

MiSeq Reporter update (-> MSRv2.3)

癸生川絵里 (Eri Kibukawa)
Bioinformatics Support Scientist
Illumina



MiSeq Reporter



© 2012 Illumina, Inc. All rights reserved.

Illumina, illuminaDx, BaseSpace, BeadArray, BeadXpress, cBot, CSPPro, DASL, DesignStudio, Eco, GAllx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, Sentrix, SeqMonitor, Solexa, TruSeq, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.

illumina®



Outline

- ▶ MiSeq Reporter (MSR) 概要へのポイント
- ▶ MSRの11ワークフロー概要
- ▶ 変更点を中心とした機能概説
- ▶ 資料のご案内
- ▶ トラブルシューティング時にお送り頂きたいファイル



MiSeq Reporter



illumina®

MiSeq Reporter のインターフェイス操作, パラメータ変更方法について

- ▶ イルミナサポートウェビナー

http://www.illumina.co.jp/events/webinar_japan.ilmn?ws=ss

- ▶ 2013/01/18 「MiSeq Reporter v2 によるMiSeqデータ解析」

- ▶ - MSR インターフェイスと基本ワークフロー毎のポイント、また

Q:インデクス配列を確認するには

にてインデクス確認方法および、MSRパラメータ設定変更の方法をカバーしておりますのでご参考下さい。

MSRで解析結果がみえない場合

- ▶ イルミナサポートウェビナー (pdfあり)
http://www.illumina.co.jp/events/webinar_japan.ilmn?ws=ss
- ▶ **2012/06/15** 「MiSeq Reporterをはじめよう」
MSRのバージョンは古いですが、

Q : MSRで解析結果がでてきません

際のトラブルシューットの仕方や、MSRサービスの再起動方法、
MiSeqのIP確認方法 等末尾にカバーしておりますのでご参考下さい。

BaseSpaceの利用方法

イルミナサポートウェビナー (pdfあり)

http://www.illumina.co.jp/events/webinar_japan.ilmn?ws=ss

2013/09/06 「BaseSpace リリース版 (MiSeq/HiSeq) 」

2012/11/22 「BaseSpace - genomics cloud computing

-- ベーススペースの使いかた」

<http://basespace.illumina.com>

MiSeq Reporter リリースノート

- ▶ http://www.illumina.co.jp/document/pdf/release_note_mcs_2-3_j.pdf

*MiSeqのリリースノートはサポートメールにてお知らせしております。
受け取られていないお客様でメールにご登録されたい方は、
テクニカルサポートまでお知らせ下さい。

http://supportres.illumina.com/documents/documentation/system_documentation/miseq/miseq-updater-v2-3-software-release-notes.pdf

(オリジナル英語版. 上記日本語版よりリンクあり)

MiSeq Reporter ワークフロー追加の歴史 (～MSR v2.3)

- Resequencing
- Custom Amplicon -> TruSeq Amplicon
- De novo assembly
- GenerateFASTQ (MSRv1.3～)
- Library QC
- 16S Metagenomics
- PCR Amplicon (MSR v2.0～)
- Small RNA
- Enrichment (MSR v2.1～)
- Targeted RNA (MSR v2.2～)
- Amplicon –DS (MSR v2.2+plug-in, MSR v2.3ワークフロー)



*バージョン間の変更はReleaseNoteをご参考下さい。

MiSeq Reporterの11ワークフロー概要

Resequencing

リファレンスへのアライメントと変異コール

Library QC

リファレンスへのアライメント。変異コールは行わない
DNAライブラリの素早いQCレポート生成に最適化

TruSeq Amplicon (Custom Ampliconより名称変更)

イルミナTSCAライブラリ用。TSCAのマニフェスト
ファイルにある領域に対しアライメントと変異コールを
行う

PCR Amplicon (Nextera-fragmented)

Nexteraを用いたPCRアンプリコン用。Manifest
ファイルで指定されたアンプリコン領域に対して
リードのアライメントと変異コールを行う

de novo Assembly

小さいゲノム(<20 Mb) をアセンブル。velvetを利用。
リファレンス配列を与えた場合は生成コンティグ(fasta)
とのドットプロットを描画

16S Metagenomics

16S rRNA bacterial 解析。Classification にはGreenGenesを
利用。v2.3からアルゴリズムとdb バージョン変更

Small RNA

コンタミ指定の配列を除いた後、miRBaseのmature miRNA
と一致する配列を検出。一致しないものは、
その他smallRNA > genome と順次アライメント。
デフォルトでAdapter 配列をマスク。

Enrichment

Nextera Rapid CaptureとTruSightパネル(Tumor以外)に対応。
リード配列はゲノム全体にアライメントされ、
変異コールは Manifestファイルで指定された領域のみ。

Targeted RNA

イルミナTargeted RNA Expression Kitライブラリ用。
マニフェストファイルにより定義したターゲット領域における
発現変動解析、二群間比較。

Amplicon-DS(double stranded)

TruSight Tumor targeted assay専用のワークフロー
Manifestファイルで指定されたアンプリコン領域に対して
リードのアライメントを行い、somatic variant callerで
変異を検出する。FFPEサンプルの体細胞性変異検出に適する。

GenerateFASTQ

その他のワークフローがFASTQ生成後に各解析を行うのに
対し、ベースコールファイルから FASTQ ファイルに変換した
ところで終了する。
3rdパーティー等その他のソフトで解析を続けたい場合に指定。

Resequencing

Resequencing

リファレンスへのアライメントと変異コール

Library QC

リファレンスへのアライメント。変異コールは行わない
DNAライブラリの素早いQCレポート生成に最適化



MiSeq Reporter



BWAとGATKのバージョンは変更ありません

MSR v2.3

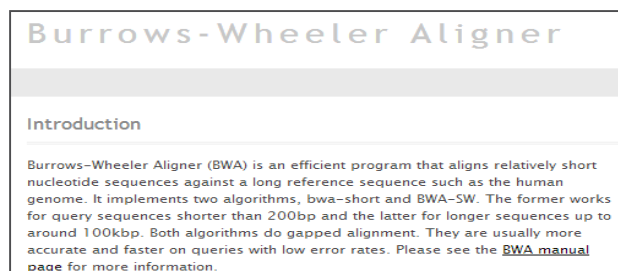


* v1.6 (MSR v2.1～同様)

<http://www.broadinstitute.org/gatk/guide/>

*使用されたGATKバージョンはvcfファイルのヘッダにも記載されています。

BWA



* v0.6.1

<http://bio-bwa.sourceforge.net/bwa.shtml> (**Manual Reference Pages**)

http://icb.med.cornell.edu/wiki/index.php/Elementolab/BWA_tutorial

GATK, BWAはスタンダードな3rd party ツールですので、インターネット上に様々な情報があります。

(オプション)

dbSNP 137が利用可能になりました (Humanのみ)

- 弊社の以下のページから配布している手順書、およびファイルで**dbSNP137**を使用した変異解析が可能になります (MSRはデフォルトでは**dbSNP131**を使用)

http://support.illumina.com/sequencing/sequencing_software/miseq_reporter/downloads.ilmn

Updating MiSeq Reporter with dbSNP version 137

This document provides instructions for updating the dbSNP information used by the MiSeq Reporter software. Use the file provided to update MiSeq Reporter with dbSNP version 137.

