2015年9月4日
イルミナ サポートウェビナー

解析に適したリード前処理を行うために

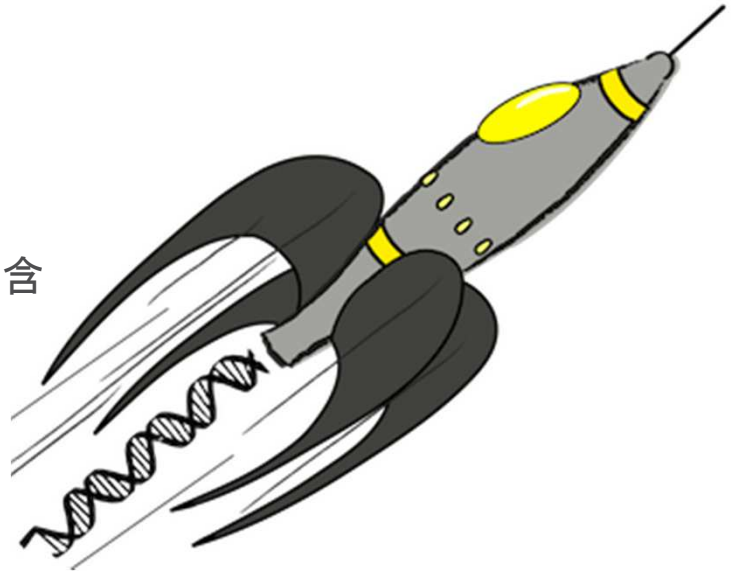
イルミナ株式会社
バイオインフォマティクス
サポートサイエンティスト
癸生川絵里 (Eri Kibukawa)

※BaseSpaceアプリ: FASTQ toolkit /smallRNA/ FASTQC

illumina®

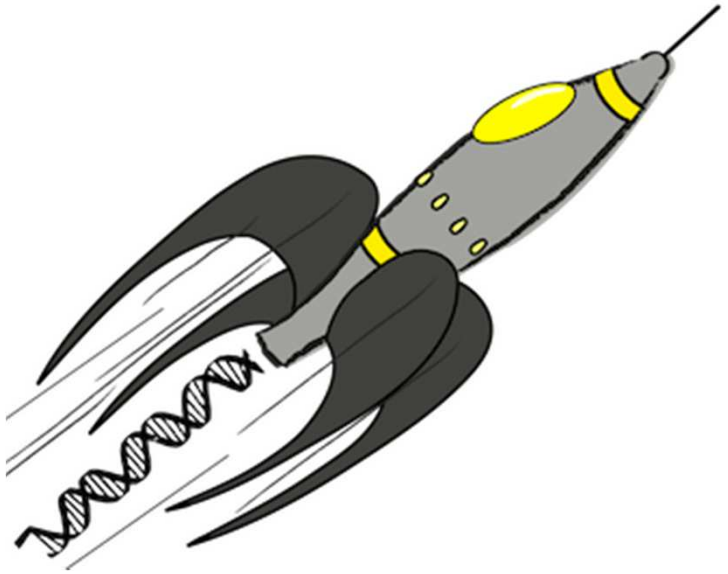
本日の内容

- イントロダクション
- アダプタートリミング
  ※smallRNA 例含
- クオリティトリミング
- ダウンサンプリング
- リードの結合
- 手元のFASTQをトリミングするには



本日の内容

- イントロダクション
- アダプタートリミング
- クオリティトリミング
- ダウンサンプリング
- リードの結合
- 手元のFASTQをトリミングするには



装置からの解析フロー

MCS/HCS/NCS/RTA
ソフトウェア
on ControlPC

画像取得 /シグナル抽出

↓

ベースコール (*.bcl)

↓

*.bcl ファイ… …変換

Fastq

BaseSpace BaseSpace ONSITE

bcl2fastq (Linux)

MiSeq Reporter (Win)

シーケンシング後

二次解析以降

ワークフロー化

再解析、その他下流解析、可視化、
アノテーション、フィルタリング
レポート生成など

illumina

FASTQフォーマット

illumina®

FASTQの生成場所・方法

MiSeq Reporter

MiSeqに内蔵されている.
64bit Win に別途インストール
も可能

BaseSpace®

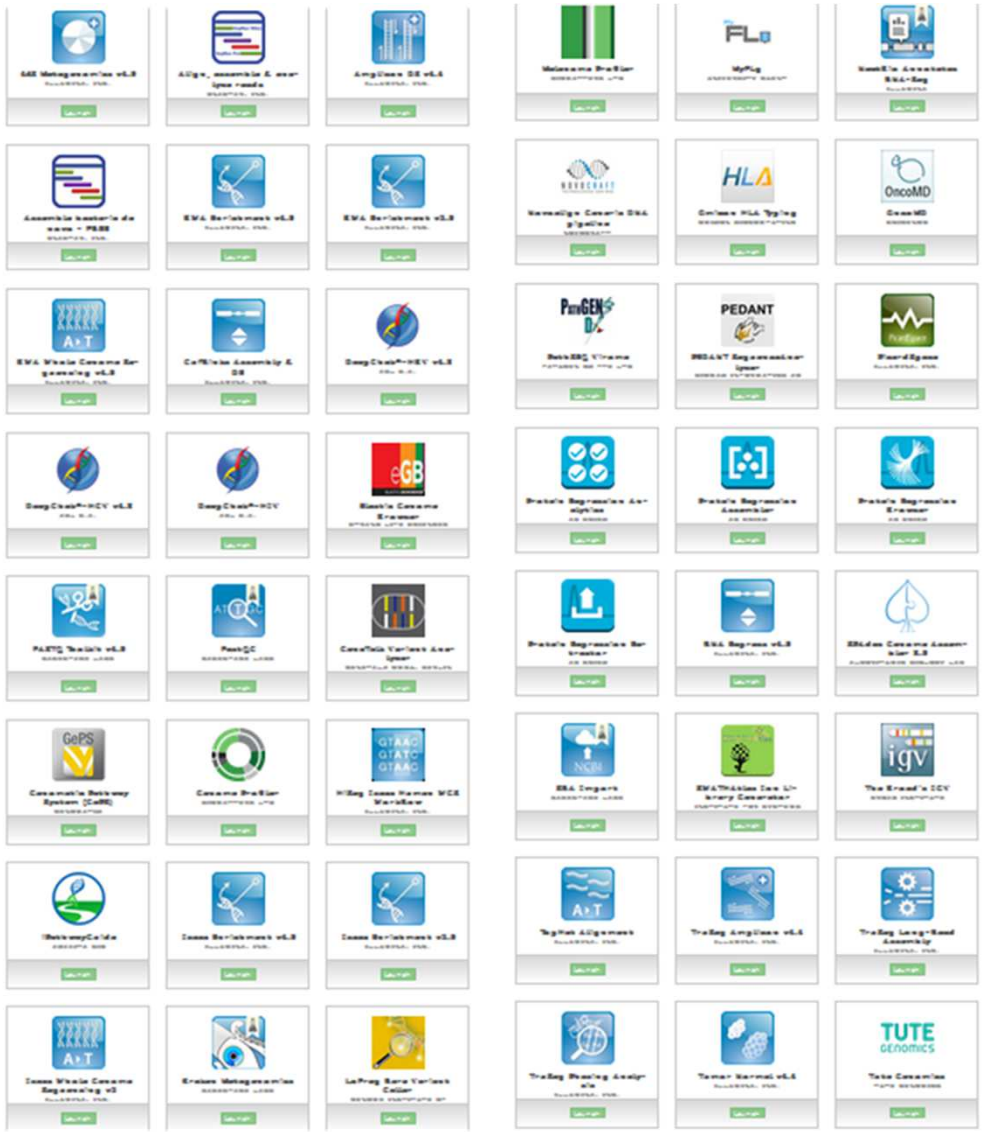
BaseSpace® ONSITE

お使いのLinux server

bcl2fastq2

illumina®

アプリ (>60)



Categories

Resequencing	Small RNA
Targeted Sequencing	De Novo Assembly
RNA-Seq	Gene Fusion Detection
ChIP-Seq	Methyl-Seq
Metagenomics	Tumor Normal
Variant Analysis	Differential Expression
Quality	Proteomics
Synthetic Long Reads	Visualization

illumina

＜イルミナコアアプリ＞


16S Metagenomics


TopHat Alignment


Cufflinks Assembly & DE


RNA Express


Variant Studio


BWA Enrichment


Isaac Enrichment


Broad IGV


TruSeq Amplicon


Amplicon DS


BWA WGS


Isaac WGS


Tumor Normal


Long Read Assembly


Long Read Phasing

＜イルミナラボアプリ＞

他


FastQC


Kraken Metagenomics


NextBio Annotates


VCAT


NextBio Transporter


SRST2


Fastq Toolkit


Velvet Assembly


Picard Space


SRA Importer


Prokka

他

＜他社製アプリ＞


SPAdes


Novo Align


Advaita


DNA Star


AB SCIEX


AB SCIEX


AB SCIEX


AB SCIEX


SWATH Atlas


MetaPhlAn


n of One


My FLQ


Lo Feq


eGB

Genomatix

Genome Profiler

OncoMD

GeneTalk

PathGEN Dx

Melanoma Profiler

TUTE

DeepCheck HIV,HBV,HCV

Pedant

他

illumina

BaseSpace Labsアプリ（準サポート)

- 人気の機能をイルミナで素早くラップ/開発したツールをご提供.
- 一方、テストやドキュメント作成は低減
- テクニカルサポートの正式サポート対象ではなく、開発者へダイレクトにお問合せ戴けるご提供形態のアプリ(basespacelabs@illumina.com).

FASTQ Toolkit

- Sub-sample reads
- Trim Adapters
- Trim Bases
- Ploy A/T trimming
- Quality Trimming
- Read Filtering
- Reverse Complement

FastQC

- Perform QC of raw sequencing data.
- Determine adapter contamination

VCAT v2.3

- Compare Variant Call Sets to standards
- Intersect variant call sets.

SRA Import v0.0.3

- Import up to 25GB of sequencing data from SRA

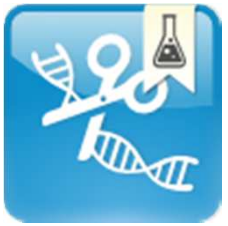
SRA Submission v0.0.3

- Deposit sequencing data in SRA.

他

illumina

FASTQ Toolkit（FASTQツールキット）



Adapter trimming（アダプタートリミング)

5’-また3’-それぞれ別にトリミングしたいアダプター配列を指定できる

Base trimming（ベーストリミング）

5‘- あるいは 3’-端から、指定長分の塩基をトリミングすることができる

Quality trimming（クオリティートリミング）

3’-端の低クオリティー配列をトリミングする用途向け．Qscore平均閾値を指定

Poly-A/T trimming（Poly-A/T トリミング）

リード終端のPoly-A/T をトリム.

Sub-sampling (サブサンプリング、またはダウンサンプリングとも呼称)

サンプルリードの一部を取り出し、より少ないサンプルリードセットをつくる

FASTQ Toolkit（FASTQツールキット）

Read filtering (リードフィルタリング)

最短/最長 塩基数や最大/最小 平均クオリティー値、最大/最小 GC含有率、
低複雑度領域などの条件を指定し指定閾値外のリードを除外

Modify reads (旧 Reverse complement)

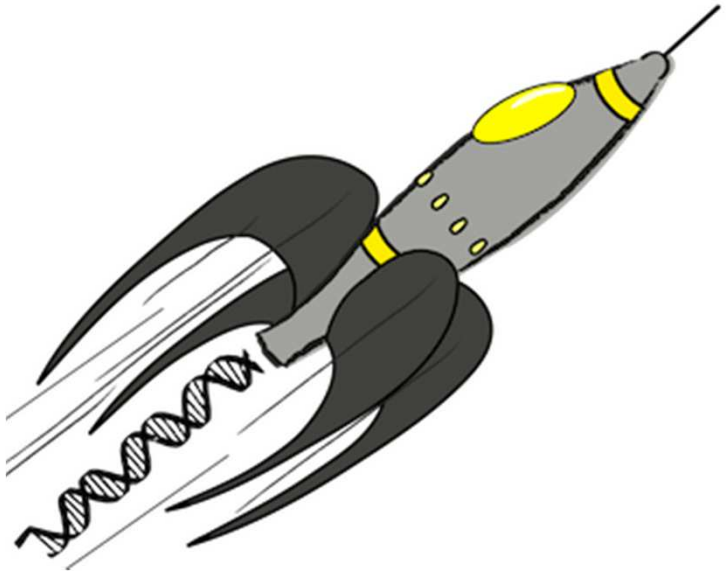
相補鎖配列取得
(Nexteraメイトペアリードからペアードエンドリード 方向への変換目的など）に加え、
他ペアードエンドリードが１つのFASTQからR1, R2への振り分け

Fix formats (フォーマット修正)

アップロードした FASTQヘッダやエンコード(Qscoreのオフセット値) 修正、
ファイル名などが規約を満たしていない事によりBaseSpaceアプリが受け付けない場合に
修正を試みるなど可能

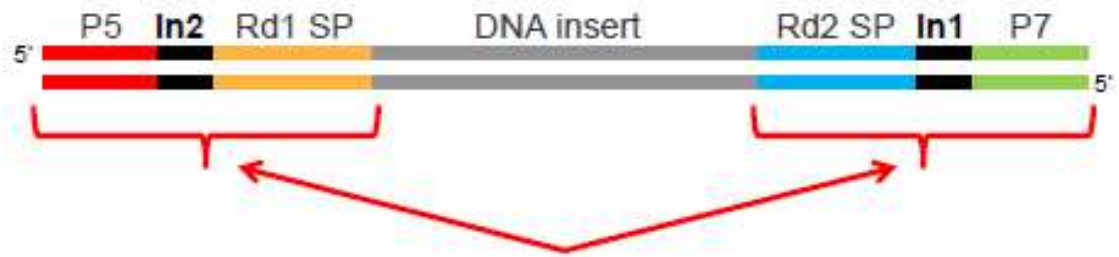
本日の内容

- イントロダクション
- アダプタートリミング
- クオリティトリミング
- ダウンサンプリング
- リードの結合
- 手元のFASTQをトリミングするには



アダプターとは

イルミナ ライブラリの構造



イルミナシーケンサー専用オリゴヌクレオチドアダプター

DNA インサート： 数百bpに断片化したDNA. 読みたい目的サンプル配列.
P5, P7 ： フローセルへの結合部位
SP ： シーケンシングプライマー結合部位
In (Index) ：複数サンプル同時解析用のバーコード（目印配列）

