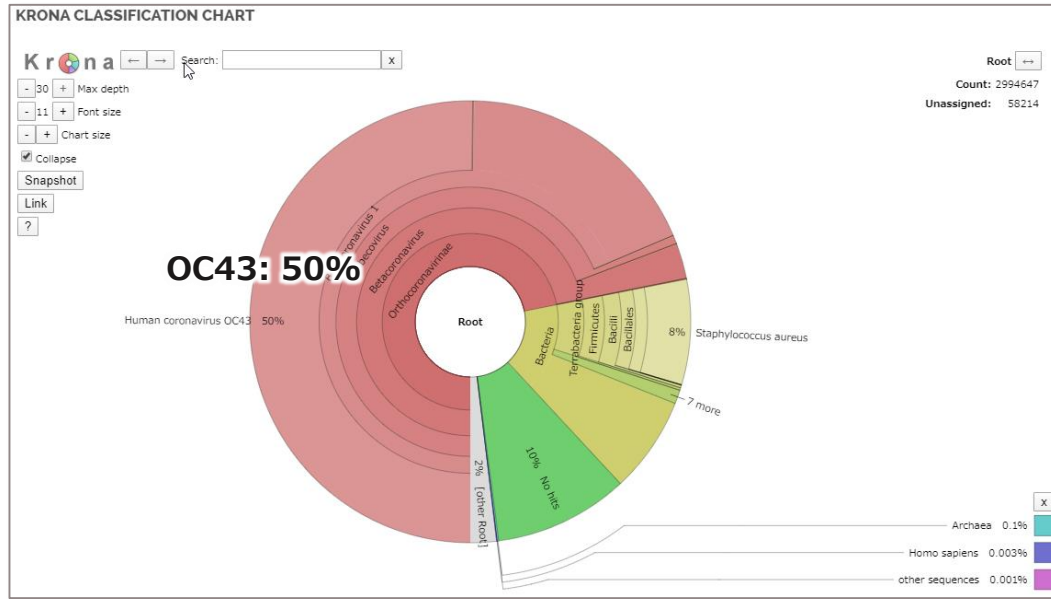
ウイルス株+UHR
RVS Enriched



ウイルス株
+UHR

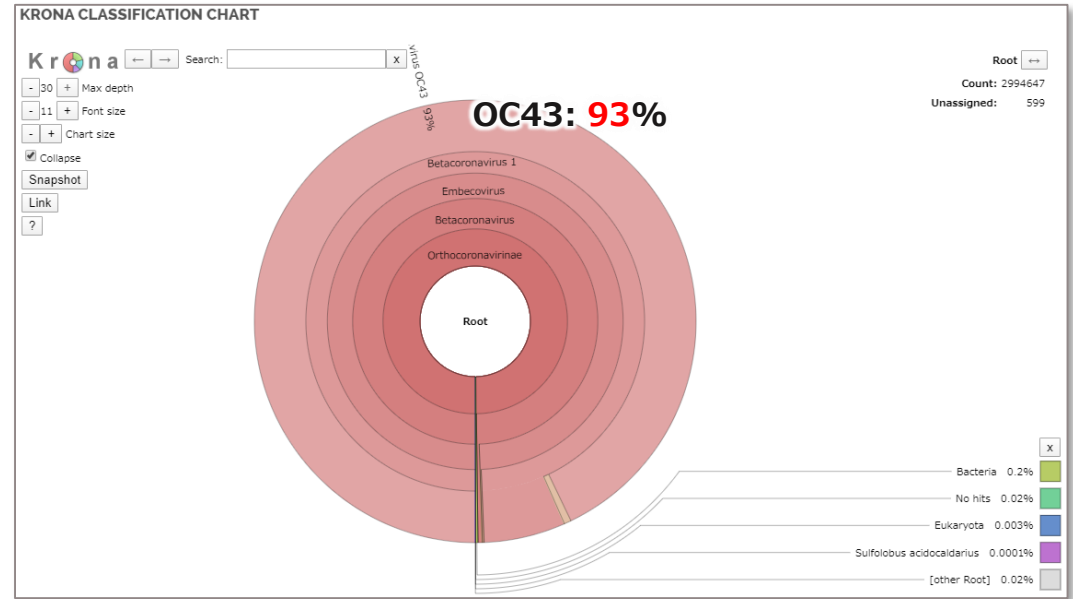


リファレンスの選択 CoV2_Enriched



MiniKraken2: **15min**

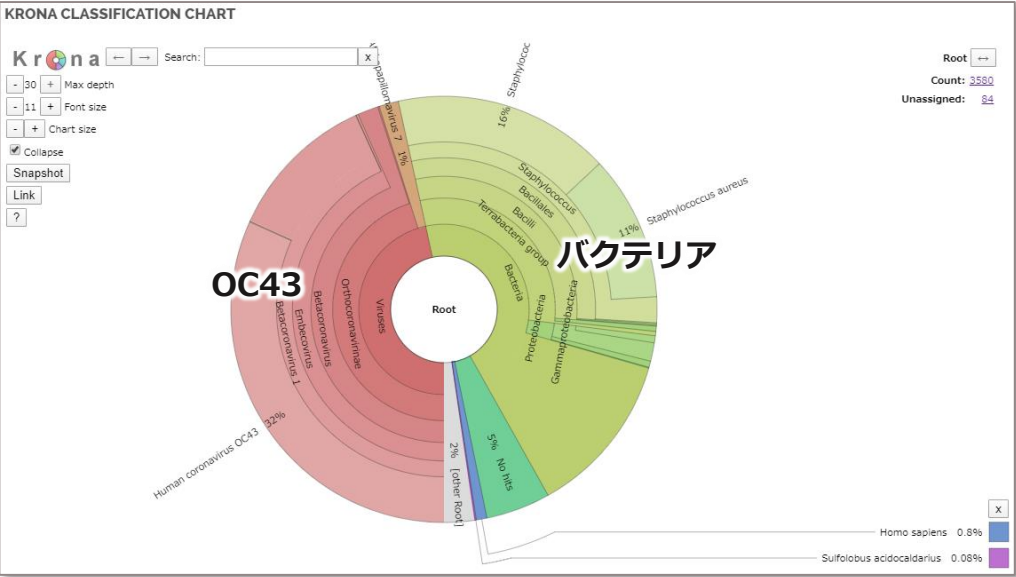
Fragments Classified	Fragments Unclassified
2699548	295099