MiSeq Reporter uses dbSNP in the variant calling step of the Resequencing, PCR Amplicon, and TruSeq Amplicon workflows. SNPs that are identified by MiSeq Reporter variant calling are cross-referenced against the dbSNP information. If a match is found in dbSNP, the dbSNP identifier (rs number) is written to the VCF file. For more information, see the associated [workflow guide](#) available MiSeq Report support page on the Illumina website.

dbSNP is the NCBI database used to annotate variants found during analysis by MiSeq Reporter. It is downloaded from UCSC using the link below. The file is not edited or altered by Illumina.

<http://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/snp137.txt.gz>

Downloads

Downloads

DESCRIPTION

FILE INFO DATE

+ MiSeq Reporter Installer

MiSeq Reporter is the bioinformatics data analysis software built into the MiSeq. This software is available as a standalone software for installation on a Windows workstation.

- MiSeq Reporter dbSNP Update

dbSNP is the NCBI database used to annotate variants found during analysis by MiSeq Reporter. This download includes a link to the updater from a secure file transfer location, and instructions for downloading and installing the update.

» Instructions for MiSeq Reporter dbSNP v137 Updater

PDF (< 1 MB) 08/16/2013

» MiSeq Reporter dbSNP v137 Updater

ZIP (1.25 GB) 08/17/2013

クリックして
ダウンロード

特定領域(Amplicon系)の解析ワークフロー

TruSeq Amplicon (Custom Ampliconより名称変更)

イルミナTSCAライブラリ用。TSCAのマニフェストファイルにある領域に対しアライメントと変異コールを行う

PCR Amplicon (Nextera-fragmented)

Nexteraを用いたPCRアンプリコン用。Manifestファイルで指定されたアンプリコン領域に対してリードのアライメントと変異コールを行う

Amplicon-DS(double stranded)

TruSight Tumor targeted assay専用のワークフロー
Manifestファイルで指定されたアンプリコン領域に対してリードのアライメントを行い、somatic variant callerで変異を検出する。FFPEサンプルの体細胞性変異検出に適する。

Enrichment

Nextera Rapid CaptureとTruSight/パネル(Tumor以外)に対応。
リード配列はゲノム全体にアライメントされ、
変異コールは Manifestファイルで指定された領域のみ。

Targeted RNA

イルミナTargeted RNA Expression Kitライブラリ用。
マニフェストファイルにより定義したターゲット領域における
発現変動解析、二群間比較。



MiSeq Reporter



特定領域の解析ワークフローで使用するマニフェストファイル

- 解析対象配列領域を指定したイルミナ独自のポジションファイル
- タブ区切りテキストファイル。このため、notepadなどで開くことができる。
- “マニフェストファイル”は総称で、ワークフローによりフォーマットの違いがあるので留意する

▶ 以下のAmplicon系ワークフローにはマニフェストファイルが必用

- TruSeq Amplicon
- PCR Amplicon
- Enrichment
- Targeted RNA
- Amplicon-DS

マニフェストファイル例；

```
[Header]
XT Manifest Version      1
ReferenceGenome Homo_sapiens¥UCSC¥hg19¥Sequence¥WholeGenomeFASTA

[Regions]
Name      Chromosome      Amplicon Start      Amplicon End      Upstream Probe Length
Amplicon 1      chr1      1      15000      10      10
Amplicon 2      chr2      2499      78907      10      10
```

▶ マニフェストファイルの拡張子や名前は認識されるので変えないよう注意する

gVCF が出力可能に

gVCFが出力されるワークフロー

-> RNAを除く 特定領域の解析ワークフローにてデフォルト出力される

TruSeq Amplicon

Enrichment

PCR Amplicon

Amplicon-DS

* OutputGenomeVCF パラメータで出力しないように設定することも可能

gVCF とは?

- VCFの拡張フォーマット
- VCF v4.1に準拠している
- 変異が検出されたサイトだけではなく、変異が検出されなかったサイトの情報がどうであったのかもvcf形式で記述する
 - Genotype
 - Read-depth
 - Genotype quality score
 - Mapping quality score
 - Filter pass status
 - Etc
- まだ業界では対応していない3rd-party toolツールもある (例 ; vcf-merge)
- gVCF は VCFに記載のあるSNPs や Indels の情報も含む



gVCF フォーマットの仕様 ;

<https://sites.google.com/site/gvcftools/home/about-gvcf>

VCF (Variant Call Format) 形式

仕様; <http://www.1000genomes.org/wiki/Analysis/Variant%20Call%20Format/vcf-variant-call-format-version-41>

```
chr1      70097194      .      A      G      383.00  PASS      SNVSB=-
47.5;SNVHPOL=2  GT:GQ:GQX:DP:DPF:AD      1/1:84:84:29:0:0,29
chr1      70097708      .      T      G      680.00  PASS      SNVSB=-
61.7;SNVHPOL=4  GT:GQ:GQX:DP:DPF:AD      1/1:150:150:51:3:0,51
chr1      70098387      .      G      C      430.00  PASS      SNVSB=-
43.8;SNVHPOL=3  GT:GQ:GQX:DP:DPF:AD      1/1:93:93:32:1:0,32
chr1      70099418      .      A      G      368.00  PASS      SNVSB=-
33.7;SNVHPOL=2  GT:GQ:GQX:DP:DPF:AD      1/1:72:72:25:0:0,25
chr1      70100533      .      C      G      495.00  PASS      SNVSB=-
56.9;SNVHPOL=2  GT:GQ:GQX:DP:DPF:AD      1/1:102:102:35:0:0,35
```

Info:

SNVSB = Strand bias estimate

SNVHPOL = SNV contextual homopolymer length

Sample Info:

GT = Genotype (0/1 = het, 1/1 = hom)

GQ = Genotype quality

GQX = Minimum of {Genotype quality assuming variant position, Genotype quality assuming non-variant position}.