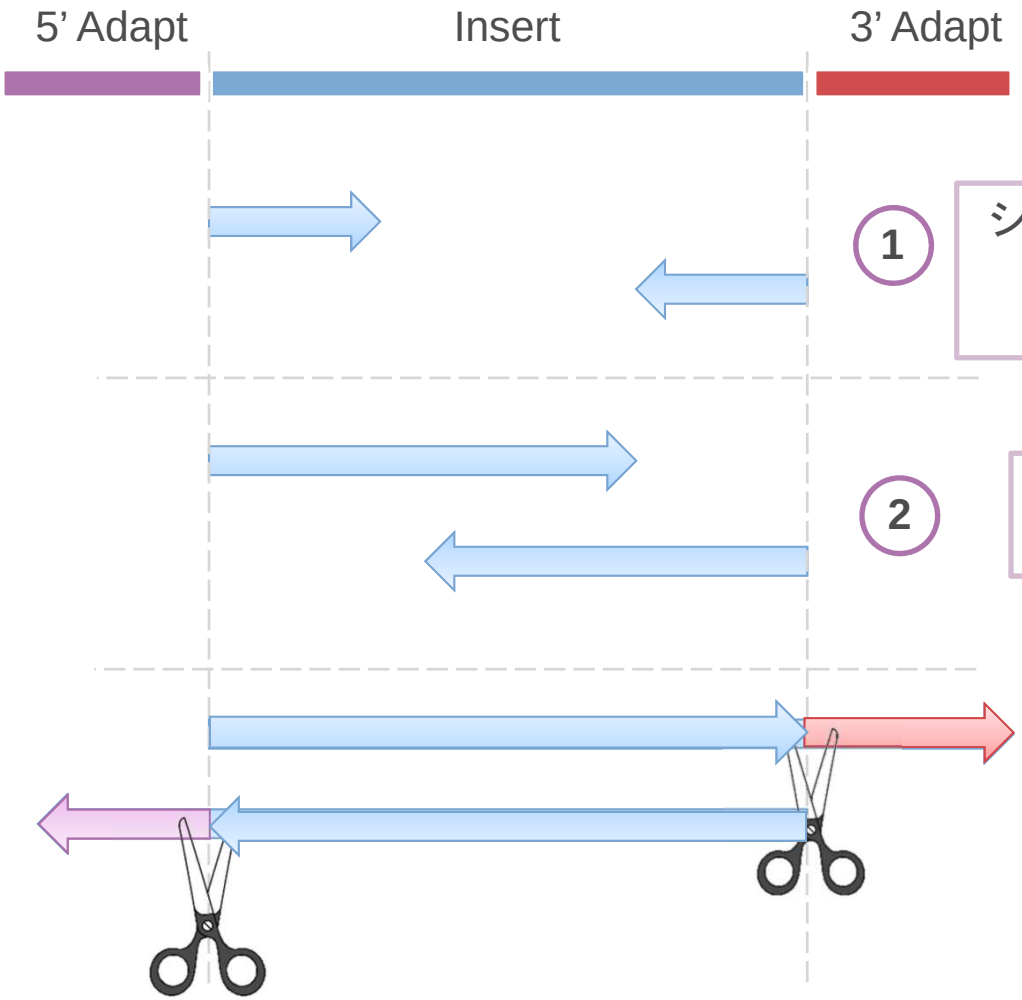
ライブラリ = DNA インサート ＋ 両端にそれぞれ別のアダプター

イルミナシーケンサーでシーケンスするため、この構造をとるようにサンプル調整する

※ 詳しくは、弊社サポートウェビナー 2015/07/10 をご参考いただけます。
SBS (Sequencing By Synthesis) ケミストリーとは何か？
http://www.illuminakk.co.jp/events/webinar_japan/support_webinar.ilmn

インサート長とアダプタートリミング

アダプターとインサート配列からなるライブラリに対する、
実際シーケンスしてリードとして得られる配列の位置関係のパターン



1 シーケンスするリード(青矢印) がインサートよりも短い場合（通常はこのパターン）

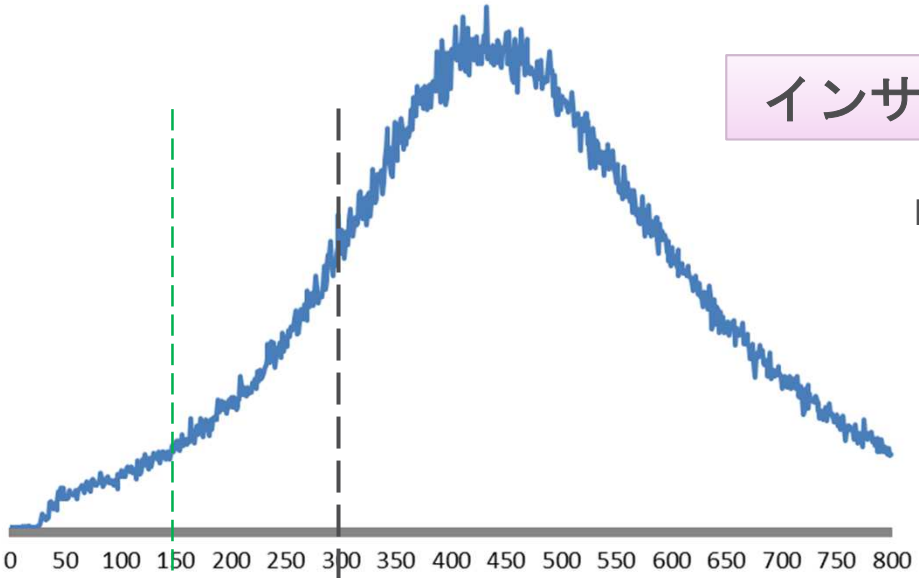
2 シーケンスするリードをオーバーラップさせた場合

3 シーケンスするリード長がインサート長 よりも長い例.

アダプター配列にまで読み超している

illumina®

インサート長の分布とアダプタートリミング



インサート分布の例

NexteraXT MiSeq v3

例) リード長150bp
シーケンスした場合

150bp

リード長150bp がインサート長 よりも長い状態である
1.8% が要トリミング

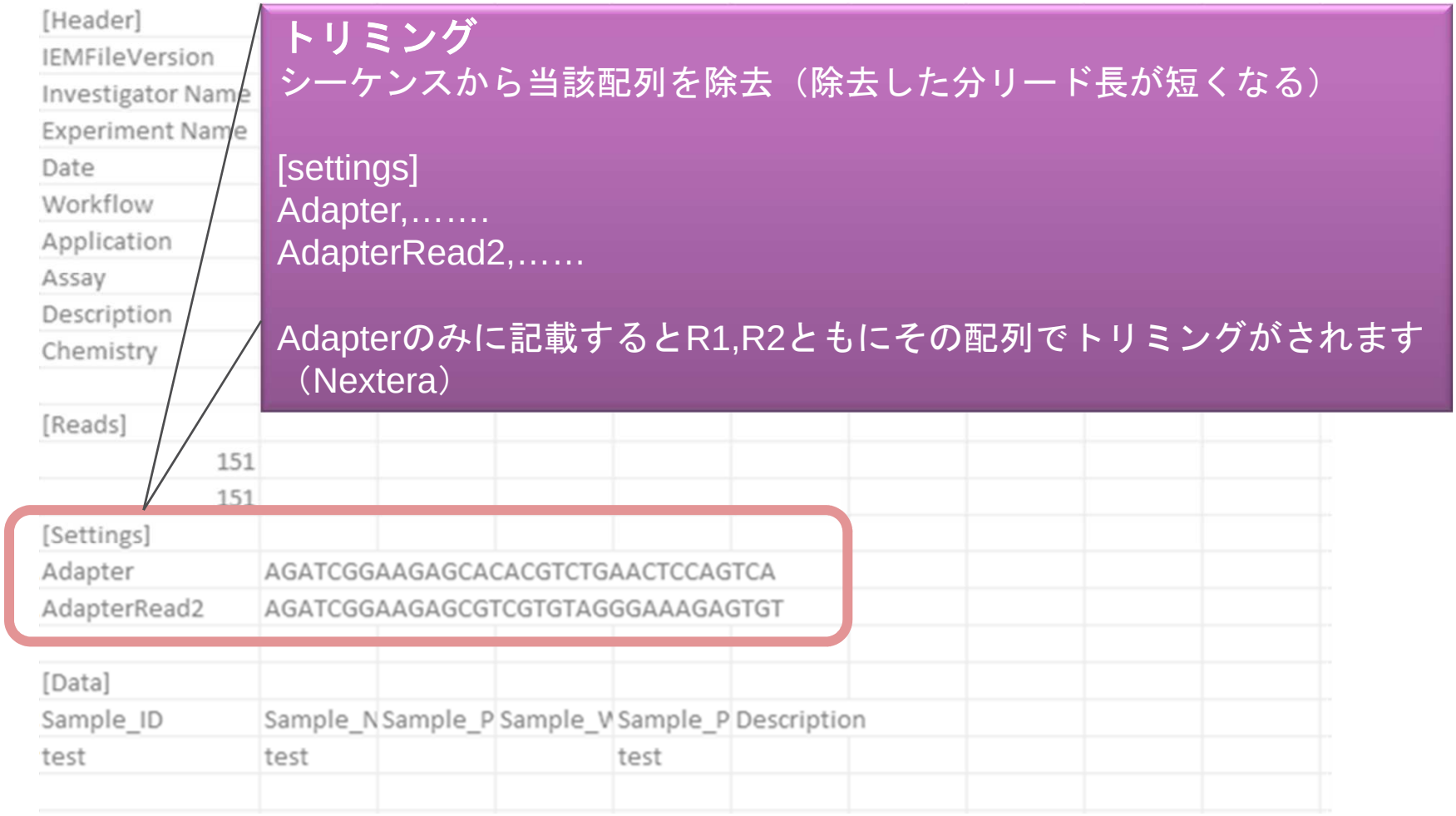
例) リード長300bpで
シーケンスした場合

300bp

リード長300bp がインサート長 よりも長い状態である
16.7% が要トリミング

illumina®

アダプタートリミングの方法
Adapter, AdapterRead2



illumina®

アダプタートリミングの例

アダプター配列
マッチ > 90%
(デフォルト)

ビフォー

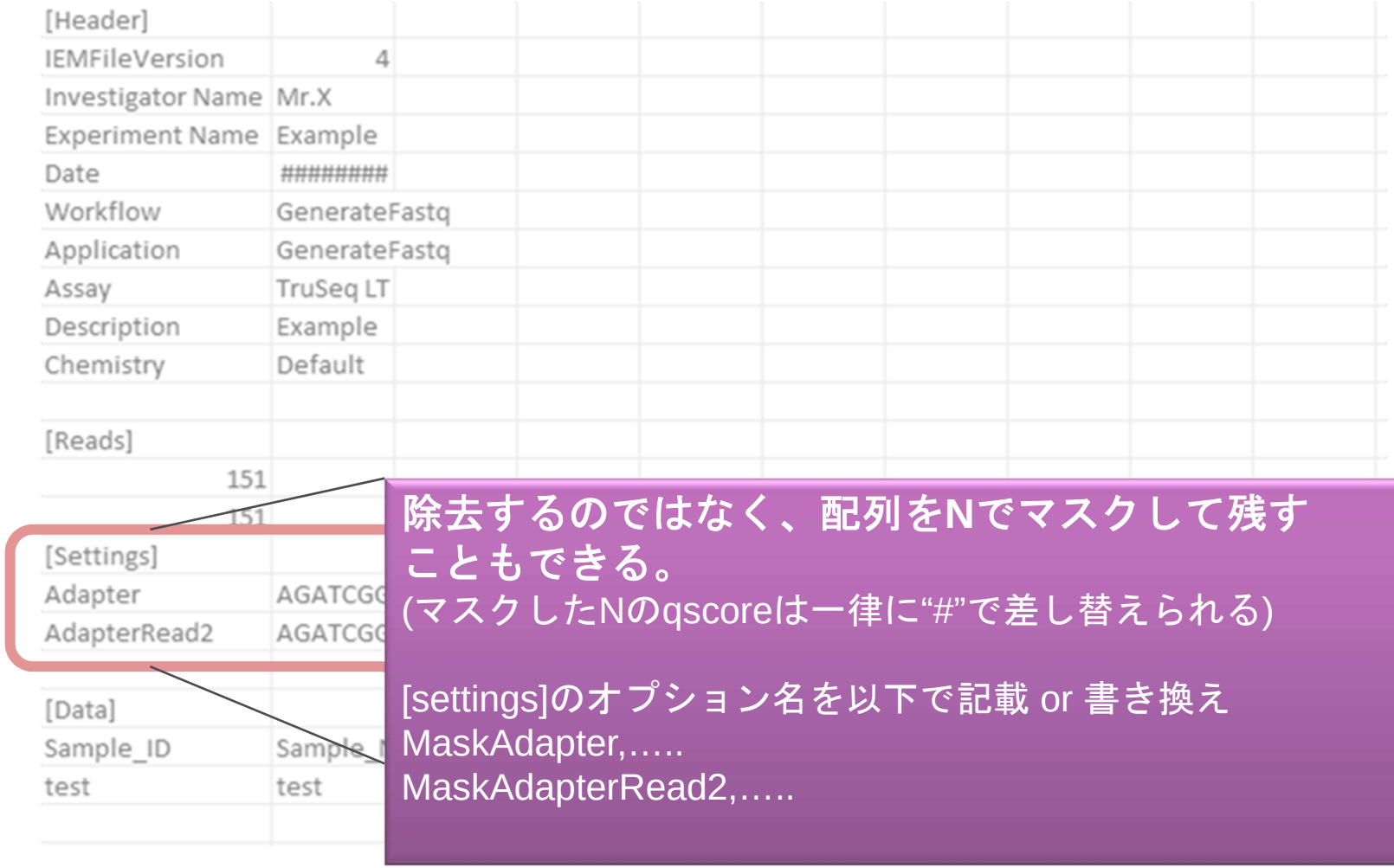
```
@M00000:71:000000000-D00LW:1:1101:16265:1658 1:N:0:1
ACTCTGCGTTGCGCTTCTGCTCGGCCTCCAGCTCACCCTCCCTGTCTCTTATACACATCTCCGAGCCCA
+
BCCCCFFCCBCCGGGGGGGGGGGGGGGHHHHHHHHHHHHHGGHHHHHHHHHHHHHHHHHHHHGGGGGGH
```