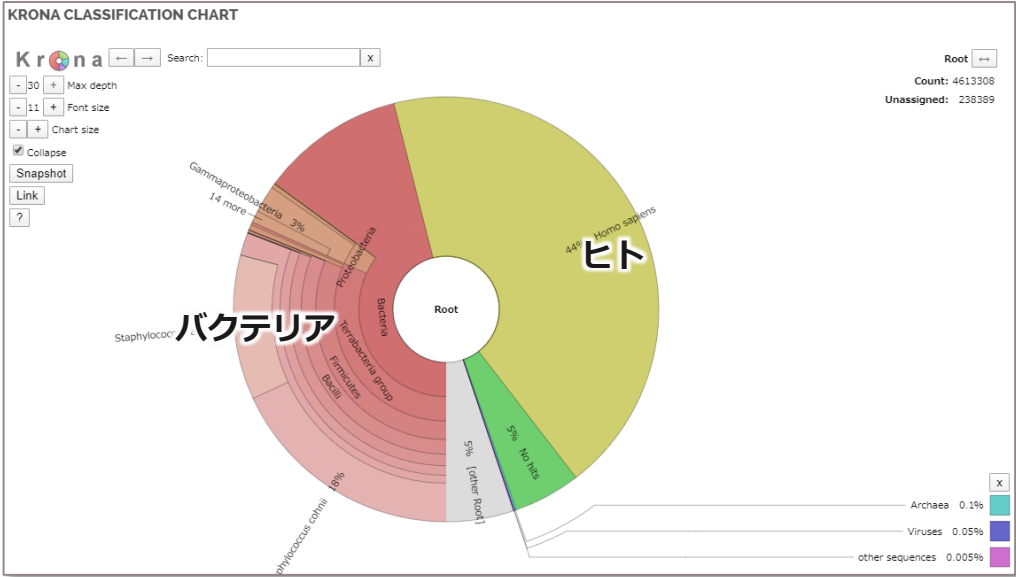
Extended Kraken2: **52min**

Fragments Classified	Fragments Unclassified
2994126	521

De-hostingの有無 CoV2+UHR



De-hosting ON **12min**
1,136 / 3,580 reads after De-hosting



De-hosting OFF **29min**
1,423 / 4,613,308 reads

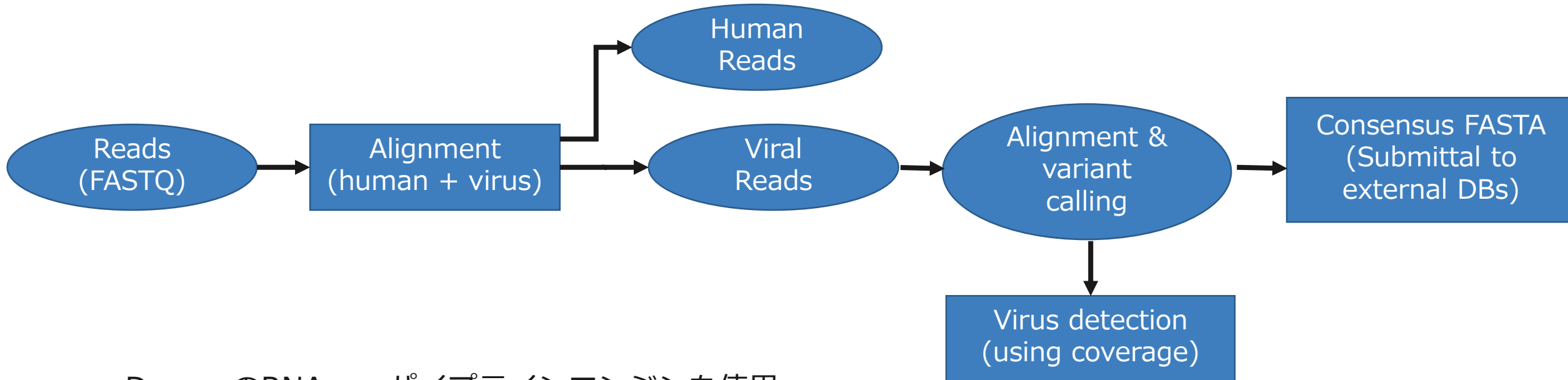
DRAGEN RNA Pathogen Detection Pipeline

エンリッチメントワークフロー



注) 資料内テスト解析には仮のデータを使用したため、実際のパイプラインの実施時とは数値などが異なる場合がございます

ワークフロー



- DragenのRNA-seqパイプラインエンジンを使用
- ヒトhg38に、下記を追加した情報を選択することができる
 - Content from Illumina Respiratory Virus Panel
 - 168 viral sequences from Seattle Flu Study
- ドラゲンアルゴリズムを使用したデュプリケートの除去やアライメントを行う
- 変異コール、コンセンサスFastaの作成を行う
- ドラゲnk-mer解析パイプラインも実装

インプットフォーム

 DRAGEN RNA Pathogen Detection v3.5.13
Illumina Inc.

Configuration

Reference ⓘ

Human (hg38) and Illumina Respiratory Virus Panel

Reference

- Human (hg38) and Illumina Respiratory Virus Panel
- Human (hg38) and Respiratory Viruses from the Seattle Flu Study

K-mer Analysis ⓘ

None

Generate k-mers from Reference

Use Existing k-mer List

K-mer Reference FASTA ⓘ

SARS-CoV-2_NC_045512.2

K-mer Analysis

K-mer解析の実行について指示

Accession list for coverage plots and FASTA generation ⓘ

NC_045512.2,MN908947.3

Accession list for coverage plots and FASTA generation
コンセンサスファスタの参照配列

結果サマリー

Percent of pathogen genome with > 5x coverage

Color Palette:

Copy CSV Excel

分類群名

サンプル名	Human coronavirus OC43	Human Respiratory syncytial virus 9320 (type B)	Human metapneumovirus (CAN97-83)	Human parainfluenza virus 1	Human coronavirus NL63	Human coronavirus 229E	Human parainfluenza virus 2	Influenza A virus (A/Hong Kong/1073/99(H9N2))
CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Detected (99.97%)	ND (4.91%)	ND (3.25%)	ND (1.54%)	ND (0.56%)	ND (1.21%)	ND (1.20%)	ND (1.16%)
CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep2	Detected (99.96%)	ND (3.98%)	ND (2.43%)	ND (1.38%)	ND (0.62%)	ND (1.22%)	ND (1.37%)	ND (0.85%)
CoVOC43_Thermo_cDNA_Synthesis_Enriched_Rep1	Detected (100.00%)	Detected (10.15%)	Detected (7.73%)	ND (3.34%)	ND (1.76%)	ND (2.40%)	ND (3.21%)	ND (2.89%)
CoVOC43_Thermo_cDNA_Synthesis_Enriched_Rep2	Detected (100.00%)	Detected (10.46%)	Detected (8.43%)	ND (4.03%)	ND (2.01%)	ND (2.87%)	ND (3.74%)	ND (2.64%)

検出の様子

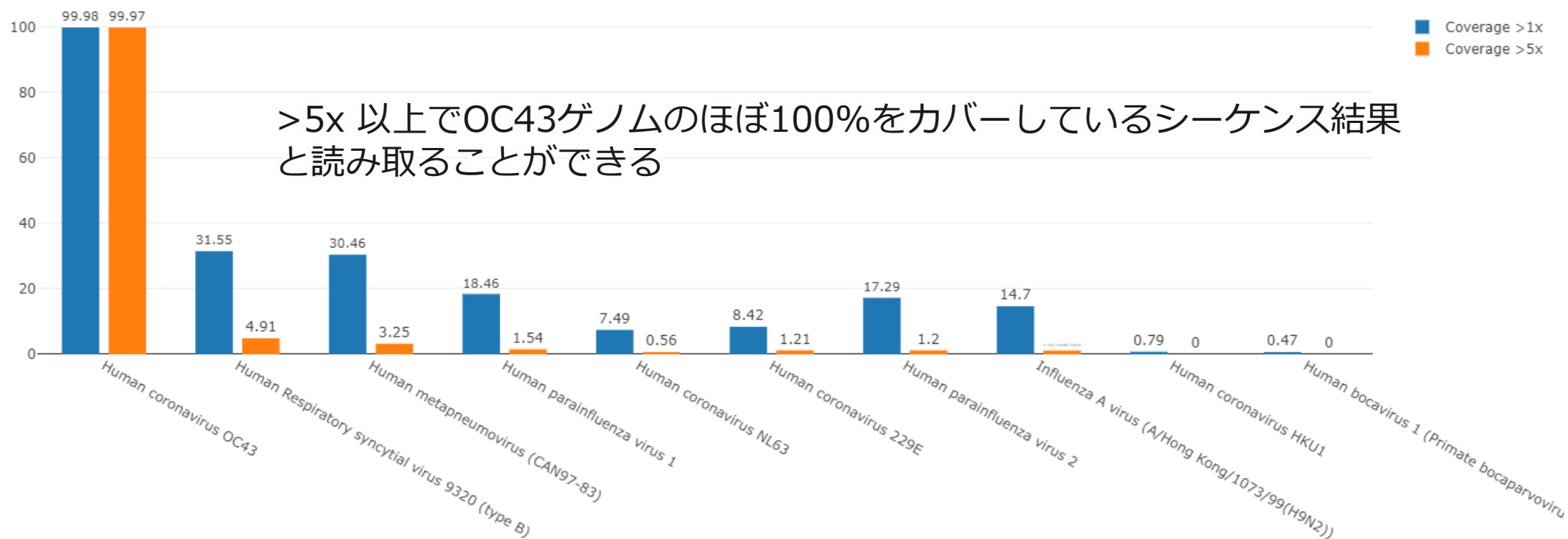
サンプル情報

Sample Information

Type	Read 1	Read 2	Total Reads	Mapped Reads
Paired	151	151	6948512	3473135



Percentage of viral genome with coverage >1x, >5x



>5x 以上でOC43ゲノムのほぼ100%をカバーしているシーケンス結果と読み取ることができる

コンセンサスFASTA

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

GenBank: MN908947.3

[GenBank](#) [Graphics](#)