DP = Filtered basecall depth used for site genotyping

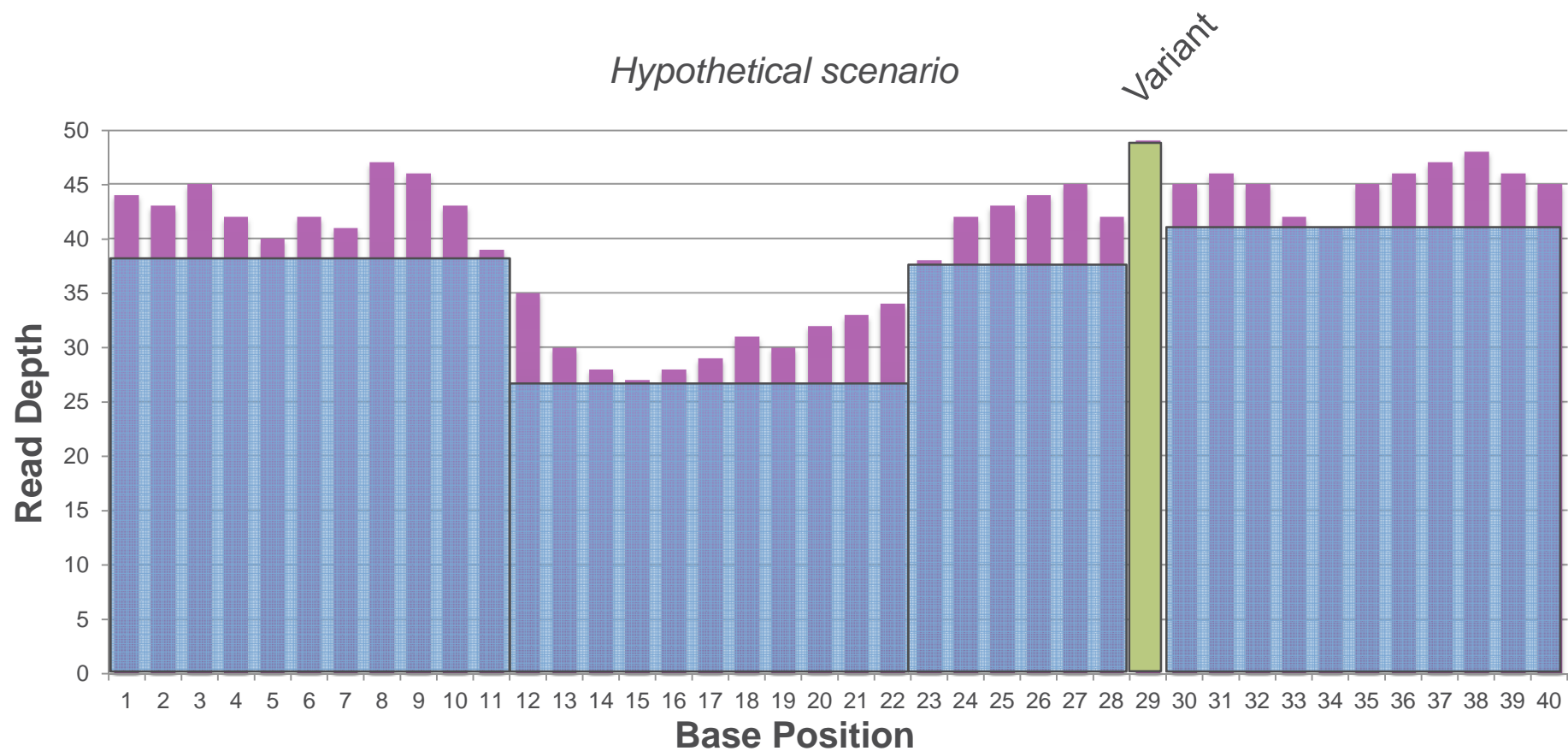
DPF = Basecalls filtered from input prior to site genotyping

AD = Allelic depths for the ref and alt alleles in the order listed

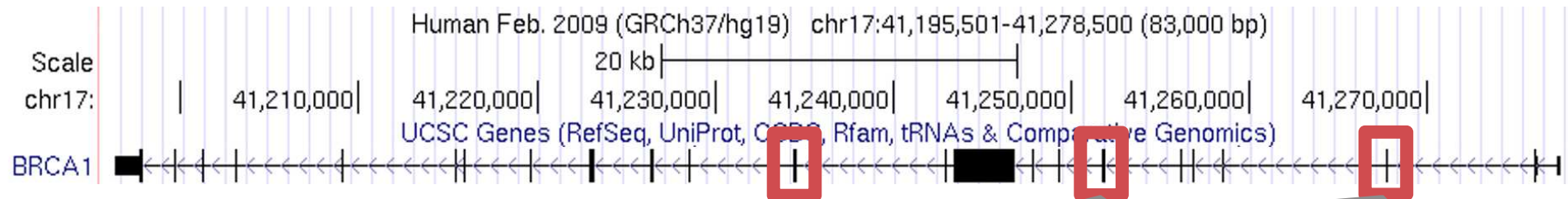
* 上記のURLまた、MSR Theory of Operationsドキュメントにも説明がございます

gVCFにおける情報圧縮

- 変異が検出されなかった領域において情報圧縮を適用している
- 情報圧縮はBlock compression と呼んでいるもので、同程度のポジションの連続とみなすことができるブロックをつくり、1行に圧縮して記述
- gVCF はbamファイルと比較して概ね~100X の容量