アフター

```
@M00000:71:000000000-D00LW:1:1101:16265:1658 1:N:0:1
ACTCTGCGTTGCGCTTCTGCTCGGCCTCCAGCTCACCCTCC
+
BCCCCFFCCBCCGGGGGGGGGGGGGGGHHHHHHHHHHHHG
```

当該アダプター配列の初頭から以降がトリムされる

アダプターマスキング
MaskAdapter, MaskAdapterRead2



[Header]	
IEMFileVersion	4
Investigator Name	Mr.X
Experiment Name	Example
Date	########
Workflow	GenerateFastq
Application	GenerateFastq
Assay	TruSeq LT
Description	Example
Chemistry	Default
[Reads]	
151	
151	
[Settings]	
Adapter	AGATCGG
AdapterRead2	AGATCGG
[Data]	
Sample_ID	Sample_
test	test

除去するのではなく、配列をNでマスクして残すこともできる。
(マスクしたNのqscoreは一律に“#”で差し替えられる)

[settings]のオプション名を以下で記載 or 書き換え
MaskAdapter,…..
MaskAdapterRead2,…..

※MiSeq Reporter、BaseSpace、bcl2fastq2等 利用時のサンプルシート設定

illumina®

アダプターマスキングで実行した例

ビフォー

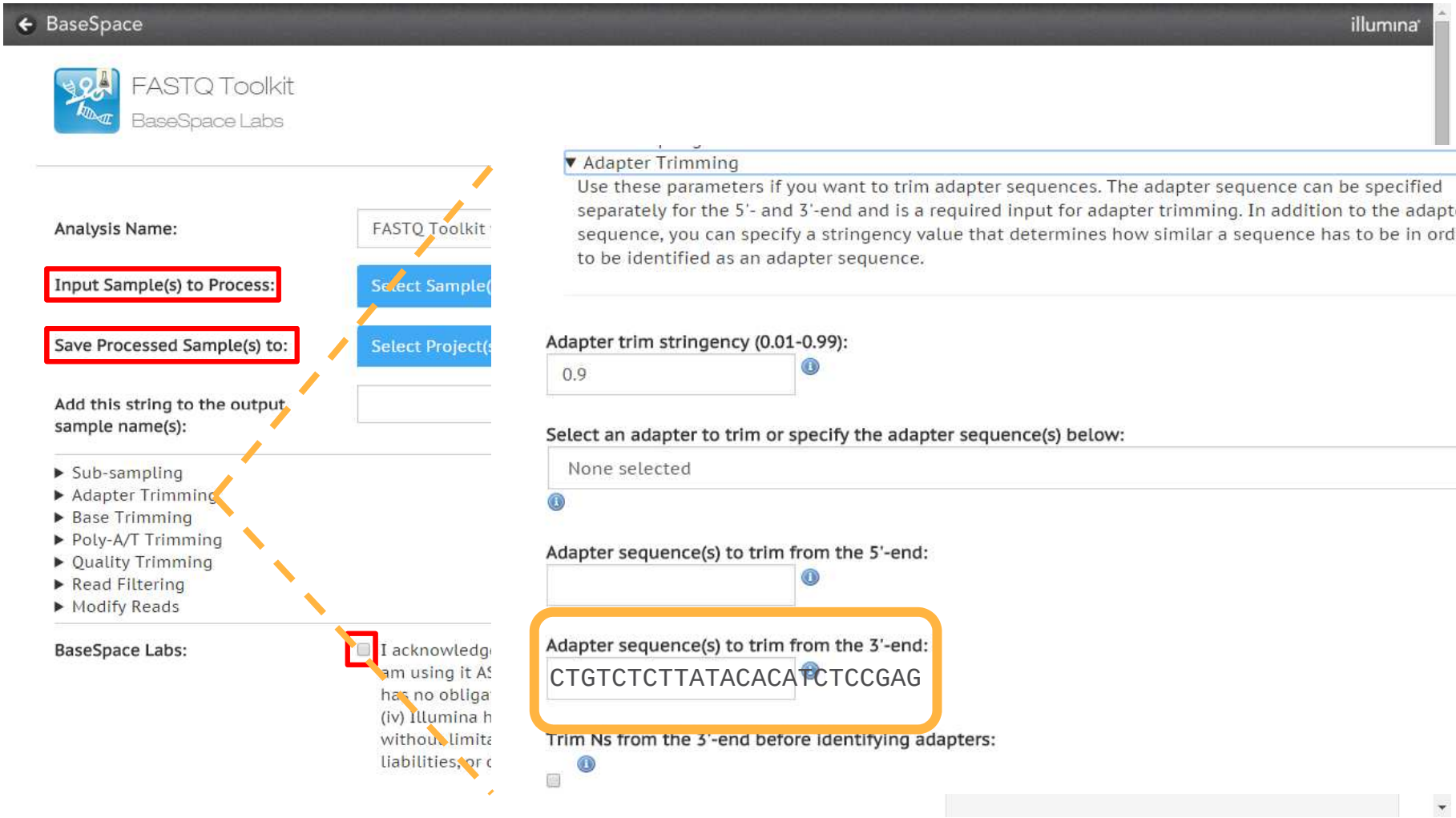
```
@M00000:71:000000000-D00LW:1:1101:16265:1658 1:N:0:1
ACTCTGCGTTGCGCTTCTGCTCGGCCTCCAGCTCACCCTCCCTGTCTCTTATACACATCTCCGAGCCCA
+
BCCCCFFCCBCCGGGGGGGGGGGGGGGHHHHHHHHHHHHHGGHHHHHHHHHHHHHHHHHHHHGGGGGGH
```

アフター

```
@M00000:72:000000000-D00LW:1:1101:16265:1658 1:N:0:1
ACTCTGCGTTGCGCTTCTGCTCGGCCTCCAGCTCACCCTCCNNNNNNNNNNNNNNNNNNNNNNNNNNNN
+
BCCCCFFCCBCCGGGGGGGGGGGGGGGHHHHHHHHHHHHHG##############################
```

アダプター配列を含むアダプター配列以降の塩基をNでマスクし、
クオリティースコアは一律２(#)で置換

BaseSpaceでトリミング目的に使えるツール
FASTQ Toolkit



その他アダプタートリミングに使える3rd-partyツールの一例

ツール名	配布場所
Trimmomatic	http://www.usadellab.org/cms/?page=trimmomatic
FASTX toolkit	http://hannonlab.cshl.edu/fastx_toolkit/ (FastQ clipper)
Seq-Prep	https://github.com/jstjohn/SeqPrep
Cut-Adapt	https://code.google.com/p/cutadapt/
PEAT	https://github.com/jhhung/PEAT アダプター配列そのものを指定せずにトリミングができる (PEの重なりから判別するため、PE必須)

参考：http://omictools.com/adapter-trimming-c402-p1.html

なぜアダプター配列トリムを検討するのか?

BWA Enrichment V2.1

1

アライメントできるリード量が増える場合がある

BWA (backtrace)

Sample	Sample Name	Total Aligned Reads	Percent Aligned Reads
1	NA12892	354,882	77.4%
2	NA12892-trim	450,007	98.2%

ただし: 使用しているアライナープログラムによる

BWA (mem)

Sample	Sample Name	Total Aligned Reads	Percent Aligned Reads
1	NA12892	457,611	99.8%
2	NA12892-trim	457,542	99.8%

なぜアダプター配列トリムを検討するのか?

2

例えばアセンブル結果の向上

Assembly metrics	Before adapter trimming	After adapter trimming
N50	21	29,791
Maximum contig	553	174,326
Assembly length	18,497,207	4,876,437
Number of contigs	1,387,508	1,115

2 x 250bp, E.coli (Nextera XT)

illumina

なぜアダプター配列トリムを検討するのか?

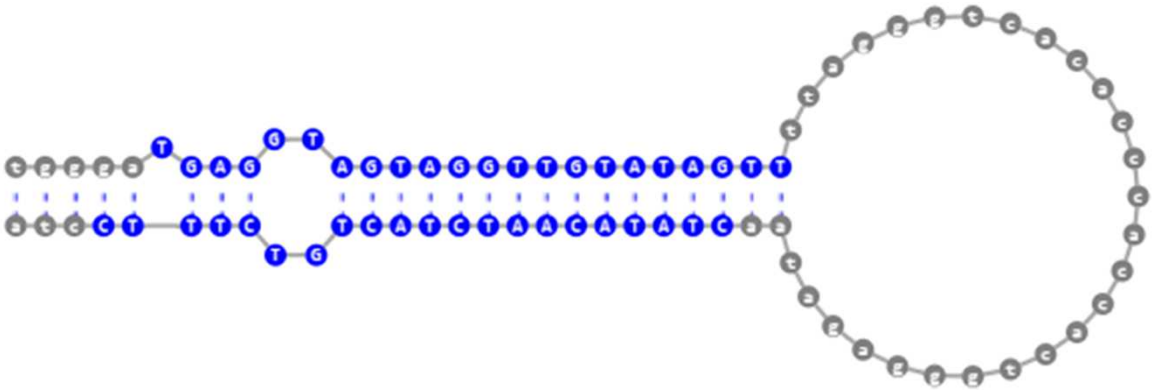
Small RNA v1.0

3

Small RNA のワークフローで必要となる

smallRNA解析では通常非常に短い配列を対象とするため、
シーケンシングのリード長の方が、smallRNAのインサート長よりも、短くなる。
したがって、アダプタートリミングが定常処理として必要となってくる 。

(例 ヒト miRNAだと例えば分布ピークが 22bpなど)



illumina

アダプタートリミングが必用となる例：

Small RNA 解析

illumina®

Small RNA のワークフロー
MiSeqの場合

内蔵のMiSeq Reporterが
自動トリム

1. Illumina Experiment ManagerウィジェットでSampleSheetを
作成する際、“smallRNA”ワークフローを選択する。
シーケンシングを開始する。

2. 生成されたFASTQファイルは自動でアダプタートリム済みとなる。
明示的にサンプルシートには記載なくとも**デフォルトでトリムが適用されている**。
TruSeq small RNA adapter (TGGAATTCTCGGGTGCCAAGG)
他のキットを使用している場合は明示的にサンプルシートに記載すれば適用される。

3. MiSeq ReporterではsmallRNAのワークフローによりレポート生成まで自動実行される。
途中で出力されたFASTQは、アダプタートリム済みのため、
BaseSpaceにアップロードするなどしてさらに後続の解析にそのまま使う事が可能。

BaseSpace Small RNA v1.0 アプリ

※ アダプタートリム済みのリードが必用

Small RNA v1.0

対応装置データ

- **HiSeq 2500/3000/4000**
- **NextSeq 500**
- MiSeq

対応ライブラリ調整キット

- TruSeq Small RNA

対応ゲノム

- Human HG19
- Mus musculus
- Rattus norvegicus

機能

- Alignment
- Classification of miRNAs, isomiRs, and piRNAs
- Novel miRNA discovery
- miRNA Precursor discovery
- Differential Expression of miRNAs, precursor groups, miRNA families, and piRNAs

内包ソフトウェアバージョン

- Isis (Analysis Software)—2.5.52.11
- Samtools 0.1.19-isis-1.0.2
- Bowtie (Aligner) 0.12.8
- miRDeep* 3.2
- DESeq2 1.0.17

Small RNAのワークフロー (GenerateFastq)
HiSeq/ NextSeq の場合

1. **smallRNAは装置からBaseSpace直アップロードの際は、留意が必要※**
   アダプター配列を自動トリムされないようにする必要がある
   サンプルシートはGenerateFASTQを指定、かつアダプタを記入しない など(HiSeq)