>MN908947.3 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

```
ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGTTCTCTAAA  
CGAACTTTAAATCTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACTCAGCAGTATAATTAATAAC  
TAATTACTCTCTTCAACACCAACCAACTAATCTCTATCTTCTCCACGCTGCTTACCTTTCTGCTCTC  
TTGCAGCCGATCATCAGCACATCTAGGTTTTGTCGGGTTGACCGAAAGGTAAGATGGAGAGCCTTGTCT  
CCTGGTTTTCAACGAGAAAACACACGTCACACTCAGTTTTGCTGTTTTACAGGTTCCGCGACGTTGCTGTA  
GTGGCTTTGGAGACTCCGTGGAGGAGGCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGG  
CTTAGTAGAAGTTGAAAAGGCGTTTTGCTCAACTTGAACAGCCCTATGTTTCAACAAAGTTCCGGAT  
GCTCGAACTGCACCTCATGGTCTGTTATGGTTGAGCTGGTAGCAGAACTCGAAGGCATTGATACGGTC  
GTAGTGGTAGACACTTGGTGTCTTGTCCCTCATGTGGCGGAAATACCAGTGGCTTACCGCAAGGTTCT  
TCTTGTAAAGACGGTAATAAAGGAGCTGGTGGCCATAGTTACGGCGCCGATCTAAAGTCATTTGACTTA
```

SRR11801823_Genome-Sequence-of-SARS-CoV-2 とのコンセンサスFASTA抜粋

```
6 ↓  
7 TTGCAGCCGATCATCAGCACATCTAGGTTTTGTCGGGTTGACCGAAAGGTAAGATGGAGAGCCTTGTCT ↓  
8 ↓
```

C > T

```
ATGAGTTCCCTGTGTTGTGGCAGATGCTGTCAAAAACTTTGCAACCAAGTATCTGAATTACTTACACC  
ACTGGGCATTGATTTAGATGAGTGGAGTATGGCTACATACTACTATTTTATGAGTCTGGTGAAGTTAAA  
TTGGCTTACATATGATTTGTTCTTTTACCTCCAGATGAGGATGAAGAAGAAGGATTTGTGAAGAAG  
AAGAGTTTGAAGCCTCAACTCAATATGAGTATGGTACTGAAGATGATTAACAAGGTAACCTTTGAAAT  
TGGTGCCACTTCTGCTGCTTCAACCTGAAGAAGAGCAAGAAGAAGATTGGTTAGATGATGATGATCAA  
CAAACTGTTGGTCAACAAGAAGCGCAGTGAAGCAATCAGACAACTACTATTCAACAATTGTTGAGGTT  
AAGCTCAATGAGAGTGAAGTACACCACTTGTGCACTATTGAAGTAAAGTTTACTGCTTATTT  
AAAACTTACTGACAATGTATACATTAATAATGCAGACATTTGGAAGAAGCTAAAAAGGTAACCAACA  
GTGGTTGTTAATGCAGCAGTATTTACCTTAAACATGAGGAGGTTTGCAGGAGCCTTAAATTAAGGCTA  
CTAACAATGCCATGCAAGTTGAATCTGATGATTACATAGCTACTAATGGACCACTTAAAGTGGTGGTAG  
TTGTGTTTTAAGCGCACACAATCTTGTAAACACTGTCTTCACTGTTGTCGGCCCAATGTTAACAAGGT  
GAAAGACTTCAACTTCTTAAAGTGTCTTAAATAATTAATCAGCAGCAAGTTCTACTTGCACCAATAT  
TATCAGCTGGTATTTTGGTGTGACCTTACATTTCTTAAAGATTTTGTAGATCTGTTTCGCACAAA  
TGCTACTTAGCTGCTTTGATAAAAACTCTATGACAAACTTGTTCAGCTTTTTGGAATGAAGAGT  
GAAAAGCAAGTTGAACAAGAATCGCTGAGATTCTAAGAGGAAAGTTAAGCCATTTAATACTGAAAGTA  
AACCTTCAAGTTGAACAGAGAAAACAAGATGATAAGAAAATCAAGCTTGTGTTGAAGAAGTTACAACA  
TCTGGAAGAAAATCAAGTTCTCACAGAAAATCTGTTACTTTATATTGACATTAATGGCAATCTTCAATCA
```

SRR11247076_RNA-Seq-of-Severe-acute-respiratory-syndrome -coronavirus-2とのコンセンサスFASTA抜粋

```
1 ↓  
2 AAAACTTACTGACAATGTATACATTAATAATGCAGACATTTGGAAGCAAGCTAAAAAGGTAACCAACA ↓  
3 ↓
```

A > C

デフォルトだとMN908947.3の配列情報を参照し、変異情報を含めたFASTAファイルとして体裁の整った29,903塩基のコンセンサスFASTAを自動生成する

アライメントと変異コール

Map/Align Metrics

マッピング/アライメントメトリクス (一部分のみ表示)		
	Total input reads	6948512
	Number of duplicate marked reads	5496499
	Number of duplicate marked and mate reads removed	NA
	Number of unique reads (excl. duplicate marked reads)	1452013
	Reads with mate sequenced	6948512
	Reads without mate sequenced	0
	QC-failed reads	0
	Mapped reads	6947188
	Mapped reads R1	3474053
	Mapped reads R2	3473135
	Number of unique & mapped reads (excl. duplicate marked reads)	1450689
	Unmapped reads	1324

アライメントと変異コール

Map/Align Metrics

マッピング/アライメントメトリクス
(一部分のみ表示)


Total Input reads	6948512
Number of duplicate marked reads	5496499
Number of duplicate marked and mate reads removed	NA
Number of unique reads (excl. duplicate marked reads)	1452013
Reads with mate sequenced	6948512


Variant Caller Metrics

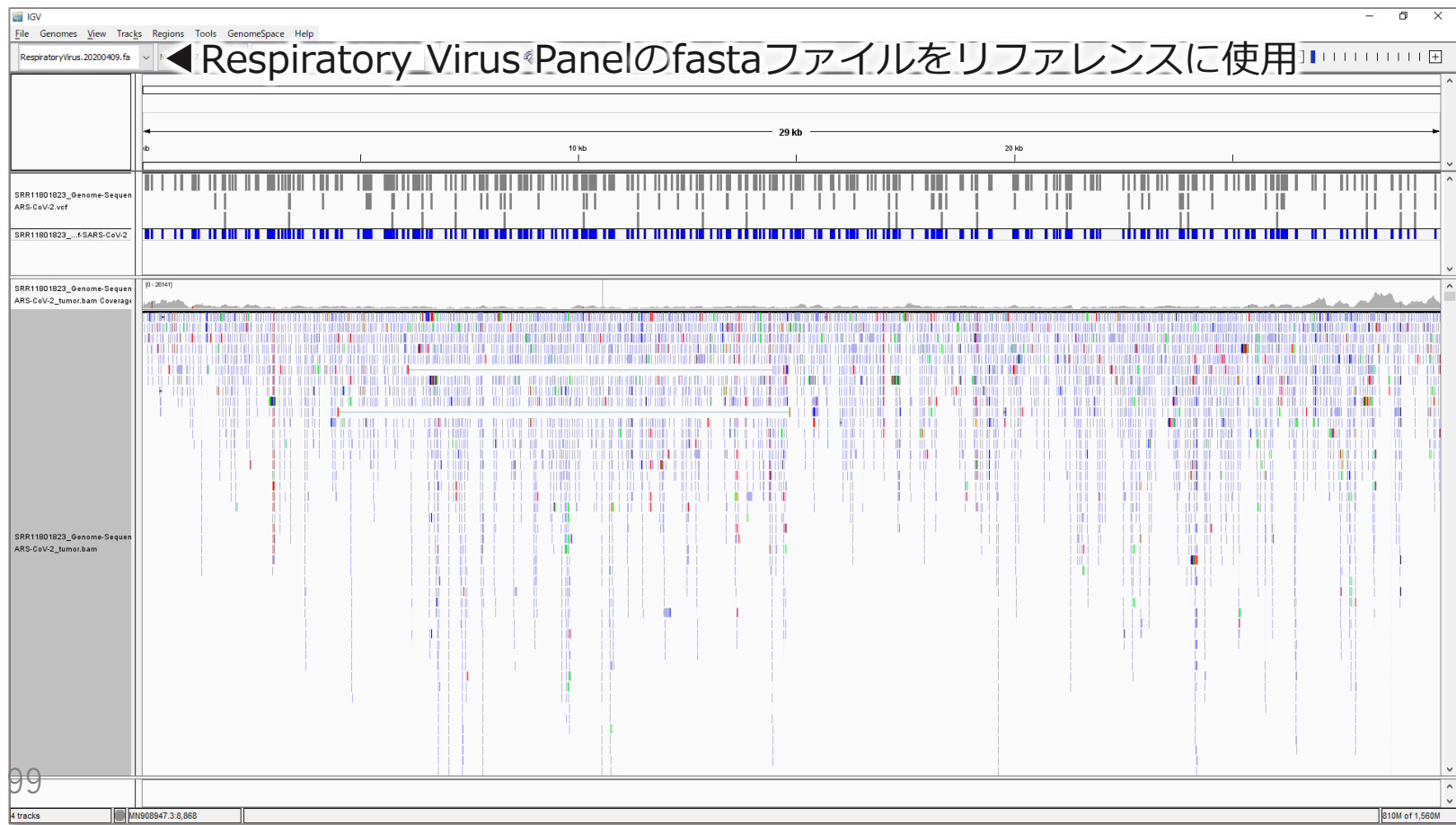
変異コールメトリクス
(一部分のみ表示)