gVCF 例: *BRCA1*



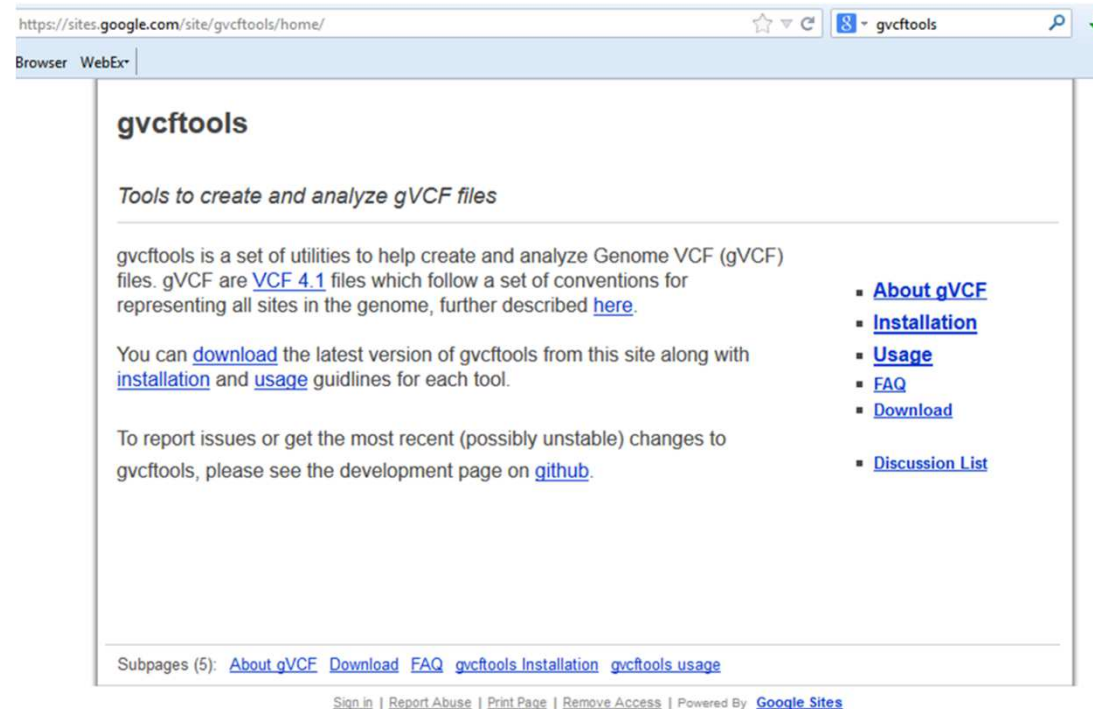
```
chr17 41267635 . . . . PASS END=41267754;min30p3a GT:GQ:DPU:DPF 0/0:152:61:0
chr17 41267755 . . . . PASS END=41268125;min30p3a GT:GQ:DPU:DPF 0/0:180:69:0
```

```
chr17 41251767 . . . . PASS END=41251814;min30p3a GT:GQ:DPU:DPF 0/0:117:44:0
chr17 41251815 . . . . PASS END=41251880;min30p3a GT:GQ:DPU:DPF 0/0:158:60:0
chr17 41251881 . . . . PASS END=41251973;min30p3a GT:GQ:DPU:DPF 0/0:174:67:0
```

```
chr17 41234248 . . . . PASS END=41234451;min30p3a GT:GQ:DPU:DPF 0/0:161:63:0
chr17 41234452 . . . . PASS END=41234469;min30p3a GT:GQ:DPU:DPF 0/0:153:64:2
chr17 41234470 . A G 261 PASS VARTYPE_SNV GT:GQ:DPU:DPF:AU 0/1:261:62:2:32,0,30,0
chr17 41234471 . . . . PASS END=41234728;min30p3a GT:GQ:DPU:DPF 0/0:163:63:0
```

Appendix: gvcftools を独自の解析系で使いたい場合

- ▶ オープンソースプロジェクト
- ▶ GATK へのプラグインとして使用し出力させる事ができる
- ▶ gVCF utility tools
- ▶ Trio の解析等に有用



<https://sites.google.com/site/gvcftools/home>

Enrichment

Enrichment

Nextera Rapid CaptureとTruSightパネル(Tumor以外)に対応。
リード配列はゲノム全体にアライメントされ、
変異コールは Manifestファイルで指定された領域のみ。

Nextera Rapid Capture と、 TruSight製品 (Tumor以外) に対応

*Tumor はAmplicon –DSワークフローをご利用下さい

- One
- Cancer
- Inherited Disease
- Autism
- Cardiomyopathy



One



Inherited
Disease



Autism



Cancer



Cardiomyopathy

- TruSightを購入の場合は**VariantStudio v2.1**アノテーションソフト
を合わせて使用可能(単年ライセンス)

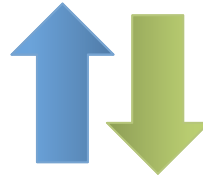
http://support.illumina.com/sequencing/sequencing_software/variantstudio.ilmn