2. FASTQ Toolkit アプリなどでアダプタートリムを行っておく

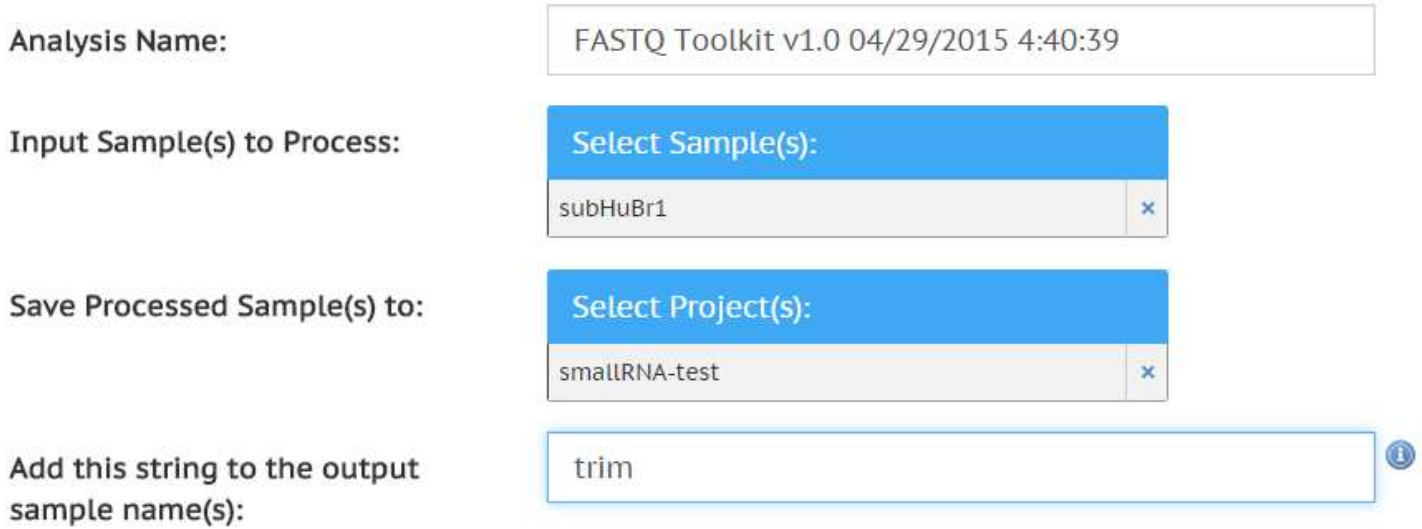
3. トリム済みのFASTQをsmallRNA v1.0アプリの入力に供する

※ BaseSpaceにおいてGenerateFastq でアダプタートリムの指定を行うと32 bp よりも短い配列は一律に Nでマスクされるため 。

Small RNAのリードを Fastq toolkitでトリムする

1 ProjectエリアのLaunch appボタンなどから “FASTQ Toolkit ” アプリを起動

2 Select Samples で入力サンプル(= fastq)を選択し “Add a string to the output sample name(s)”にファイル名に別名を付けるための文字列を入力

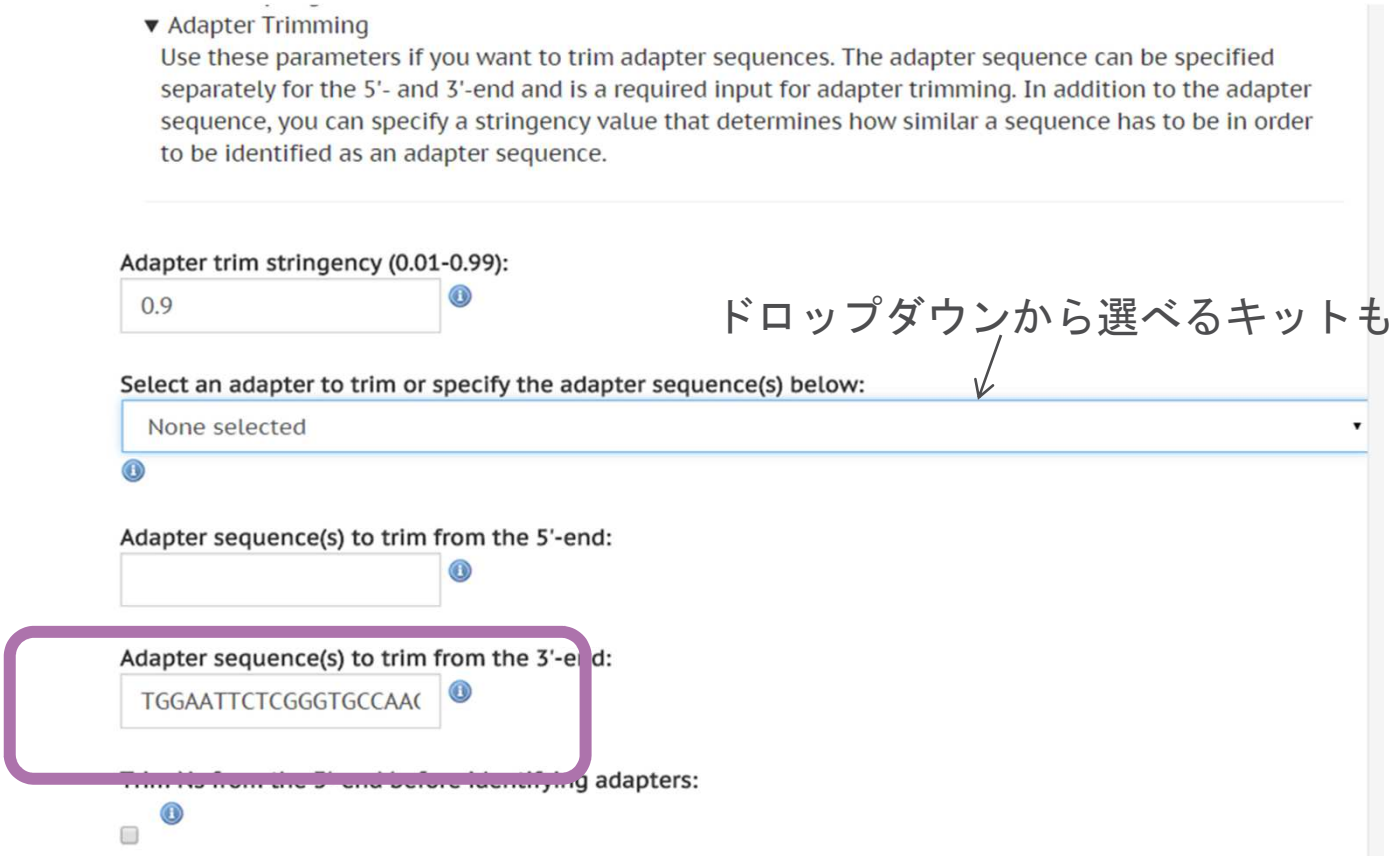


例: 上記のようにtrimを入れておくと、トリム後のサンプル名(fastqファイル名)が “subHuBr1trim”となる. オリジナルとの区別のため.

TruSeq Small RNAのリードを Fastq toolkitでトリムする

3 トリムしたいアダプター配列を選ぶ:

“Adapter trimming” > “Adapter sequences(s) to trim from the 3’ end”:
“TGGAATTCTCGGGTGCCAAGG” (This is the TruSeq smallRNA adapter)

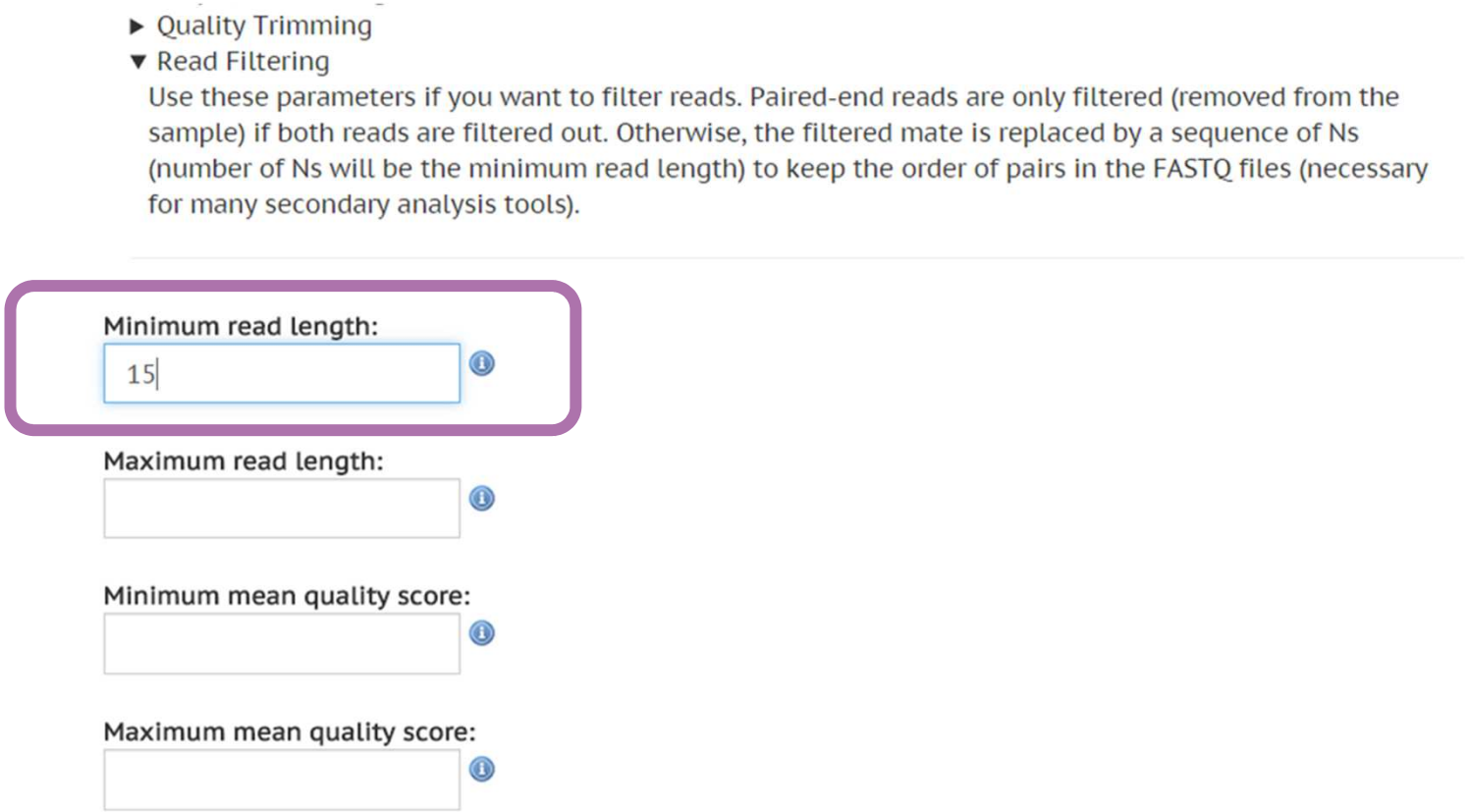


illumina®

TruSeq Small RNAのリードを Fastq toolkitでトリムする

4 最低リード長を入力

“Read Filter” > “Minimum Read length: 15” (変更可能)

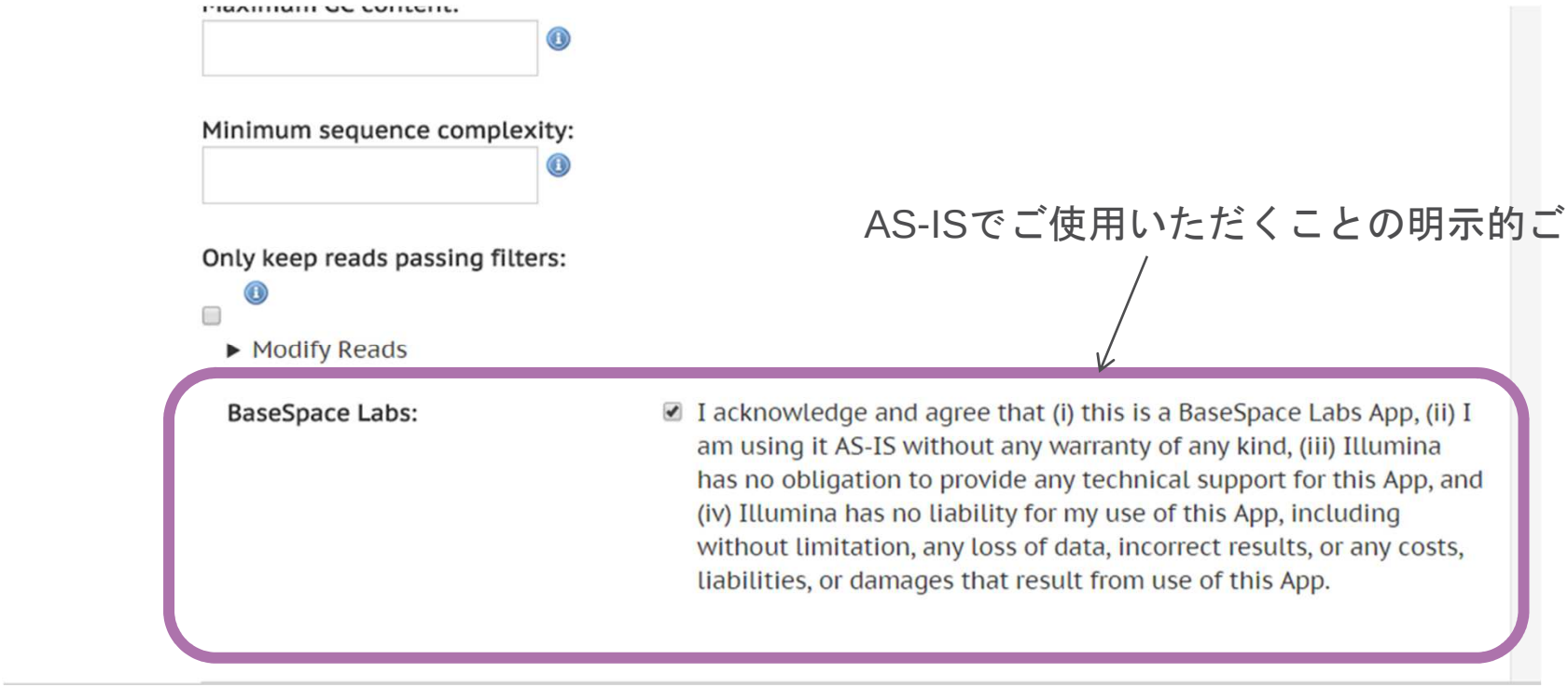


Note, that leaving as default will result in conversion of sequences <32bp to “N” strings