VARIANT CALLER SUMMARY		Number of samples	1
VARIANT CALLER SUMMARY		Reads Processed	2403780
VARIANT CALLER SUMMARY		Child Sample	NA
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Total	483
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Biallelic	483
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Multiallelic	0
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	SNPs	476
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Insertions (Hom)	0
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Insertions (Het)	2
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Deletions (Hom)	0
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Deletions (Het)	5
VARIANT CALLER PREFILTER	CoVOC43_Illumina_cDNA_Synthesis_Enriched_Rep1	Indels (Het)	0

アライメントと変異コール

 SRR11801823_Genome-Sequence-of-SARS-CoV-2_tumor.bam	2020-05-18 09:34	bam	SRR11801823_Genome-Sequence-of-SARS-CoV-2_tumor.bam
---	------------------	-----	---

 SRR11801823_Genome-Sequence-of-SARS-CoV-2.vcf.gz	2020-05-18 09:34	gz	vc/SRR11801823_Genome-Sequence-of-SARS-CoV-2.vcf.gz
--	------------------	----	---



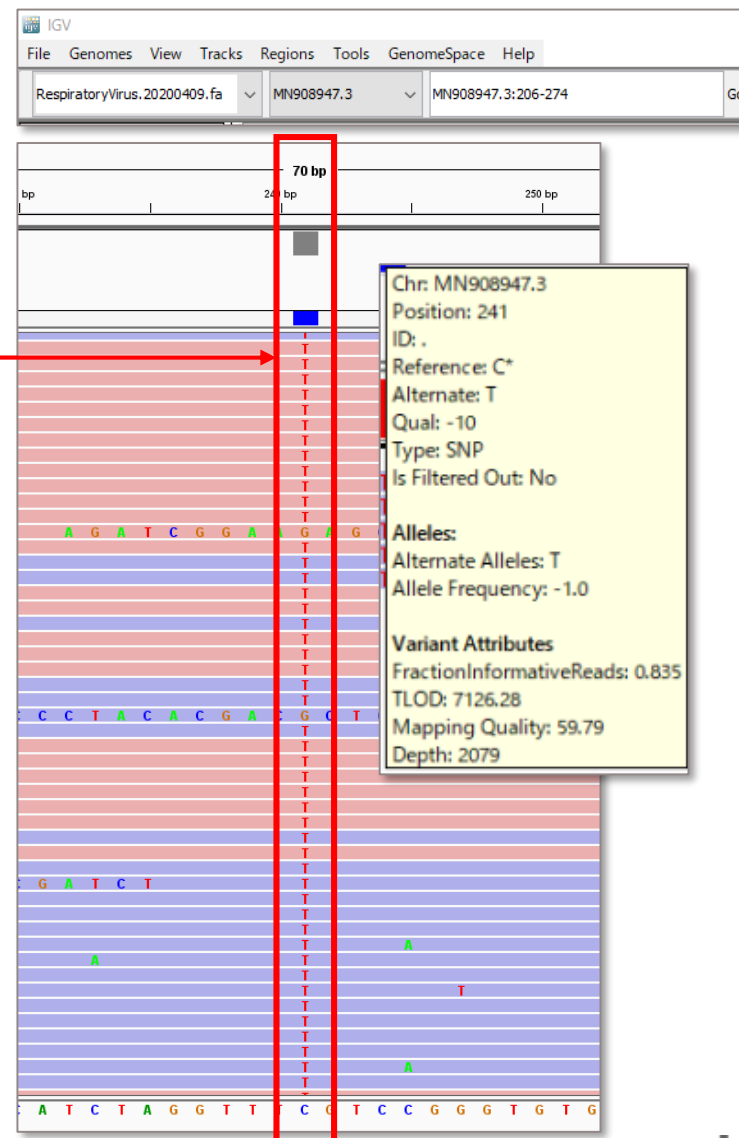
アライメントと変異コール

SRR11801823_Genome-Sequence-of-SARS-CoV-2
とのコンセンサスFASTA抜粋

```
6 ↓  
7 TTGCAGCCGATCATCAGCACATCTAGGTTTTGTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCCTTGTC ↓  
8 ↓
```

C > T

Fastaファイルに記載がない、マッピングメトリクス
の詳細を確認することも可能



DRAGEN Enrichment workflow

エンリッチメントワークフロー



注) 資料内テスト解析には仮のデータを使用したため、実際のパイプラインの実施時とは数値などが異なる場合がございます

インプットフォーム



Reference ⓘ

Custom

Custom DRAGEN Reference File ⓘ

Select Dataset File(s)

RespiratoryVirus_hg38_DNA.v8.tar

Custom FASTA Reference File ⓘ

Select Dataset File(s)

RespiratoryVirus.20200409.fa

Targeted Regions ⓘ

Custom BED (select BED file below)

Custom Target BED File ⓘ

Select Dataset File(s)

RespiratoryVirus.20200409.bed

Custom DRAGEN Reference File

Respiratory Virus Reference File プロジェクト内
- RespiratoryVirus_hg38_DNA.v8.tar

Custom FASTA Reference File

Respiratory Virus Reference File プロジェクト内
- RespiratoryVirus.20200409.fa
- RespiratoryVirus_hg38.20200409.fa

Target Regions, Custom Target BED File

Respiratory Virus Reference File プロジェクト内
- RespiratoryVirus.20200409.bed

アウトプット

SAMPLE INFORMATION ⁱ

Total PF Reads	Unique PF Reads	Percent Q30 Bases	Mean Target Coverage Depth
1,025,784	580,990	95.16%	2390.1

◀リード情報

ENRICHMENT SUMMARY ⁱ

Target Manifest	Total Length of Targeted Reference	Number of Targeted Regions	Padding Size
RespiratoryVirus.20200409_used.bed	596,713 bp	118	150 bp

Note: All enrichment values are calculated without padding unless otherwise stated. If any targeted region overlaps another region, the region positions will be adjusted to remove overlaps (check manifest file in the output directory for differences). The total length of the targeted reference excludes no-call positions (Ns) in the reference sequence and might result in a smaller value than expected from the targeted regions coordinates alone.