VariantStudio v2.1

- アノテーションを加えるにはネットワークが必用
- アノテーションを加えるには認証にBaseSpaceアカウントが必要



VariantStudio
Annotation Database



Import VCF or gVCF Files

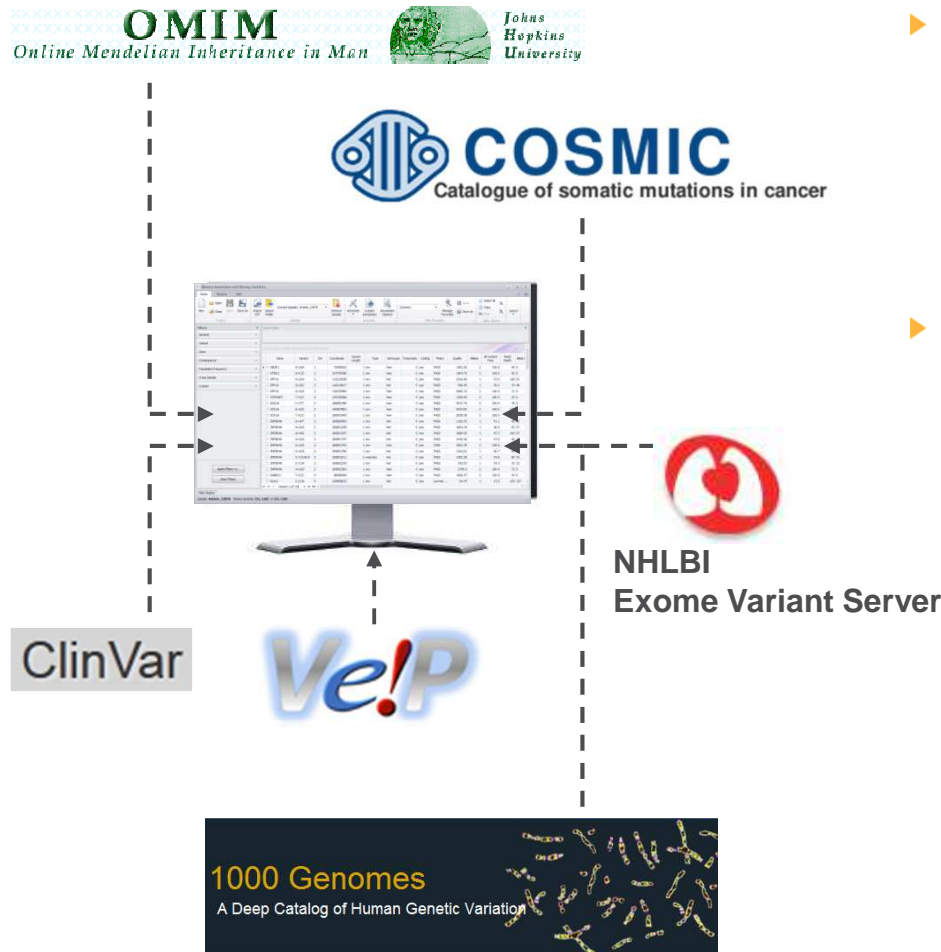


Export Report of interpreted variants

- イルミナシーケンサー出力のvcf v4.1 を入力ファイルとする

Illumina VariantStudio Desktop Client

Variant Studio v2.1 によるアノテーション



- ▶ 以下各レベルのアノテーション付加
 - Variant level
 - Transcript level
 - Gene level
- ▶ アノテーションのカテゴリ
 - Transcript consequence
 - Functional impact
 - Overlap with functional elements
 - Population allele frequencies
 - Disease association and description
 - Literature searches 等

Targeted RNA v2.2~

Targeted RNA

イルミナTargeted RNA Expression Kitライブラリ用。
マニフェストファイルにより定義したターゲット領域における
発現変動解析、二群間比較。



MiSeq Reporter



illumina®

サンプルシートの作成

- ▶ ワークフロー名 ; Targeted RNA
- ▶ 供与のマニフェストファイル名を記入する
- ▶ マニフェストには発注時DesignStudioで選択したターゲットをカバーする一連のTargeted Oligo Poolが記載されている)

[Header]					
Customer Name	ILLUMINA, INC.				
Product Type	15032433				
Date Manufactured	25/10/2012				
Lot	35105				
DesignStudio ID	NA				
Target Plexity	212				
[Probes]					
Target Region Name	Target Region ID	Target ID	Species	Build ID	Chromosome
MPL1_2	MPL1_2.chr1.43815008	MPL1_2.chr1	Homo sap	hg19	chr1
NRAS1_7	NRAS1_7.chr1.1152565	NRAS1_7.chr1	Homo sap	hg19	chr1
NRAS8_13	NRAS8_13.chr1.115258	NRAS8_13.chr1	Homo sap	hg19	chr1
ALK1	ALK1.chr2.29432664.29	ALK1.chr2	Homo sap	hg19	chr2
ALK2	ALK2.chr2.29443695.29	ALK2.chr2	Homo sap	hg19	chr2
IDH1_1_2	IDH1_1_2.chr2.2091131	IDH1_1_2.chr2	Homo sap	hg19	chr2
ERBB4_1_2	ERBB4_1_2.chr2.212284	ERBB4_1_2.chr2	Homo sap	hg19	chr2

Targeted RNA マニフェストファイルの例

- ▶ ノーマライゼーション用の遺伝子を指定することができる (Normalize列)

Targeted RNA

イルミナTargeted RNA Expression Kitライブラリ用。
マニフェストファイルにより定義したターゲット領域における
発現変動解析、二群間比較。