illumina®

TruSeq Small RNAのリードを Fastq toolkitでトリムする

5 “BaseSpace Labs Apps” Agreement にチェックを入れて承諾する



AS-ISでご使用いただくことの明示的ご了承

Continueボタンを押し、実行を開始する

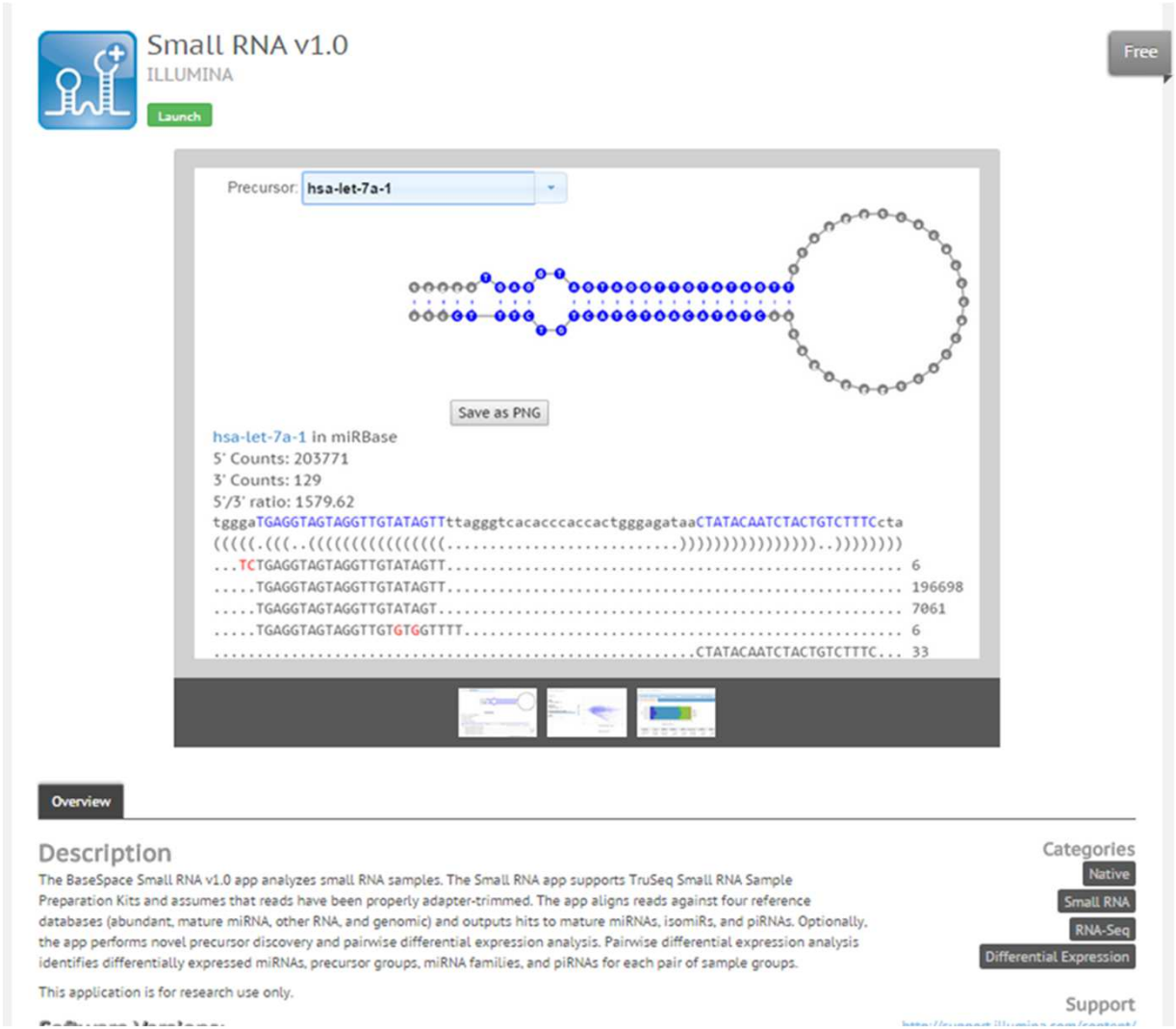
illumina

TruSeq smallRNA のリードを Fastq toolkitでトリム結果のレポート（ビフォーアフター)

（レポートの一部抜粋）

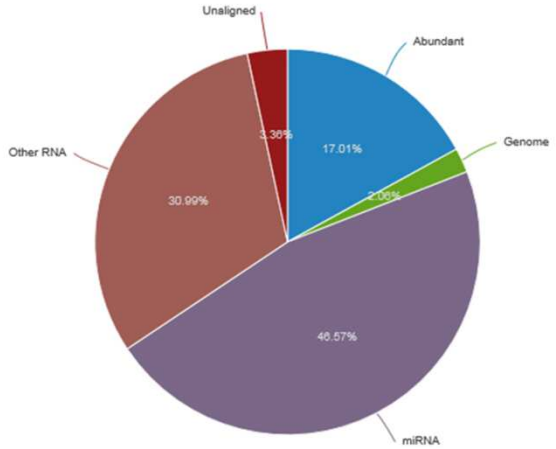
illumina®

BaseSpace Small RNAアプリ



illumina®

Small RNA アプリ結果のレポート



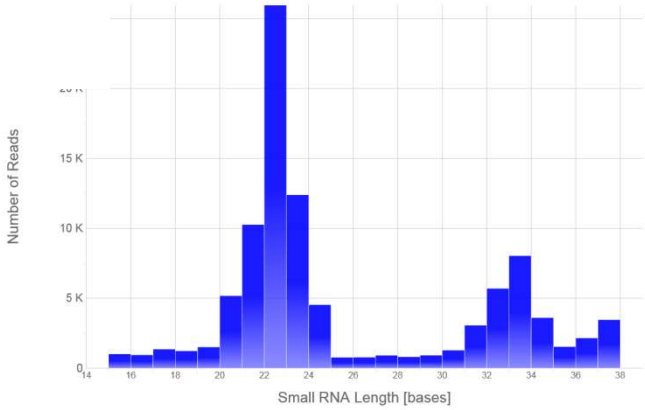
ces

ads are listed for each type. If there are less than 10 sequences with
sted.

ecords are counted. A read must align to the start of a reference
the same length of the reference sequence. No mismatches are

ads	Total Number of Reads
281	22,967

hsa-miR-22-3p
hsa-let-7a-5p
hsa-miR-9-5p
hsa-miR-27b-3p
hsa-let-7f-5p
hsa-miR-143-3p
hsa-miR-127-3p
hsa-miR-126-5p

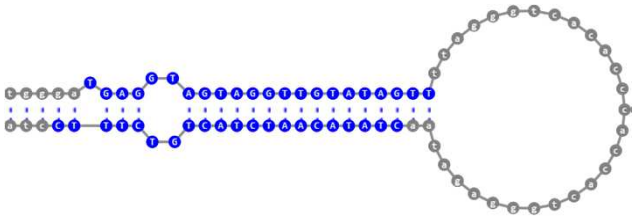


The full list is available at Hits.txt.

IsomiRs (Known Precursor)

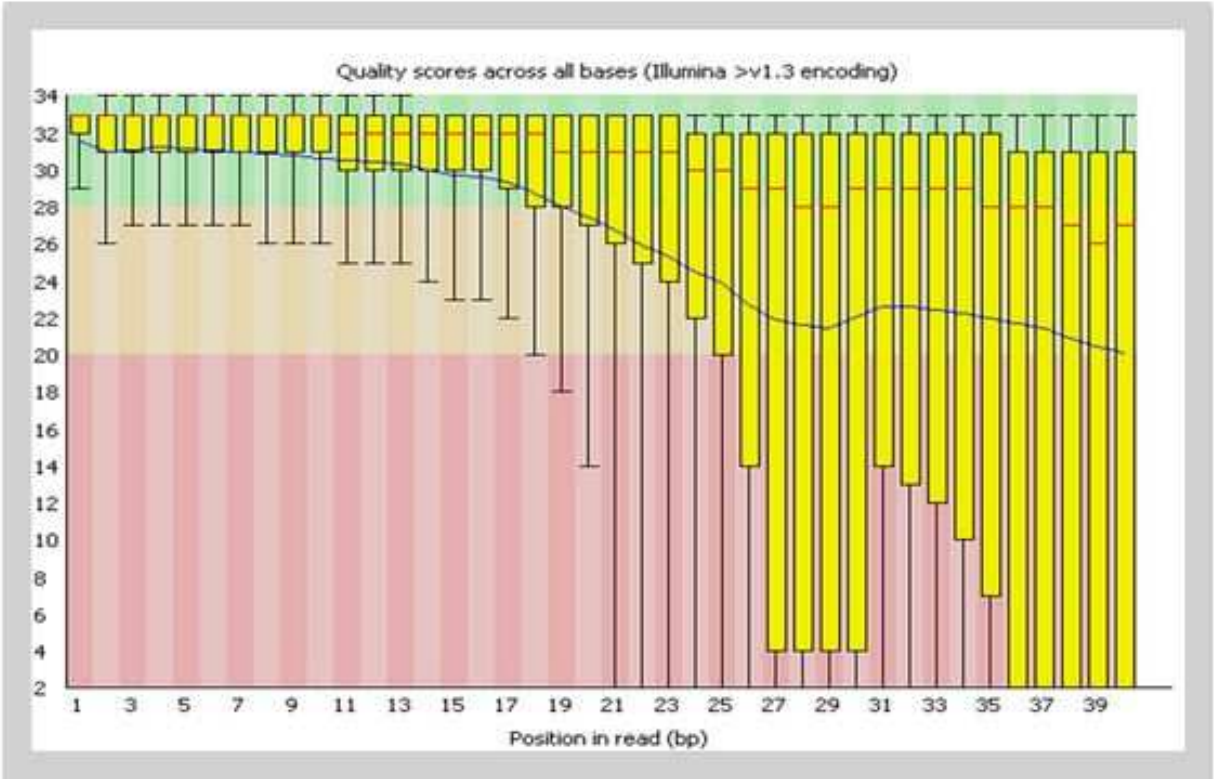
A read is counted as isomiR if it is a subsequence on the same strand of a pre
miRNA sequence. The precursor sequences were obtained from miRBase. An i
position}_{sequence read}_{precursor ID}", where the start position is 0-bas
Mismatches are allowed. IsomiR hits are filtered to remove artifacts of PCR an

Number of IsomiRs (Known Precursor) with Reads	Total Number of R
464	22



illumina®

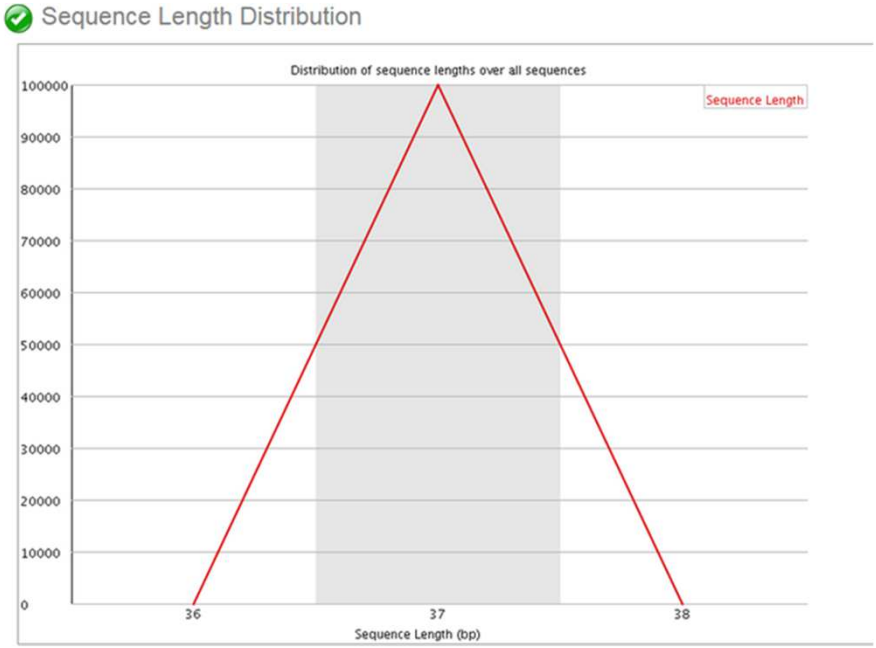
Small RNA
このFASTQリードはトリムされたものか？
– FastQCアプリ