Read Level Enrichment ⁱ

Total Aligned Reads	Percent Aligned Reads	Targeted Aligned Reads	Read Enrichment	Padded Target Aligned Reads	Padded Read Enrichment
580,388	99.90%	639,293	110.15%	639,293	110.15%

◀濃縮ステップのメトリクス

Base Level Enrichment ⁱ

Total Aligned Bases	Percent Aligned Bases	Targeted Aligned Bases	Base Enrichment	Padded Target Aligned Bases	Padded Base Enrichment
71,668,731	99.87%	71,472,294	99.73%	71,472,294	99.73%

アウトプット

SAMPLE INFORMATION ⁱ

Total PF Reads	Unique PF Reads	Percent Q30 Bases	Mean Target Coverage Depth
1,025,784	580,990	95.16%	2390.1

◀リード情報

ENRICHMENT SUMMARY ⁱ

Target Manifest	Total Length of Targeted Reference	Number of Targeted Regions	Padding Size
RespiratoryVirus.20200409_used.bed	596,713 bp	118	150 bp

Note: All enrichment values are calculated without padding unless otherwise stated. If any targeted region overlaps another region, the region positions will be adjusted to remove overlaps (check manifest file in the output directory for differences). The total length of the targeted reference excludes no-call positions (Ns) in the reference sequence and might result in a smaller value than expected from the targeted regions coordinates alone.

Read Level Enrichment ⁱ

Total Aligned Reads	Percent Aligned Reads	Targeted
580,388	99.90%	639,293

Base Level Enrichment ⁱ

Total Aligned Bases	Percent Aligned Bases	Targeted
71,668,731	99.87%	71,472,293

SMALL VARIANTS SUMMARY ⁱ

	SNVs	Insertions	Deletions
Total Passing	38	24	5
Het/Hom Ratio	0	0	0
Ts/Tv Ratio	1.11	-	-

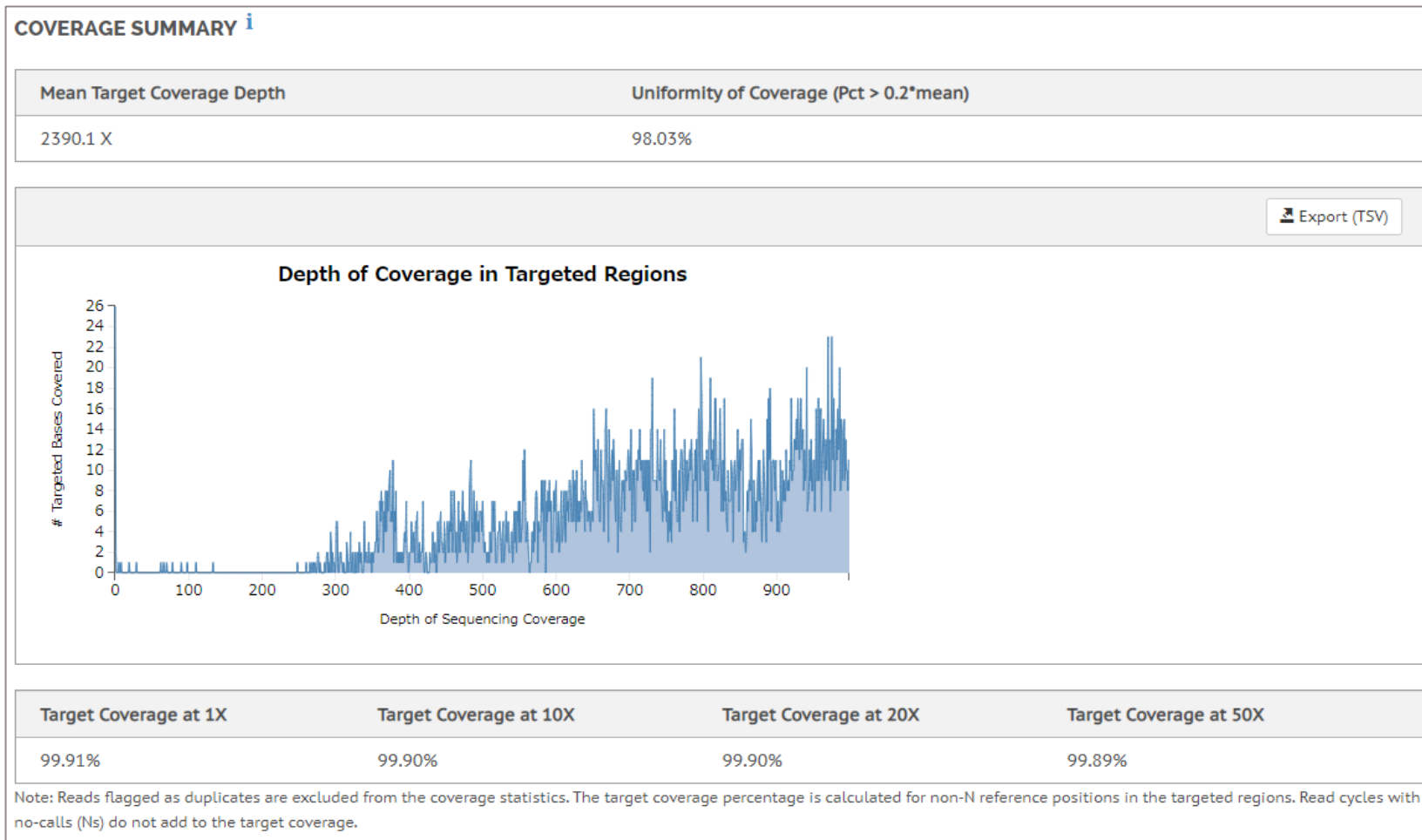
VARIANTS TABLE ⁱ

📄 Load Variants Table

Please note that for samples with a large number of variants, loading the variants table may result in slow performance. You can alternatively view the variants table content outside of the browser by downloading the "varianttable.txt" file available under "Output Files".

SNV, indelメトリクス▶

アウトプット



◀リードカバレッジ情報

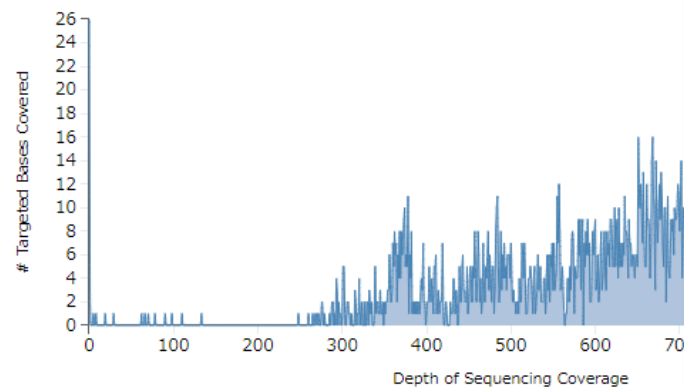
アウトプット

COVERAGE SUMMARY ⁱ

Mean Target Coverage Depth

2390.1 X

Depth of Coverage in Targeted Regions



Target Coverage at 1X

99.91%

Target Coverage at 10X

99.90%

Note: Reads flagged as duplicates are excluded from the coverage statistics. The target coverage is based on the number of unique reads. No-calls (Ns) do not add to the target coverage.

FRAGMENT LENGTH SUMMARY ⁱ

Fragment Length Median

132 bp

Minimum

35 bp

Maximum

477 bp

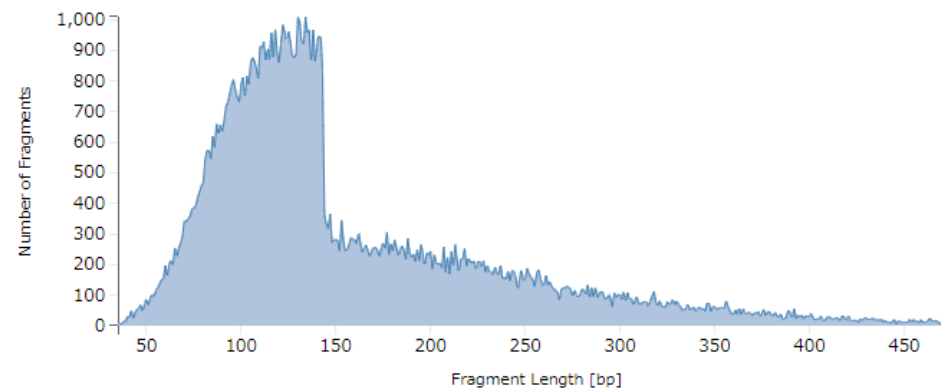
Standard Deviation

69 bp

Note: The minimum and maximum are calculated from values within approximately three standard deviations (excluding the lower and upper 0.15% of the data) to account for potential outliers.