サンプルシート

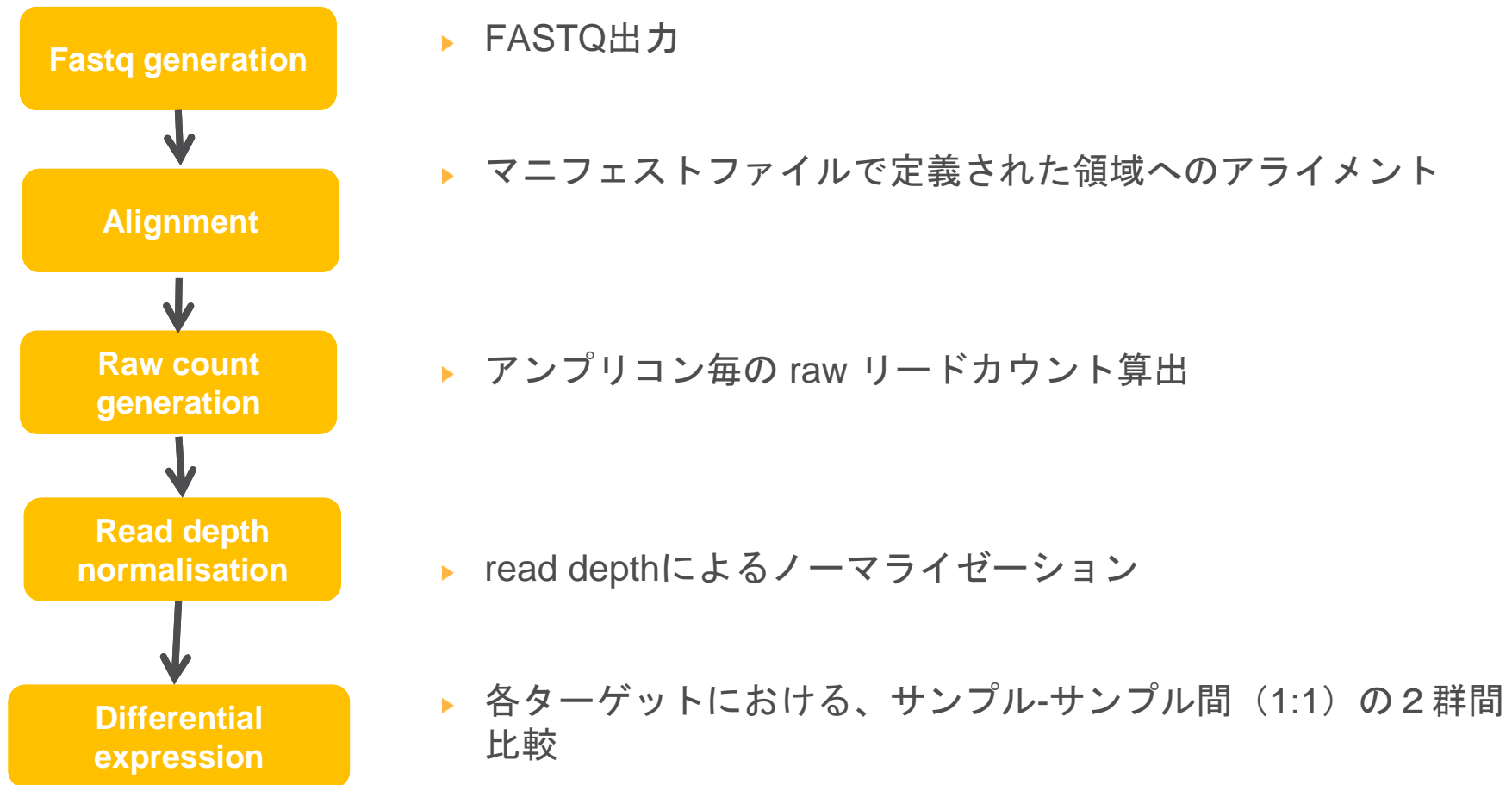


サンプルシート例と SampleID, Sample_Name の留意点

- ▶ マニフェストは“A,B,C....”と マニフェストファイル名
- ▶ 各サンプルIDで各サンプルが区別されます
- ▶ レプリケートは Sample name 列でグルーピングされる
- ▶ 発現変動解析は 2 群間解析で、“Sample Name 1” 対 “Sample Name 2” と 1 : 1 でなされる
- ▶ => SampleID と Sample_Name を注意して記入する必要がある

Date	3/19/2013						
Workflow	Targeted RNA Seq						
Assay	Targeted RNA Seq						
Description	TaqmanTissTitr_LW						
Chemistry	Amplicon						
[Manifests]							
A	TruSeq CRT Manifest TC0021987-CRT_Taqman						
[Reads]							
	50						
[Data]							
Sample_ID	Sample_Name	index	I7_Index_ID	index2	I5_Index_ID	Manifest	Normalize
UHR100-1	UHR100	ACAGTG	RPI5	TGAACCTT	A501	A	GAPDH
UHR100-2	UHR100	ACAGTG	RPI5	TGCTAAGT	A502	A	GAPDH
UHR100-3	UHR100	ACAGTG	RPI5	TGTTCTCT	A503	A	GAPDH
UHR10Br90-1	UHR10Br90	GCCAAT	RPI6	TGAACCTT	A501	A	GAPDH
UHR10Br90-2	UHR10Br90	GCCAAT	RPI6	TGCTAAGT	A502	A	GAPDH
UHR10Br90-3	UHR10Br90	GCCAAT	RPI6	TGTTCTCT	A503	A	GAPDH
UHR20Br80-1	UHR20Br80	CAGATC	RPI7	TGAACCTT	A501	A	GAPDH
UHR20Br80-2	UHR20Br80	CAGATC	RPI7	TGCTAAGT	A502	A	GAPDH
UHR20Br80-3	UHR20Br80	CAGATC	RPI7	TGTTCTCT	A503	A	GAPDH
UHR40Br60-1	UHR40Br60	ACTTGA	RPI8	TGAACCTT	A501	A	GAPDH
UHR40Br60-2	UHR40Br60	ACTTGA	RPI8	TGCTAAGT	A502	A	GAPDH
UHR40Br60-3	UHR40Br60	ACTTGA	RPI8	TGTTCTCT	A503	A	GAPDH
UHR80Br20-1	UHR80Br20	GATCAG	RPI9	TGAACCTT	A501	A	GAPDH
UHR80Br20-2	UHR80Br20	GATCAG	RPI9	TGCTAAGT	A502	A	GAPDH
UHR80Br20-3	UHR80Br20	GATCAG	RPI9	TGTTCTCT	A503	A	GAPDH
Br100-1	Br100	GGCTAC	RPI11	TGAACCTT	A501	A	GAPDH
Br100-2	Br100	GGCTAC	RPI11	TGCTAAGT	A502	A	GAPDH

ワークフローの流れ



* 解析手法は DESeq 1.8.3をベースとしている:

http://watson.nci.nih.gov/bioc_mirror/packages/2.10/bioc/html/DESeq.html

(Anders and Huber, Genome Biol., 11: R106, 2010)、

正規化法 : RLE法(relative log expression)

Targeted RNA ワークフロー

グラフィカルな表示以外にも、以下がファイルとして出力されているので必要に応じて利用できる

- **TargetHitsPerSample_M<N>.tsv** (./Data/Intensities/Basecalls/Alignment)

Raw カウント、レプリケート対アンプリコンのテーブル表記

- **TargetedRNASeqGeneExpression.tsv** (./Data/Intensities/Basecalls/Alignment)

群間比較図の元データ

- ***.fastq.gz** (./Data/Intensities/Basecalls)

- ***.bam** (./Data/Intensities/Basecalls/Alignment)