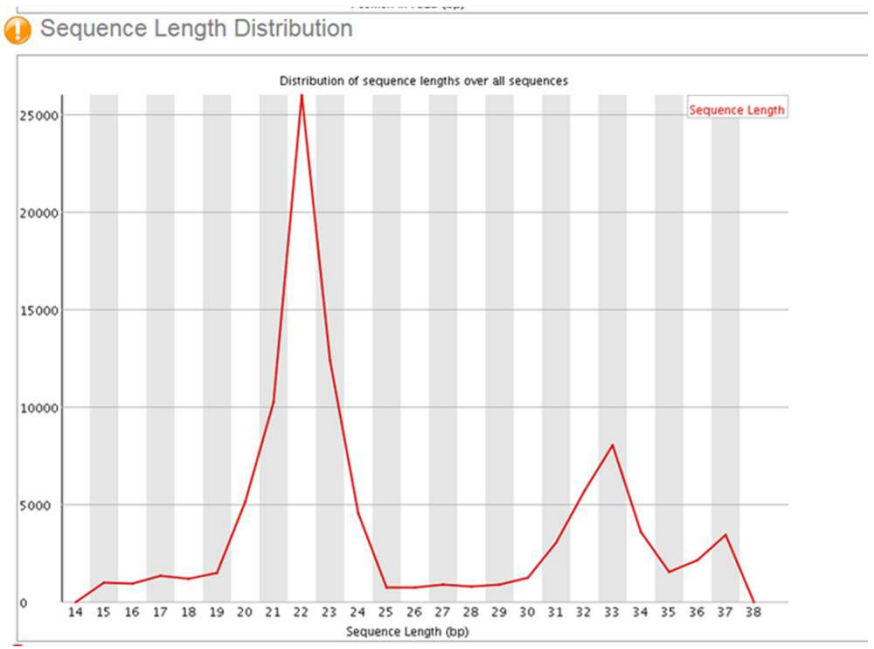
illumina®

Small RNA
このFASTQリードはトリムされたものか？
– FastQCアプリ

トリムされていない



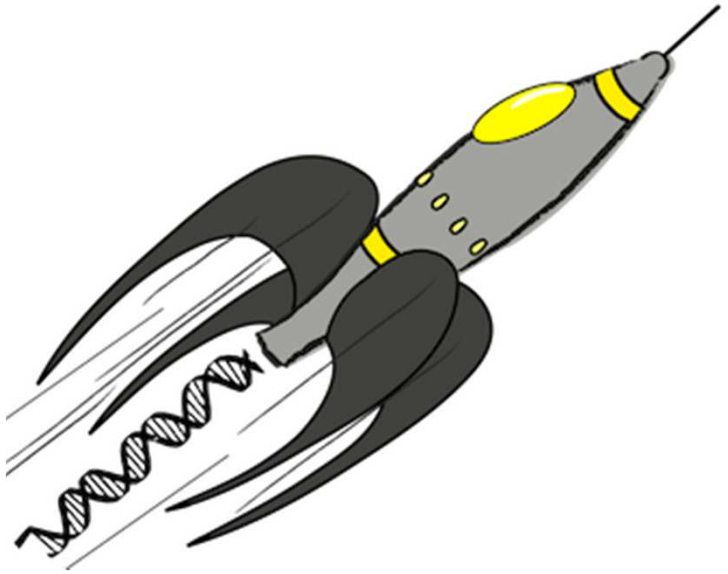
トリムされている



illumina®

本日の内容

- イントロダクション
- アダプタートリミング
- クオリティトリミング
- ダウンサンプリング
- リードの結合
- 手元のFASTQをトリミングするには



クオリティースコア(qscore) によるトリミング

とはなにか？

3’末端のクオリティーの平均に基づきトリミングする

3’末端からのスライディングウインドウのアプローチをとり、枠をスライドさせながら平均クオリティーが閾値を下回ったときに以降をトリムするものが多い

どういういう時に行うものなのか?

後続の解析でベースコールのクオリティがシビアに影響するような解析の場合。例えば– de novoアセンブリ、 リードの結合、 リードからの分類(メタゲノム解析など)

逆に、どのようなときは使われないもの?

- リシーケンシング解析．ほとんどのアライメントツールは塩基のqscoreも計算に入れており (i.e. BWA，Isaac) 、末端に低 qscore 配列がある場合はソフトウェア的に省く処理が実装されている等

Qスコアによるトリミング
GenerateFastq in MSR/ BaseSpace /bcl2fastq2)

[Header]					
IEMFileVersion	4				
Investigator Name	Mr.X				
Experiment Name	Example				
Date	4/05/2015				
Workflow	Resequencing				
Application	Resequencing				
Assay	TruSeq LT				
Description	Example				
Chemistry	Default				
[Reads]					
151					
151					
[Settings]					
Adapter	AGATCGGAAGAGCACACGTCTGAAC				
AdapterRead2	AGATCGGAAGAGCGTCGTGTAGGGA				
QualityScoreTrim	20				
[Data]					
Sample_ID	Sample_Name	Sample_Plate			
test	test			test	

QualityScoreTrim

[settings]
QualityScoreTrim,<qualityScore>

illumina®

Qスコアによるトリミングの例
QualityScoreTrim,20

ビフォー

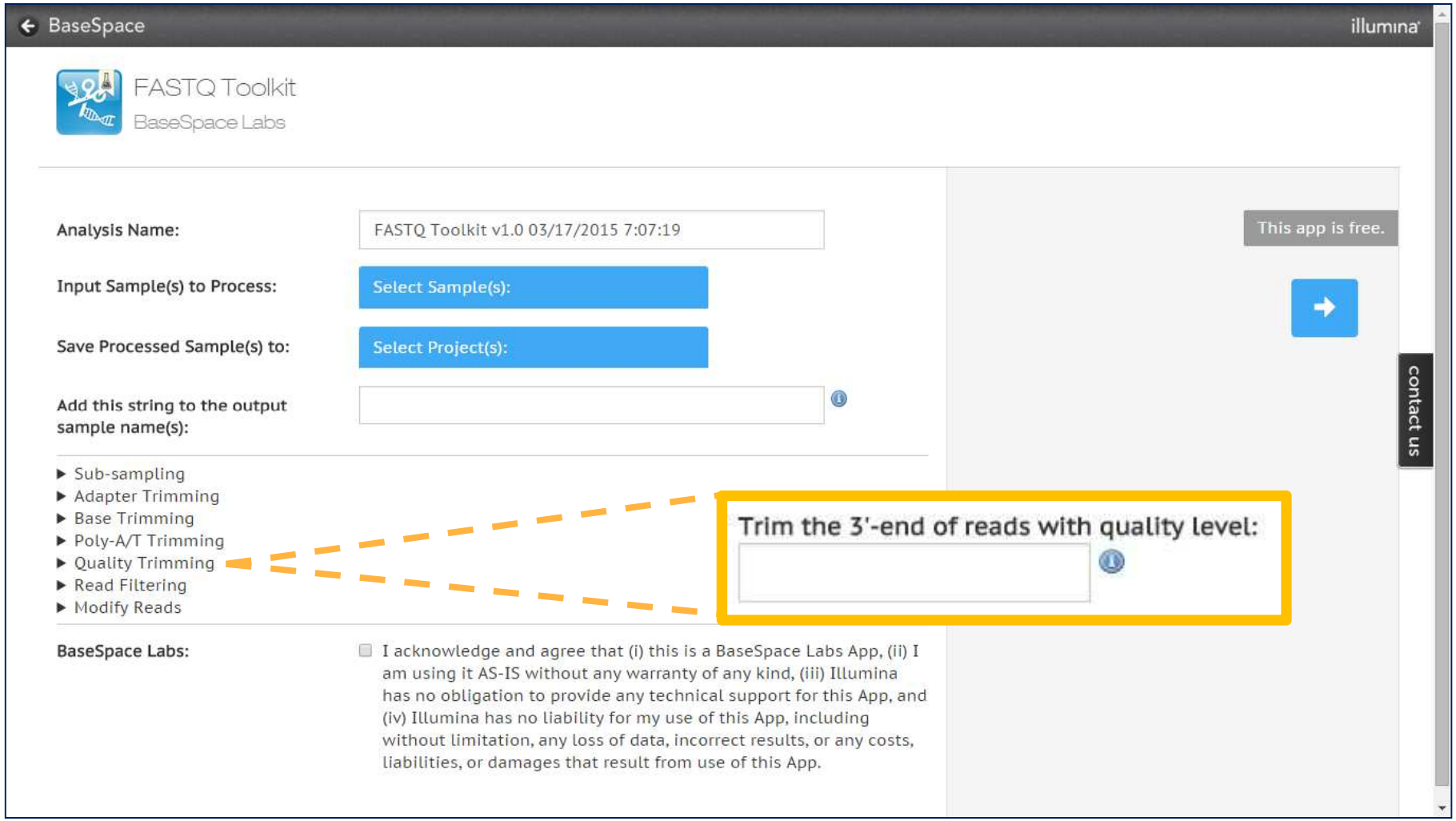
```
@M00000:72:000000000-D00LW:1:1101:22420:18334 1:N:0:1
CACCAAGGGCCTGGGGTGTCAATGGCGGGGCTTGTGACTGCACAAAAGGGGCCTCCCGCAGGGGCTCCCGCC
+
BBBBBBFBBBBBGGGGEEFGGGHHHHGGG00>10B355@BB3@3BG1?E1///1B11/////////?////
```

アフター

```
@M00000:72:000000000-D00LW:1:1101:22420:18334 1:N:0:1
CACCAAGGGCCTGGGGTGTCAATGGCGGGGCTTGTGACTGCACAAAAGG
+
BBBBBBFBBBBBGGGGEEFGGGHHHHGGG00>10B355@BB3@3BG1?E
```

Q	ASC
13	.
14	/
15	0
16	1
18	3
20	5
22	7
25	9
30	?
31	@
32	A
33	B

BaseSpace アプリによる Quality トリミング
FASTQ Toolkit



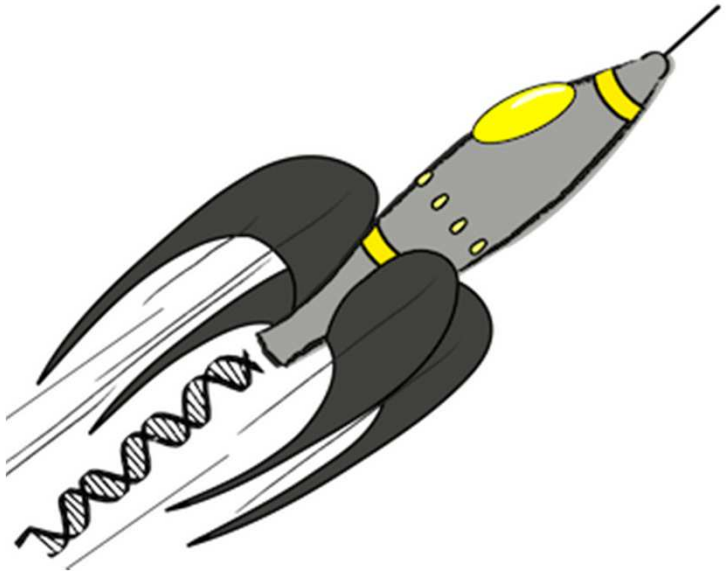
Quality トリミング
3rd- party ツール例

ツール名	URL
Trimmomatic	http://www.usadellab.org/cms/?page=trimmomatic
Trim-Galore	http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
FASTX toolkit	http://hannonlab.cshl.edu/fastx_toolkit/ (FastQ clipper)

参考：http://omictools.com/adapter-trimming-c402-p1.html

本日の内容

- イントロダクション
- アダプタートリミング
- クオリティトリミング
- ダウンサンプリング
- リードの結合
- 手元のFASTQをトリミングするには



ダウンサンプリング (サブサンプリング)

とはなにか？

- リード量が多すぎるときに一部のリードを取り出す（サブセットをつくる）

なぜあえてサンプリングによりリード量を減らすのか?

- トラブルシュートなどで素早くリードを検分(QC)したいとき、全リードで分析するとあまりに大量で解析時間がかかるため、負荷軽減、時間短縮をねらって.

- 解析環境や解析ツール、サンプル特異性によって解析系が大量リードの処理に耐えない場合がある. このエラーを回避し解析を進めるために入力リード量を減らす必要が生じる場合がある.
  例）メモリー不足で落ちる、ディスク領域が足らないなど

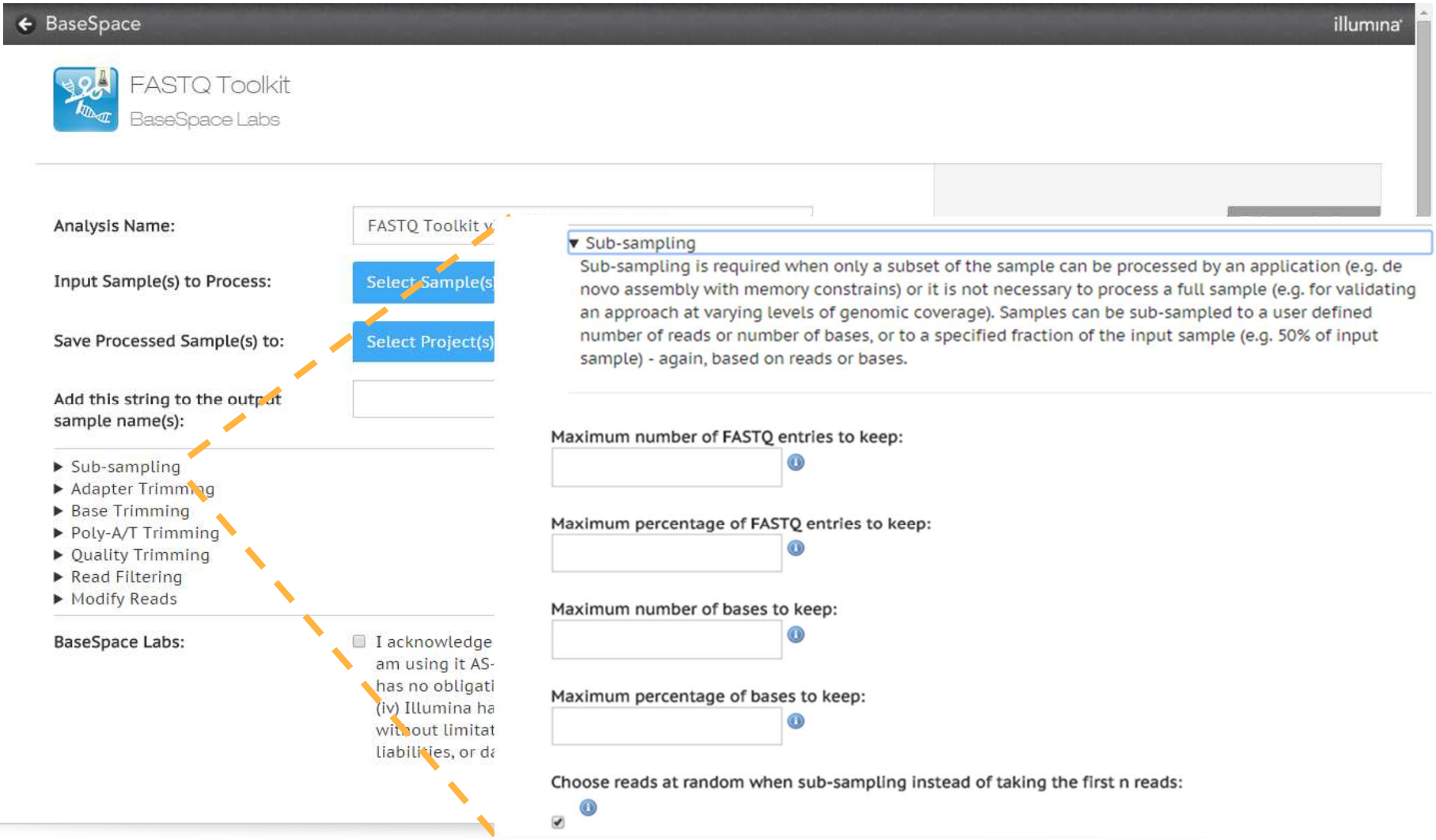
- BaseSpaceのアプリでも入力データ量の制限を明記しているものがある.
  こういったアプリや3rd-partyツールの入力制限に合わせるため.