[Export \(CSV\)](#)

Fragment Length Distribution



DUPLICATE INFORMATION ⁱ

Percent Duplicate Aligned Reads

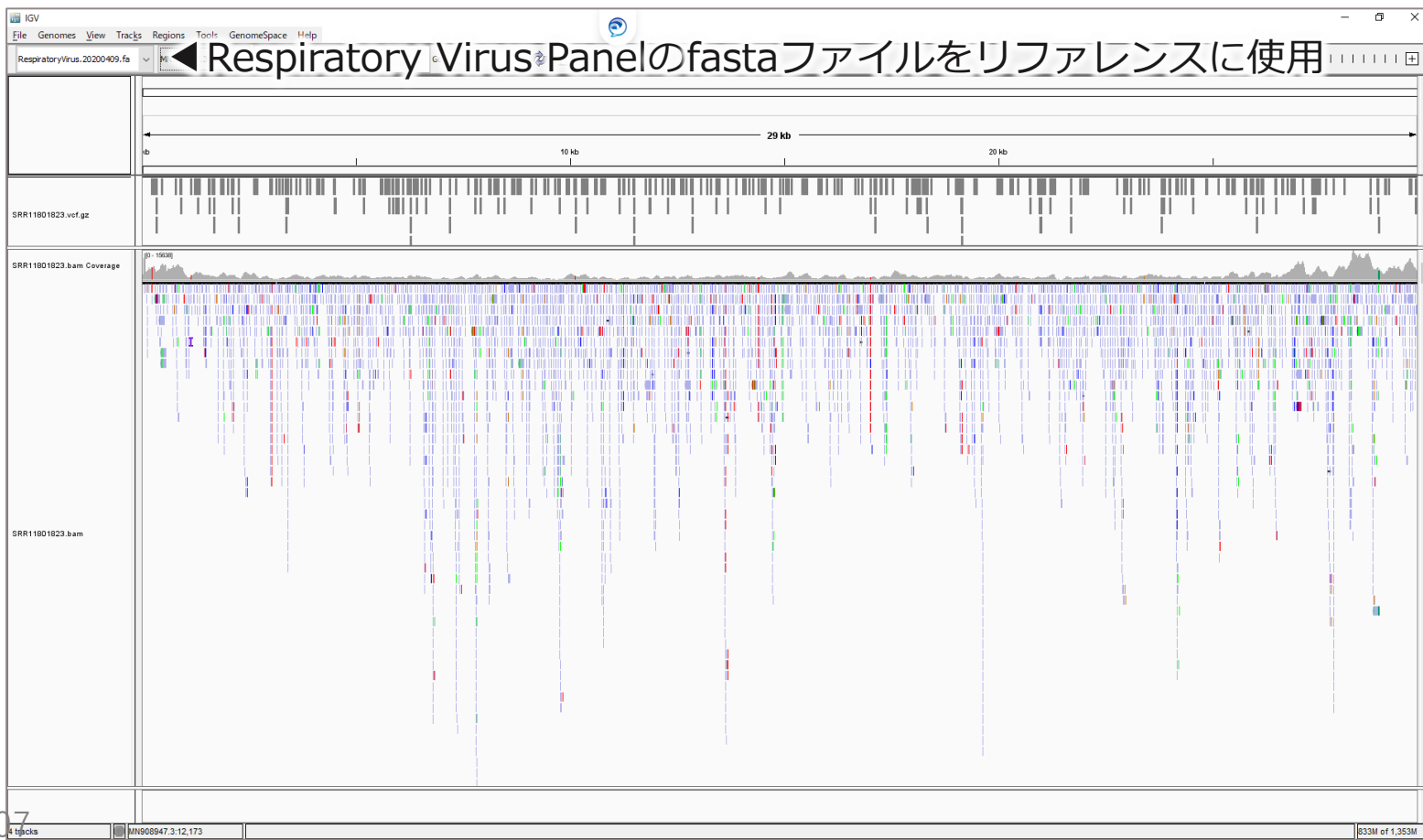
43.39%

インサート長解析▶

アライメントと変異コール

 SRR11801823.bam	2020-05-18 13:10	bam	SRR11801823.bam
---	------------------	-----	-----------------

 SRR11801823.vcf.gz	2020-05-18 13:10	gz	SRR11801823.vcf.gz
--	------------------	----	--------------------



DNA Amplicon app

アンプリコンワークフロー



注) 資料内テスト解析には仮のデータを使用したため、実際のパイプラインの実施時とは数値などが異なる場合がございます

プロダクトファイルの取得とインポート

<https://jp.illumina.com/products/by-brand/ampliseq/community-panels/sars-cov-2.html>

(現在BaseSpace上にプロダクトファイルはありません)

AmpliSeq for Illumina SARS-CoV-2 Research Panel

This sequencing panel offers a quick, accurate, and cost-effective method to research the SARS-CoV-2 virus. [Read More...](#)

製品の選択

Panel

0 AmpliSeq™ SARS-CoV-2 Research Panel for Illumina® ⓘ
20020496 [Sign in to see pricing and favorite products.](#)

製品に関する文献

- AmpliSeq for Illumina SARS-CoV-2 Panel Manifest File
product_file | ZIP< 1 MB
- AmpliSeq for Illumina SARS-CoV-2 Panel BED File
product_file | BED< 1 MB
- AmpliSeq for Illumina SARS-CoV-2 Custom Reference Genome File
product_file | ZIP< 1 MB

Upload Files

Select the type of files you want to upload



インプットフォーム

The screenshot shows the 'DNA Amplicon v2.1.1' configuration form. It includes sections for 'Analysis Name', 'Save Results To', 'Biosample(s)', 'Custom Genome (Optional)', 'Targeted Amplicons', and 'Custom Manifest File'. Red dashed boxes and arrows highlight specific fields: 'Analysis Name' (containing 'DNA Amplicon 05/18/2020 7:09:30'), 'Custom Genome (Optional)' (containing 'MN908947v3.fasta'), 'Targeted Amplicons' (containing 'Custom Panels (select manifest file below)'), and 'Custom Manifest File' (containing 'SARS-CoV-2.dna_manifest.20200408.txt').

Custom Genome

ウェブからダウンロードしたSARS-CoV-2パネル用のFASTAファイルを選択

Targeted Amplicons

カスタムパネルを選択

Custom Manifest File

ウェブからダウンロードしたSARS-CoV-2パネル用のマニフェストファイルを選択

アウトプット

SAMPLE INFORMATION ⁱ			
Total PF Reads	Percent Q30 Bases	Percent On-target Aligned Reads	Autosome Call Rate
1,025,784	95.16%	100%	0%

AMPLICON SUMMARY ⁱ	
Number of Amplicon Regions	Total Length of Target Regions
242	30,253 bp

Read Level Statistics ⁱ		
Read	Total Aligned Reads	Percent Aligned Reads
1	5,289	1.03%
2	5,275	1.03%

Base Level Statistics ⁱ				
Read	Percent Q30	Total Aligned Bases	Percent Aligned Bases	Mismatch Rate
1	96.08	690,747	0.89%	0.48%
2	94.23	689,647	0.89%	0.51%