アライメント結果を収めたファイル。リファレンスがゲノムではなく、

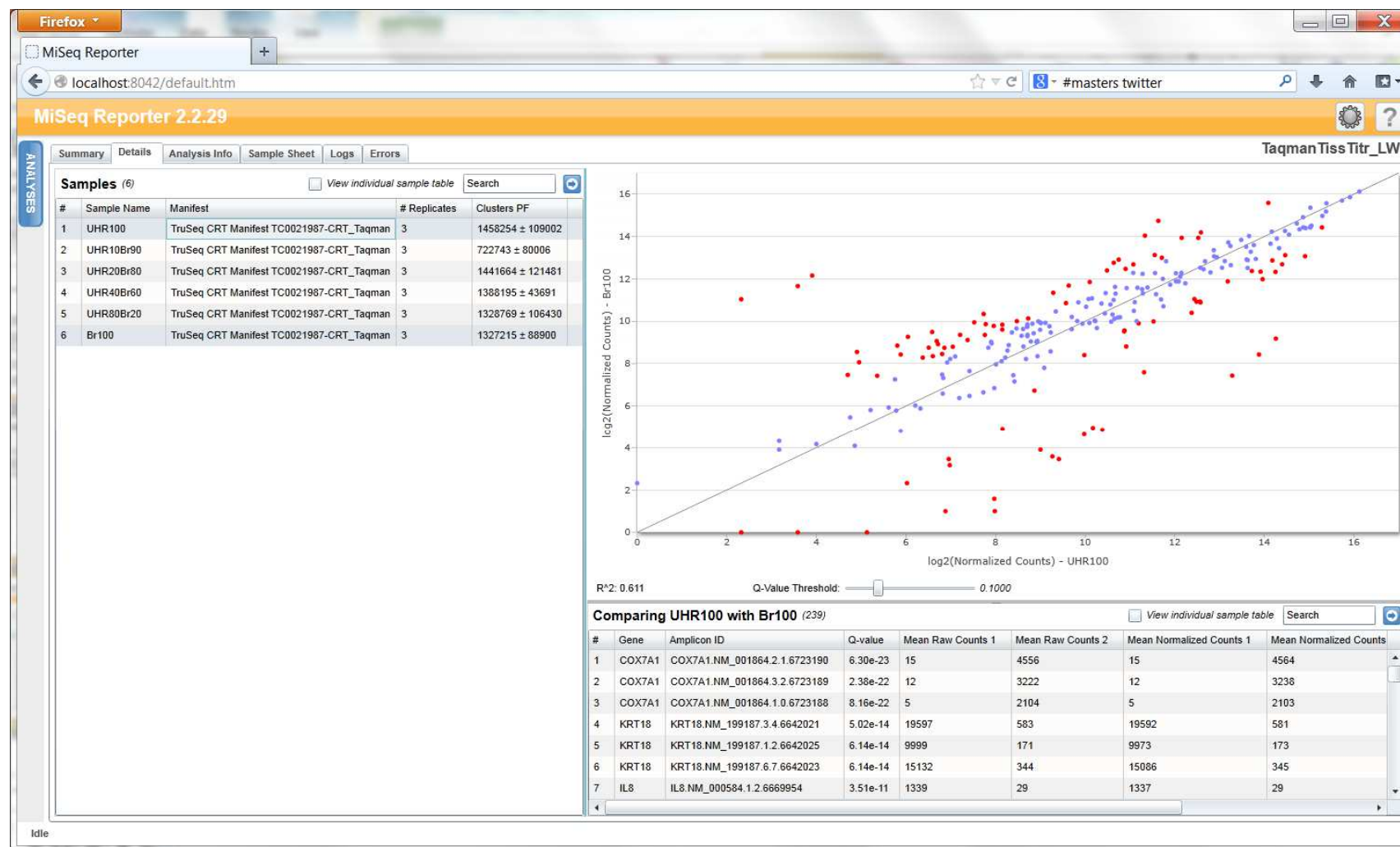
マニフェストのアンプリコンであることに注意する

- 詳細は、以下をご参考下さい。

MiSeq Reporter Targeted RNA Workflow Reference Guide (15042903 B)

MiSeq Reporter User Guide (rev B)

Targeted RNA ワークフロー



変異はコールされませんのでvcfはありません
ランのマージはできません

Amplicon-DS

Amplicon-DS(double stranded)

TruSight Tumor targeted assay専用のワークフロー
Manifestファイルで指定されたアンプリコン領域に対して
リードのアライメントを行い、somatic variant callerで
変異を検出する。FFPEサンプルの体細胞性変異検出に適する。



MiSeq Reporter



illumina®

Amplicon -DS ワークフロー



Tumor

- ・ TruSight Tumor パネルをお使いの場合に利用するワークフロー
- ・ Variant callは弊社のsomatic variant callerを利用
(to detect somatic variants below 5% frequency)
- Strand を認識する (DS: double stranded) ためストランド特異的な変異コールが利用できる
- 結果は標準のBAMファイルとVCFファイル

Technote : 「**Amplicon - DS Somatic Variant Caller**」 を必ずまずご発注前にご確認下さい。

- TruSightを購入の場合は**VariantStudio v2.1**アノテーションソフトを合わせて使用可能(単年ライセンス)

http://support.illumina.com/sequencing/sequencing_software/variantstudio.ilmn

Amplicon -DS ワークフロー

- ▶ Amplicon -DSワークフローではMSRで表示されないが、以下がAlignmentN配下に出力されているので必要に応じて確認できる

-AmpliconCoverage_Mx.tsv

マニフェスト中のアンプリコン毎・サンプル毎のdepthリスト

-.vcf, *.gvcf

変異コールの結果を収めたファイル

-.bam

アライメント結果を収めたファイル

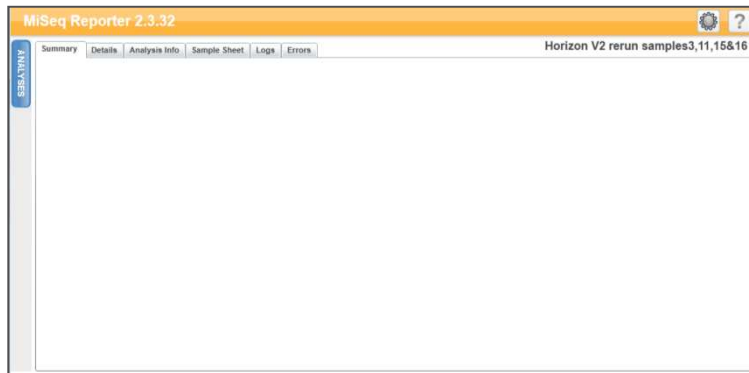
詳細は、以下をご参考下さい。

MiSeq Reporter Amplicon - DS Workflow Reference Guide (15042903 B)

MiSeq Reporter User Guide (rev B)