- 入力量で解析結果がどのように影響されるかなどの解析条件検討.

イルミナでサブサンプリングをするには

- BaseSpace FASTQ toolkit アプリ

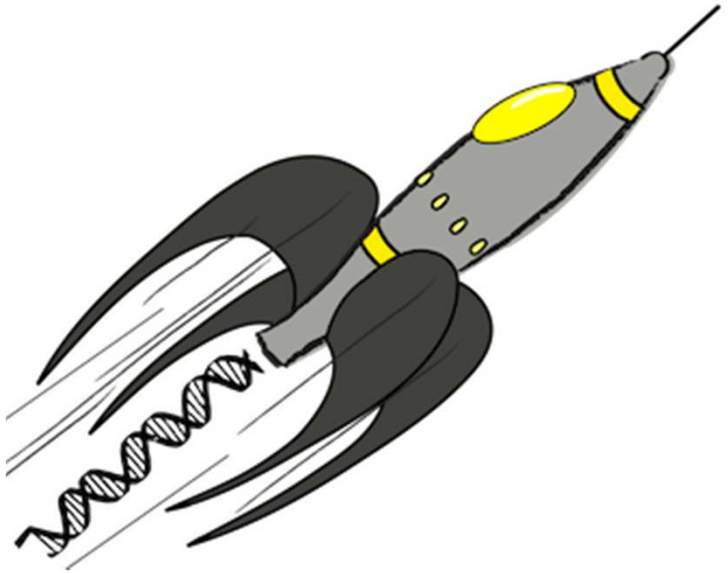
BaseSpace App: FASTQ Toolkitによるサブサンプリング



illumina

本日の内容

- イントロダクション
- アダプタートリミング
- クオリティトリミング
- ダウンサンプリング
- リードの結合
- 手元のFASTQをトリミングするには



リードのマージ　(結合、join、stitch など呼称さまざま)

とはなにか?

- **重複領域を頼りにリードをつなぎ合わせること**

  狭義では、ペアードエンドのR1とR2をつなぎ合わせること

通常はある程度クオリティーの良い塩基のオーバラップが一定長以上あること
を条件とし、つなぎあわせる処理を行う (Q15以上の塩基が連続25bp以上など)

どういういう時に行うものなのか?

- リードを長くすることが大切な場合
- indel 検出の向上に使えることもある
- 以降の解析ツールがシングルエンドしか受け付けない様なものの場合
  (一部のメタゲノム解析ツールなど)
- ほとんどのリードがオーバラップするようなデザインで読んだもの

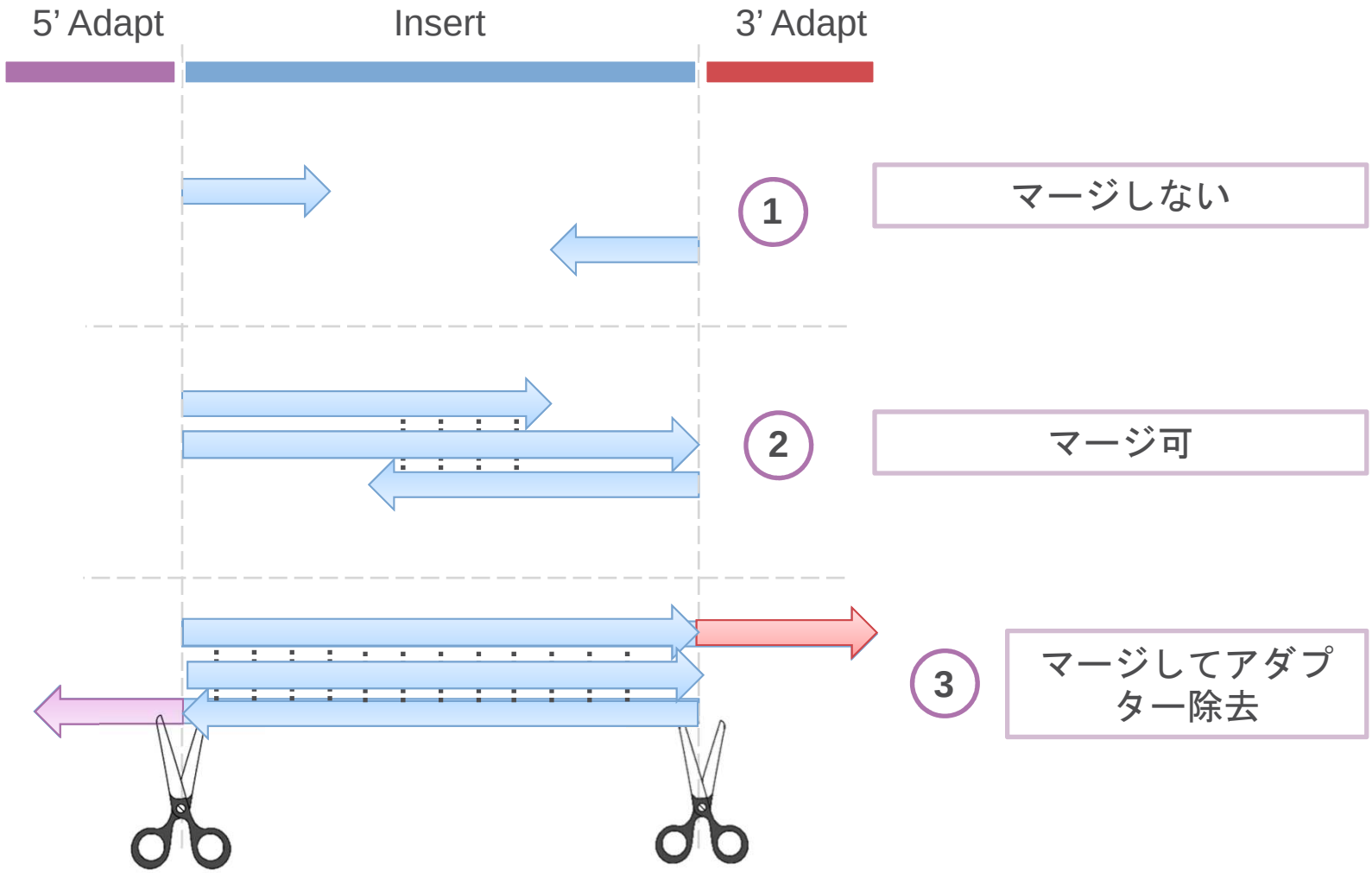
逆に、適さないときは?

- クオリティーの良い塩基のオーバラップがない
- 一部のリードしかオーバラップがない場合 (設計外)
- オーバーラップ領域にリピート配列が予想されるとき

イルミナでリードのマージをするには

- MiSeq ReporterではStitch Readという機能でR1,R2のマージ可能 (一部ワークフロー)

リードマージの概念図



illumina

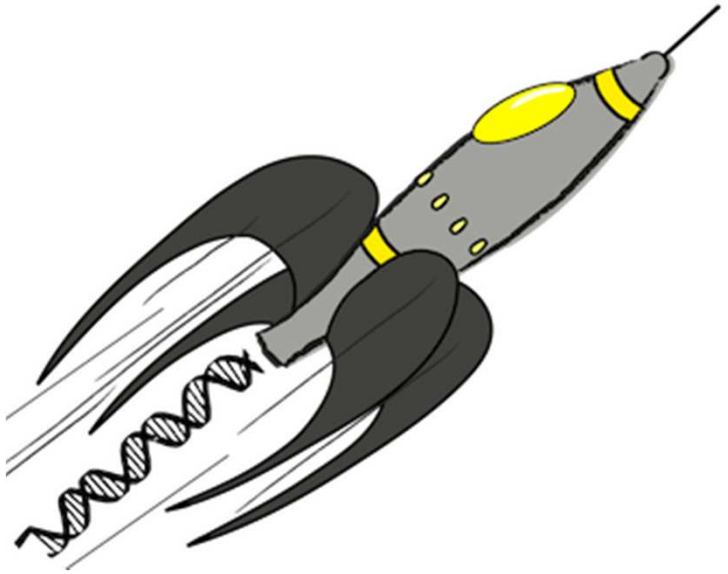
リードマージができるツールの一例
3rd-partyツール

ツール名	URL
FLASH	http://ccb.jhu.edu/software/FLASH/
Panda-seq	https://github.com/neufeld/pandaseq
Seq-Prep	https://github.com/jstjohn/SeqPrep
PEAR	http://sco.h-its.org/exelixis/web/software/pear/
FASTQ-Join	https://code.google.com/p/ea-utils/wiki/FastqJoin

等

本日の内容

- イントロダクション
- アダプタートリミング
- クオリティトリミング
- ダウンサンプリング
- リードの結合
- 手元のFASTQをトリミングするには



BaseSpace データ取り込みパターン

お手持ちの
Illumina FASTQ,
VCF

FASTQやVCFの
アップロード

※フォーマット等条件があるため、
基本的にはランからのアップロードを推奨
詳細はBaseSpace UserGuideをご参考下さい。

BaseSpace

クラウドにある公開デモデータの
取込みから

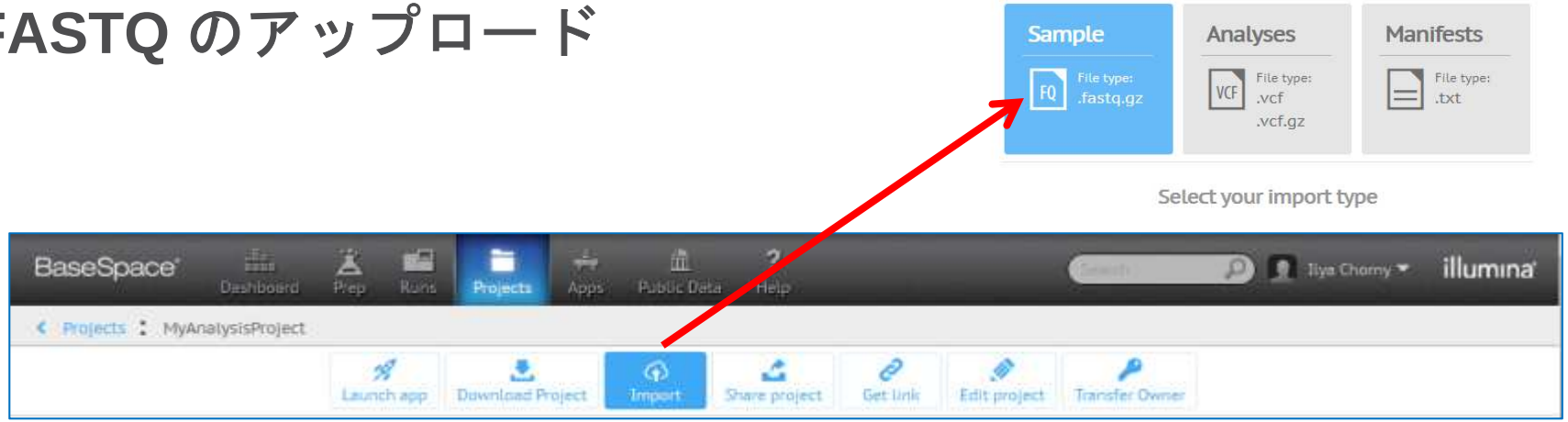
BaseSpace™

SRA Import Labアプリから

☆ラン中のデータをアップロードして自動開始！

※ (SRP*/ERP*/DRP*), experiments (SRX*/ERX*/DRX*), samples (SRS*/ERS*/DRS*), runs (SRR*/ERR*/DRR*), or submissions (SRA*/ERA*/DRA*)対応。ただしイルミナデータのみ、1回のimportは25GBまで。

FASTQ のアップロード



規約： ☆ イルミナリードのみに対応しており、ファイル名が以下のようなイルミナ標準である

SampleName_SampleNumber_Lane_Read_FlowCellIndex.fastq.gz

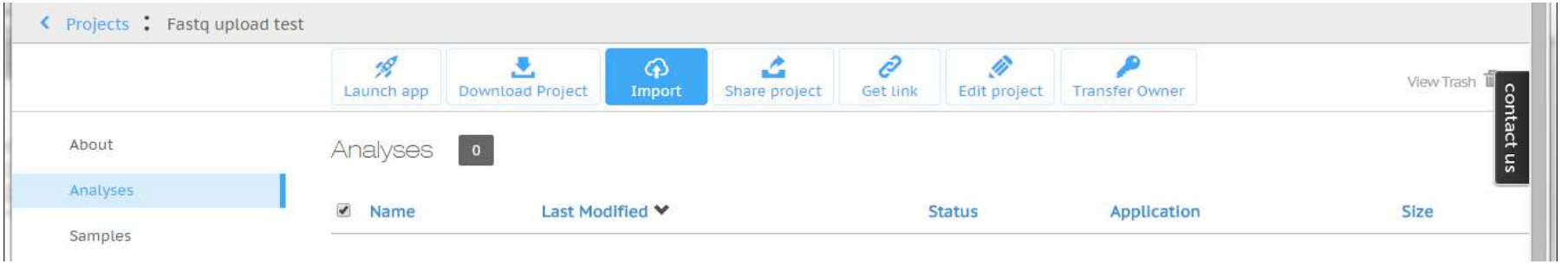
☆ gzipされている
☆ クオリティスコアの数が塩基数と一致している
☆ 各リードのヘッダが以下のようなイルミナ標準を満たしている