◀リード情報

◀アライメントメトリクス

アウトプット

SNV, indelメトリクス
▼

SAMPLE INFORMATION ⁱ

Total PF Reads	Percent Q30 Bases	Percent
1,025,784	95.16%	100%

AMPLICON SUMMARY ⁱ

Number of Amplicon Regions
242

Read Level Statistics ⁱ

Read	Total Aligned Reads
1	5,289
2	5,275

Base Level Statistics ⁱ

Read	Percent Q30	Total Aligned Bases
1	96.08	690,747
2	94.23	689,647

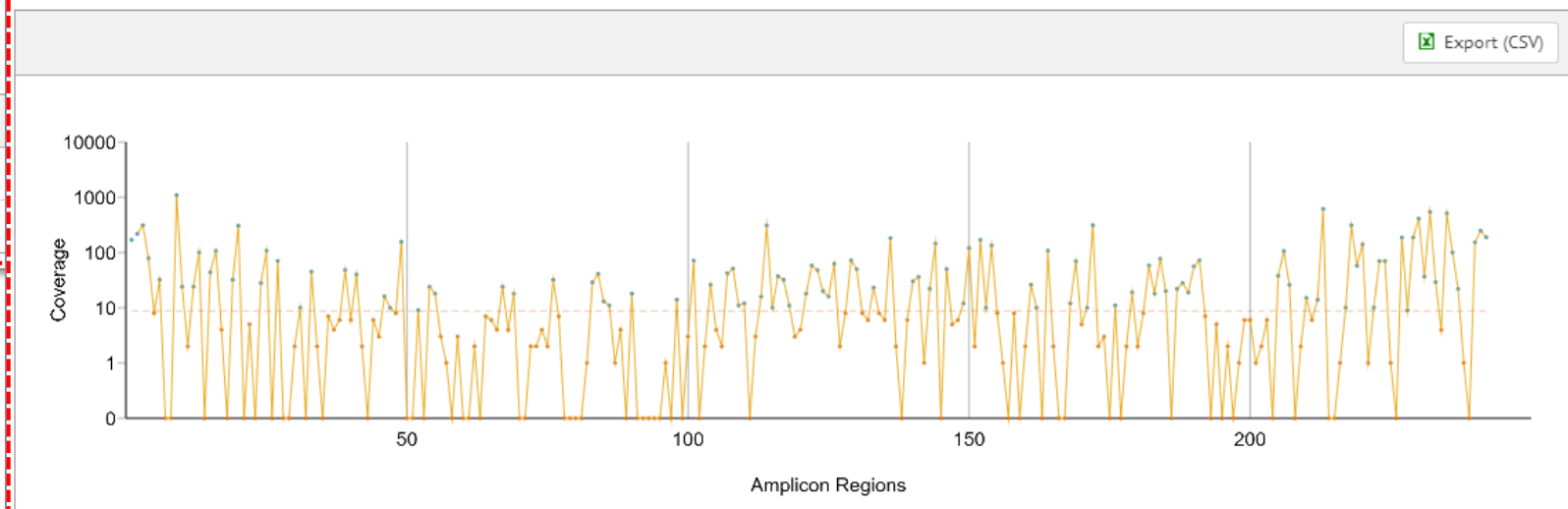
SMALL VARIANTS SUMMARY ⁱ

	SNVs	Insertions	Deletions
Total Passing	15	7	3
Het/Hom Ratio	2.75	6	-
Ts/Tv Ratio	0.667	-	-

COVERAGE SUMMARY ⁱ

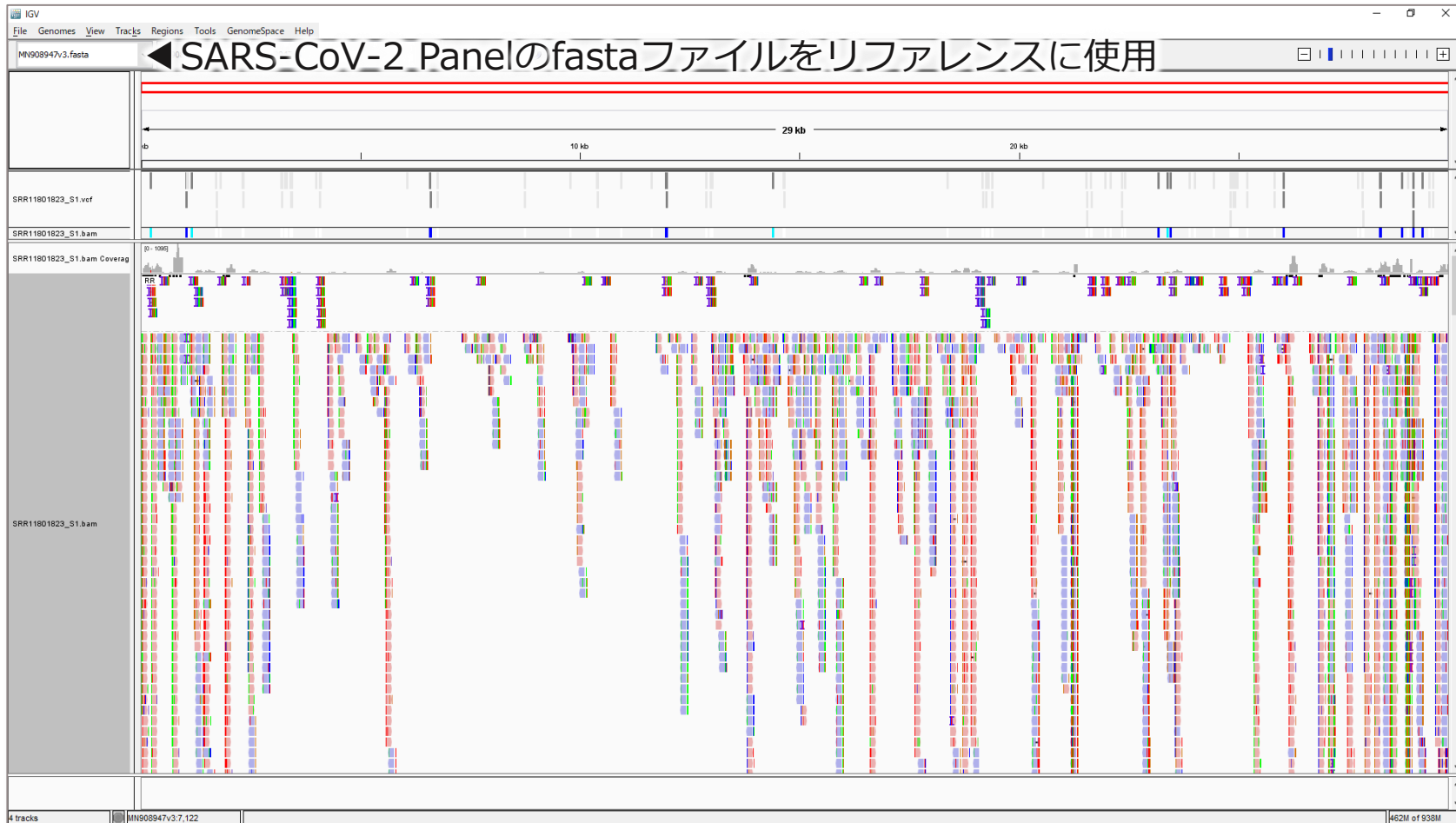
Amplicon Mean Coverage	Uniformity of Coverage (Pct > 0.2*mean)
43.7	45.87

Coverage by Amplicon Region ⁱ



アウトプット

 SRR11801823_S1.bam	2020-05-17 18:59	bam	SRR11801823_S1.bam
 SRR11801823_S1.vcf.gz	2020-05-17 18:59	gz	SRR11801823_S1.vcf.gz



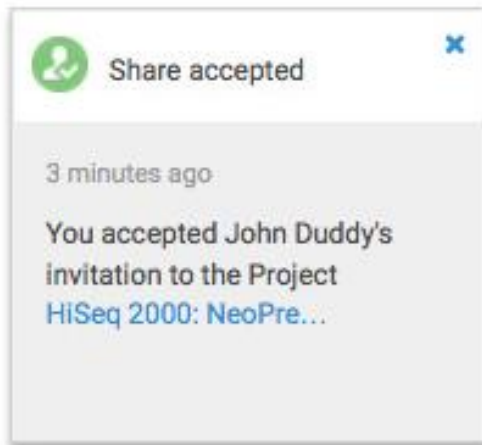
Collaboration Tools



BaseSpaceで行うデータシェア

Easy Sharing & Collaboration

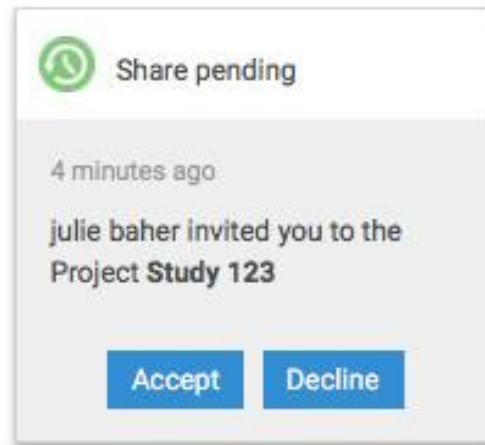
Push-button submission to public infectious disease databases



Share accepted

3 minutes ago

You accepted John Duddy's invitation to the Project [HiSeq 2000: NeoPre...](#)



Share pending

4 minutes ago

julie baher invited you to the Project **Study 123**

[Accept](#) [Decline](#)




SRA Import App



new GISAID Submission App

SRAインポートアプリ



SRA Import v0.0.5
BaseSpace Labs

Configuration

Analysis Name
SRA Import 05/07/2020 10:34:41

SRA Accession ⓘ
SRX7857604

Save Results To
Select Project
Covid19_analysis

BaseSpace Labs Disclaimer
BaseSpace Labs
 I acknowledge and agree that (i) this is a BaseSpace Labs App, (ii) I am using it AS-IS without any warranty of any kind, (iii) Illumina has no obligation to provide any technical support for this App, and (iv) Illumina has no liability for my use of this App, including without limitation, any loss of data, incorrect results, or any costs, liabilities, or damages that result from use of this App.

Launch Application

← SRAのアクセッション番号を入れる

SRX7857604: RNA-Seq of Severe acute respiratory syndrome coronavirus 2
1 ILLUMINA (Illumina MiSeq) run: 61,329 spots, 11.1M bases, 2.5Mb downloads

Design: random hexamer first and second strand synthesis followed by Nextera XT
Submitted by: UNIVERSITY OF WASHINGTON
Study: Washington SARS-CoV-2 isolate sequences
[PRJNA610428](#) • [SRP251618](#) • [All experiments](#) • [All runs](#)
[show Abstract](#)

Sample:
[SAMN14308028](#) • [SRS6266706](#) • [All experiments](#) • [All runs](#)
Organism: [Severe acute respiratory syndrome coronavirus 2](#)

Library:
Name: WAB_UW5
Instrument: Illumina MiSeq
Strategy: RNA-Seq
Source: VIRAL RNA
Selection: unspecified
Layout: SINGLE

Runs: 1 run, 61,329 spots, 11.1M bases, [2.5Mb](#)

Run	# of Spots	# of Bases	Size	Published
SRR11247076	61,329	11.1M	2.5Mb	2020-03-05

ID: 10283692

GISAIDサブミッションアプリ

Lab



Automated run upload

BaseSpace Sequence Hub



DRAGEN RNA
Pathogen Detection

Create
consensus FASTAs



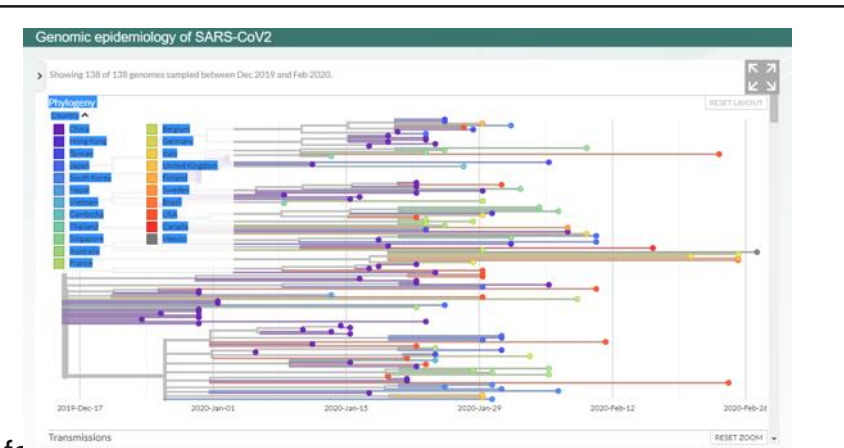
GISAID
Submission App

Collect metadata:
location, etc.

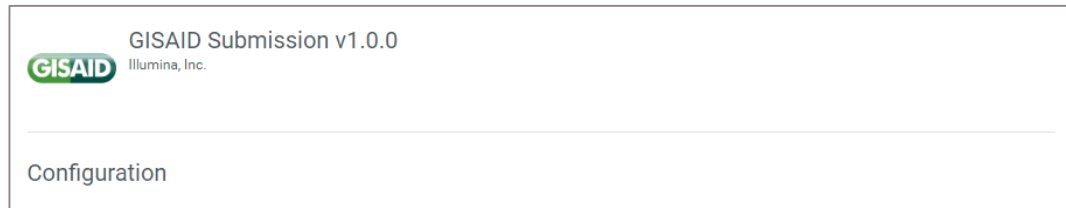
Create output files
in desired formats

Automated data
upload

GISAID



GISAIDサブミッションアプリ



ご使用にはGISAIDアカウントの取得が必要です

The image shows the main configuration area of the GISAID Submission v1.0.0 application. It contains several sections, each highlighted with a red dashed box and connected to explanatory text by a red dashed arrow. The sections are: "Metadata file (CSV)", "Choose Source of Consensus FASTAs", "Consensus sequences for upload (FASTA)", "Variant files (VCF)", and "Reference genome that was used to call variants (hCoV-19/Wuhan/WIV04/2019|EPI_ISL_402124 recommended)".

Metadata file

CSV形式のメタデータを指定

Choose Source of Consensus FASTAs

既存のコンセンサスFastaか、リファレンスとvcfを指定してコンセンサスFastaを作成するか

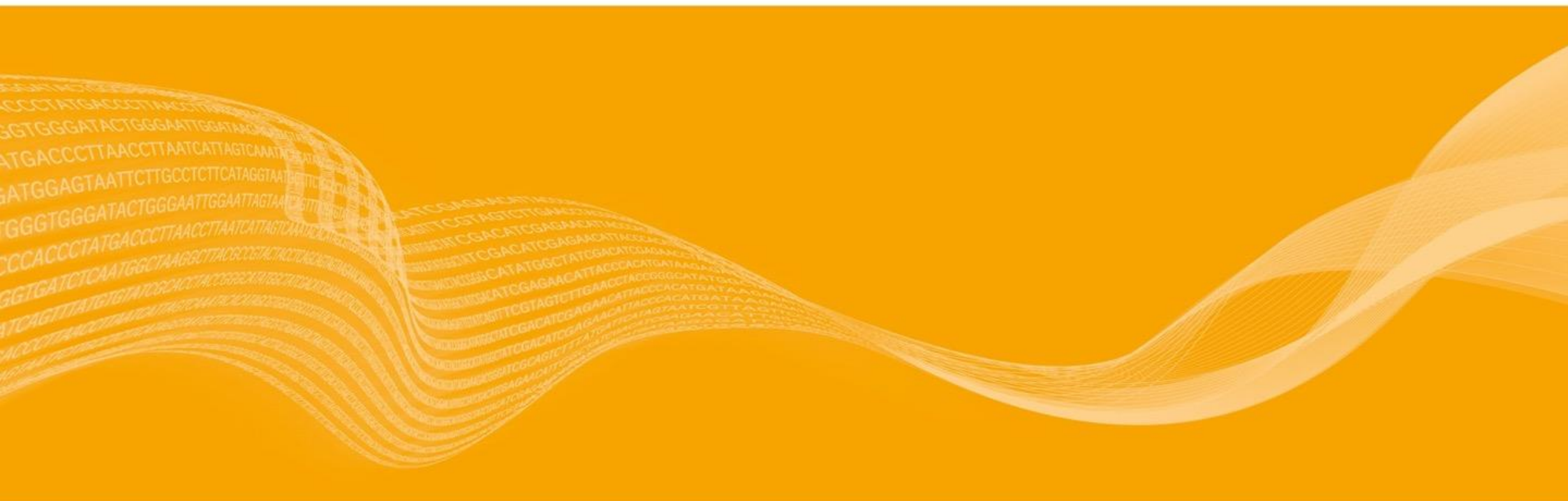
Consensus sequences for upload

既存のFastaを指定する際に使用

Variant files, Reference genome

変異情報を指定

4. 本日のまとめ



本日のまとめ

- イルミナ社はSARS-CoV-2を含めた、様々なウイルスゲノムシーケンスに対応したライブラリ調製キット、シーケンサー、解析ツールをご提供しています
- TruSeq Stranded Total RNAによるShotgun Metagenomics法
- 新製品のRespiratory Virus Oligo Panel、AmpliSeq for Illumina SARS-CoV-2 Research Panelを用いてターゲットシーケンスを手軽に実現
- BaseSpaceには、検出、同定、変異解析に使用できるアプリがある
- 公共データベースとの共有もアシストするサービスも提供

ご清聴ありがとうございました。