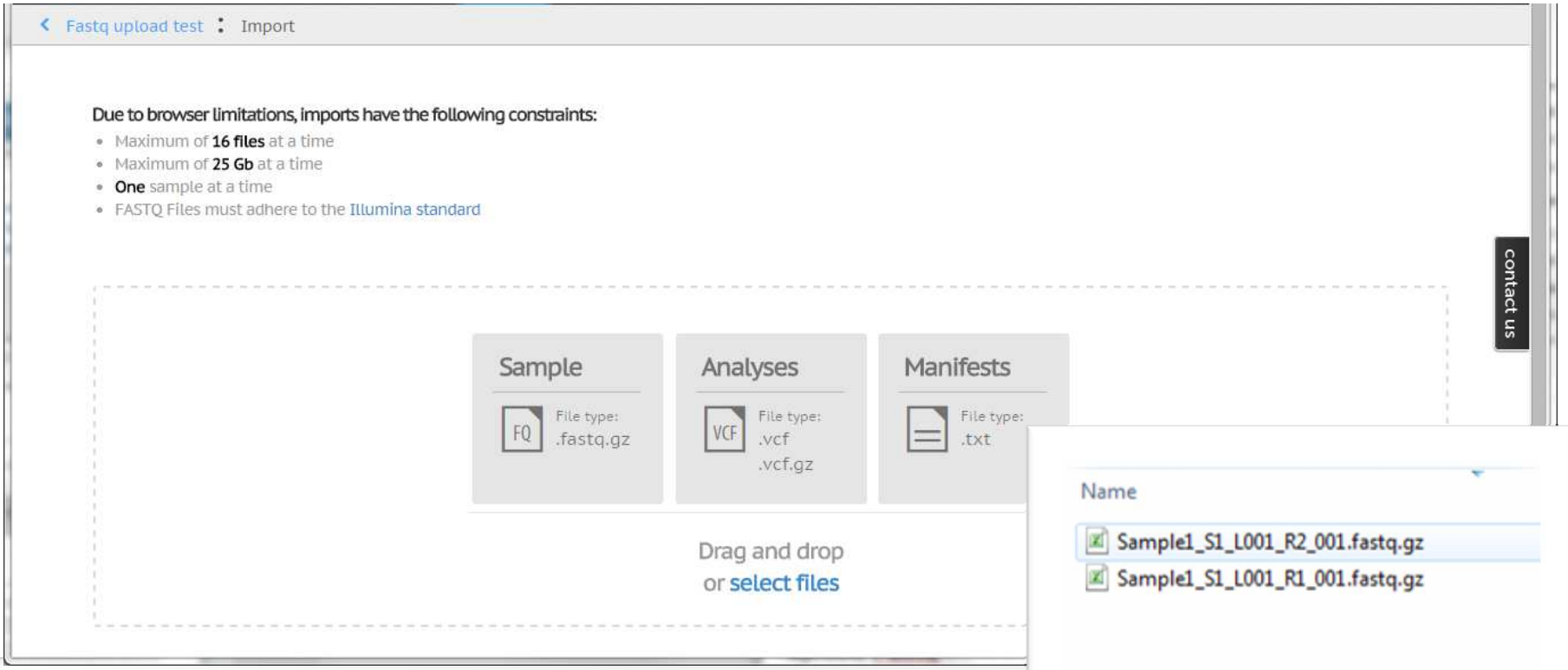
@Instrument:RunID:FlowCellID:Lane:Tile:X:Y ReadNum:FilterFlag:0:SampleNumber

ペアードエンドリードの場合さらに；
☆ R1とR2でヘッダがペアとして揃ったリード（ReadNumが1と2）が等数ある
☆ R1, R2ともにPF (Pass Filter)したリード（FilterFlagがN）のみ
☆インポート可能な最大サイズは25GByteまで
☆最大16ファイル/サンプル
☆１サンプル単位で逐次インポート（* Completeになってから次の処理を開始下さい）

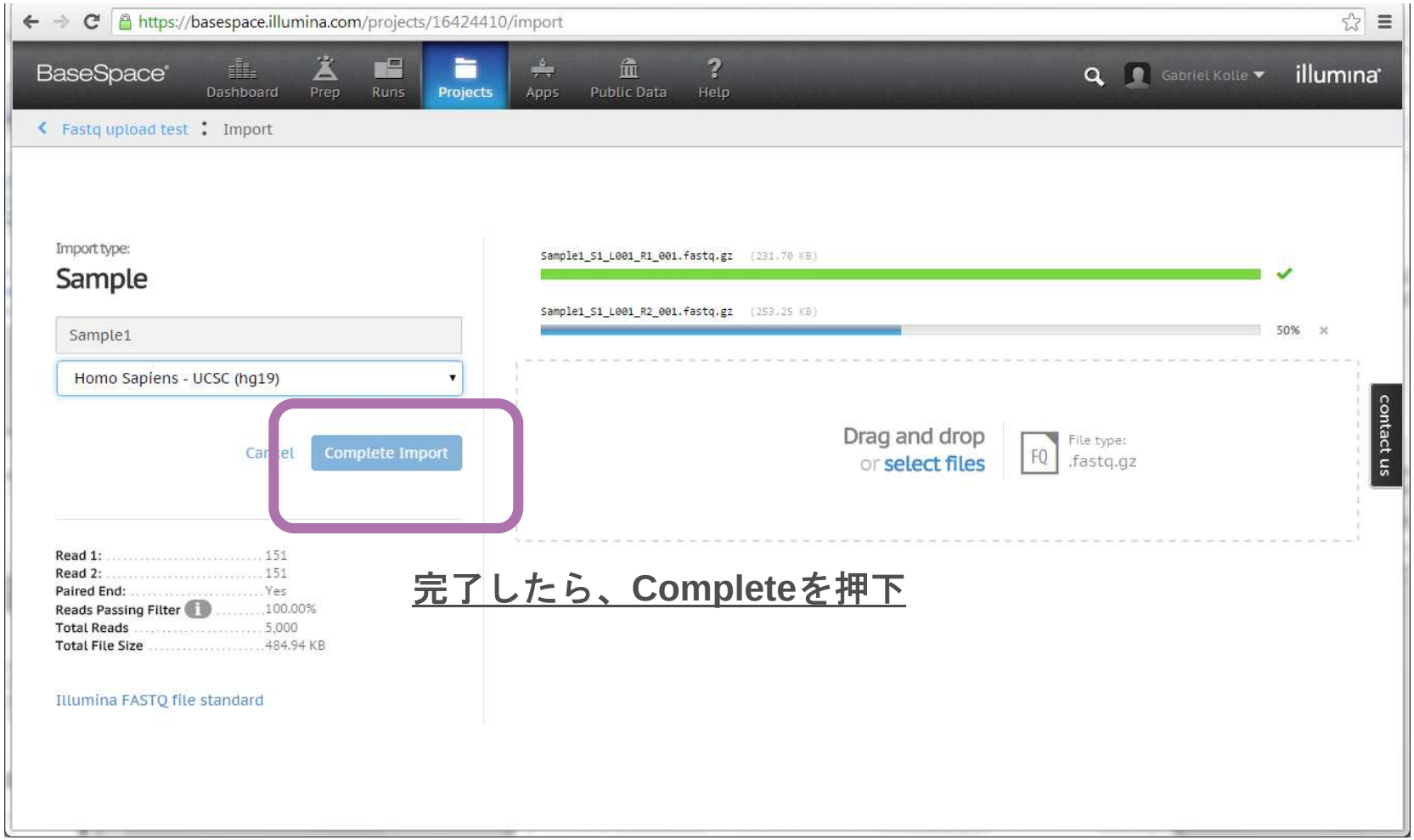
http://support.illumina.com/content/dam/illumina-support/documents/documentation/software_documentation/basespace/basespace-user-guide-15044182-e.pdf p.54

FASTQ のアップロード



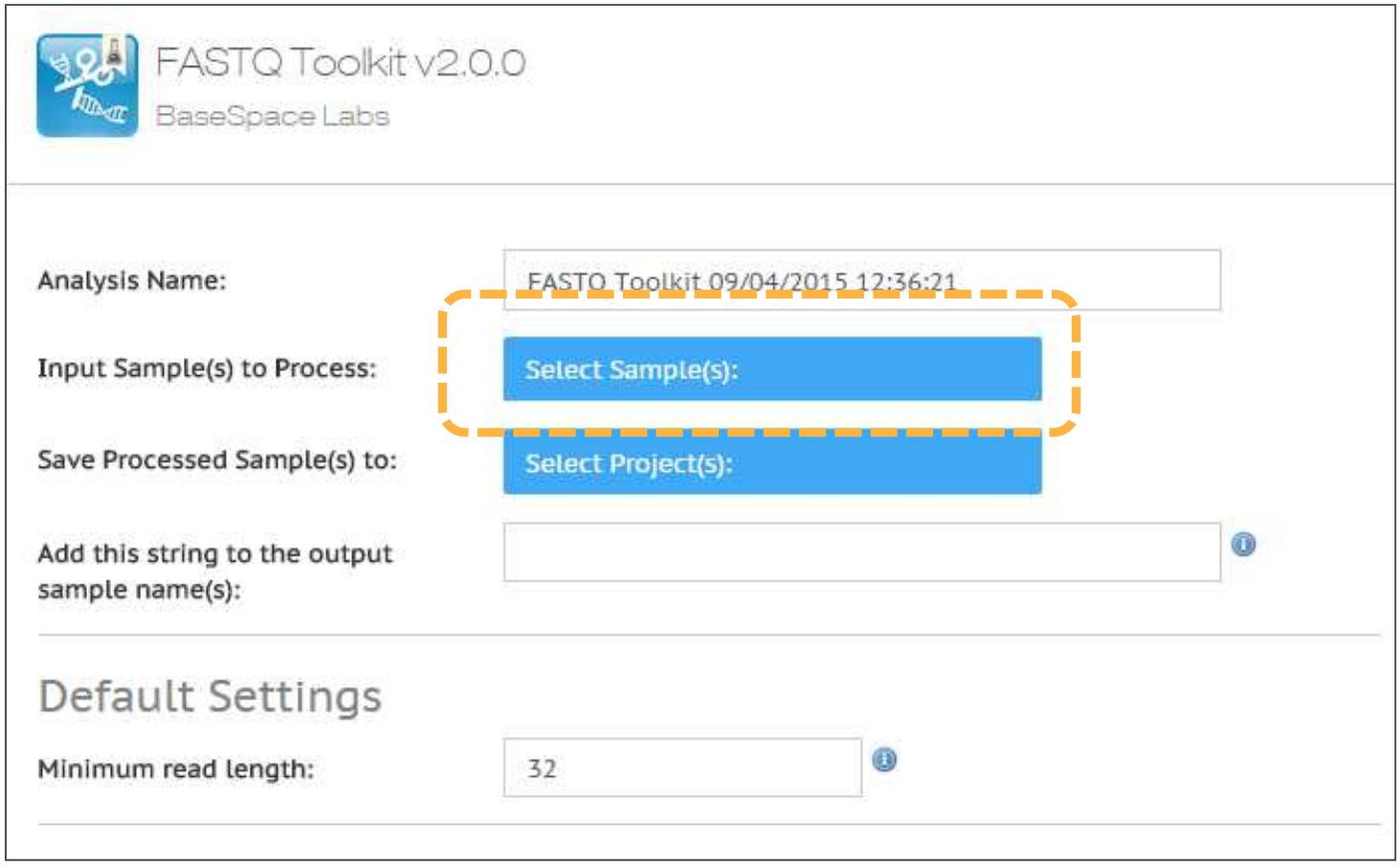


FASTQ のアップロード



完了したら、Completeを押下

FASTQ Toolkit の開始画面から、先ほどアップロードした FASTQをSelect Sample(s): から選択し、トリミングを開始



illumina®

ご参考；

Adapter trimming sequences テクニカルブルテン

https://my.illumina.com/MyIllumina/Bulletin/qFYNf9hn_kW5SyEZwGOUrA/adapter-sequences-for-use-with-casava-or-bcl2fastq

Nextera メイトペアのアダプタートリミング

http://res.illumina.com/documents/products/technotes/technote_nextera_matepair_data_processing.pdf

MiSeq Reporter GenerateFastq ワークフローガイド

http://support.illumina.com/content/dam/illumina-support/documents/documentation/software_documentation/miseqreporter/miseq-reporter-generatefastq-workflow-guide-15042322-b.pdf

bcl2fastq 変換ソフトウェア:

http://support.illumina.com/downloads/bcl2fastq_conversion_software.html

ご参考；

BaseSpace
basespace.com

BaseSpace Fastq Toolkit:
▪App について: http://www.illumina.com/informatics/research/sequencing-data-analysis-management/basespace/basespace-apps/fastq-toolkit-962962.html
▪紹介ブログ: http://blog.basespace.illumina.com/2014/12/22/rounding-out-2014-with-new-apps-for-the-basespace-platform-2/
▪サポートアドレス: basespacelabs@illumina.com

BaseSpaceコアアプリ各ワークフローのフローチャート図は各ユーザガイドにあります
support.illumina.com/downloads/basespace_core_apps_user_guides.html

BaseSpace最新News
blog.basespace.illumina.com　　#RSS 購読可能

ヘルプセンター（ウェブヘルプ）
help.basespace.illumina.com

サポートウェビナーにご参加いただき
ありがとうございました。

本日のセッション終了後のご質問は、
techsupport@illumina.com
で承ります。