Amplicon -DS ワークフロー 留意点

Summary、およびDetailsタブ



➤ Summaryタブ



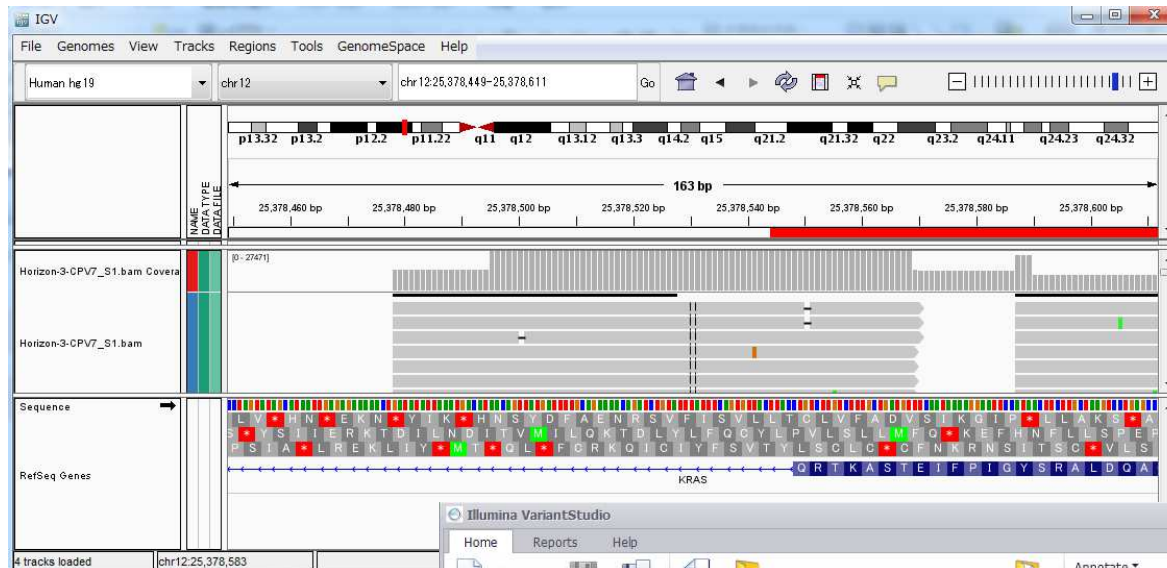
➤ Detailsタブ

* 現在、MSRにはAmplicon -DSワークフロー用の結果視覚化機能が備わっていないので、Summary、Detailsタブには左図のように何も表示されない

Alignment = **Smith-Waterman**
Variant calling = **Somatic variant caller**

Amplicon -DS ワークフロー

- 解析結果を視覚的に表示させたい場合は、Broad IGVや、illumina VariantStudio2.1を使用する。

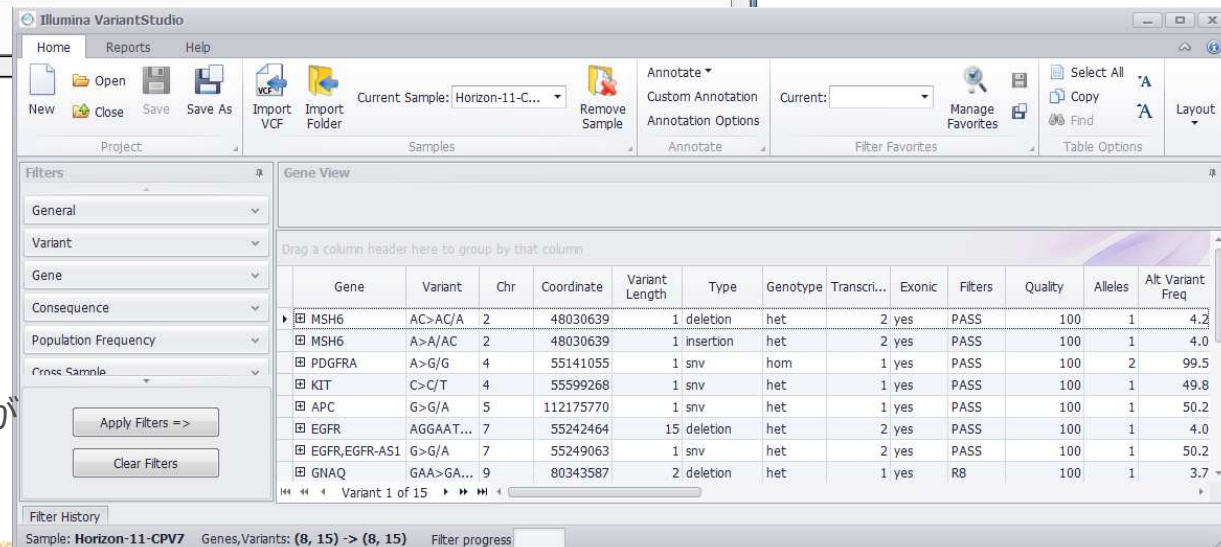


← アライメント結果の確認
(*.bamファイルの閲覧)



illumina
VariantStudio

VCFファイルを読み込ませ、
アノテーションを加えることが
可能



Stitch Read 機能について

- GenerateFASTQ、Amplicon-DS、TruSeqAmpliconワークフローにおいて、PE readに10塩基以上のオーバーラップがある場合、オーバーラップ部分を使用しリードペアを繋げて(Stitch)利用
 - GenerateFASTQ ワークフローでは、readは1つのFASTQとして出力され、stiching report も出力される。
 - Amplicon-DS、TruSeqAmplicon ワークフローでは、アライメント段階で内部的に利用されており、stitched read も中間ファイル扱いでFASTQとしては利用できない。またstitched readエントリとしてはbamファイルの書き出しに含まれない。
- デフォルトではオフ。
- Amplicon-DS、TruSeqAmpliconワークフローにおいては、通常の解析処理が終わり、お時間のある場合のみご自分のサンプルでオプション有との結果比較を行って頂き、結果の僅かな向上が認められるかをご確認下さい。結果として変わらない場合も多くあります。

(MSR v2.3.32 現在)

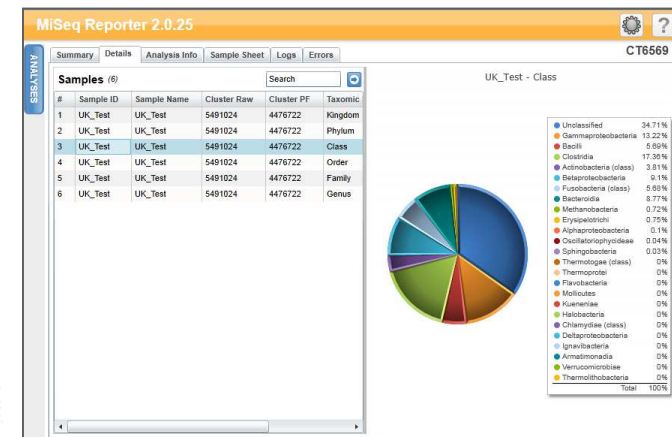
Metagenome

16S Metagenomics

16S rRNA bacterial 解析。Classification にはGreenGenesを利用。v2.3からアルゴリズムとdb バージョン変更



MiSeq Repor



Metagenomicsワークフロー

▶ v2.3からアルゴリズム変更

アルゴリズムは、Wang et al, 2007 で報告されたRDP classifier アルゴリズムをベースとしたものになります。
これまでより高速なアルゴリズムが実装され、解析時間が以前より短くなります。

* RDP アルゴリズムについては、以下をご参考下さい。
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950982/>

▶ Green Genes分類データベースの更新

- 分類に使用するGreen Genes のデータベースを更新し、Green Genes 13.5 (May 2013) が解析に使用されます。
- これまでの genus レベルまでの分類に加え、本バージョンからは、[species レベルまでの分類ができるようになりました。](#)デフォルトではspecies レベルまでの分類を行います。

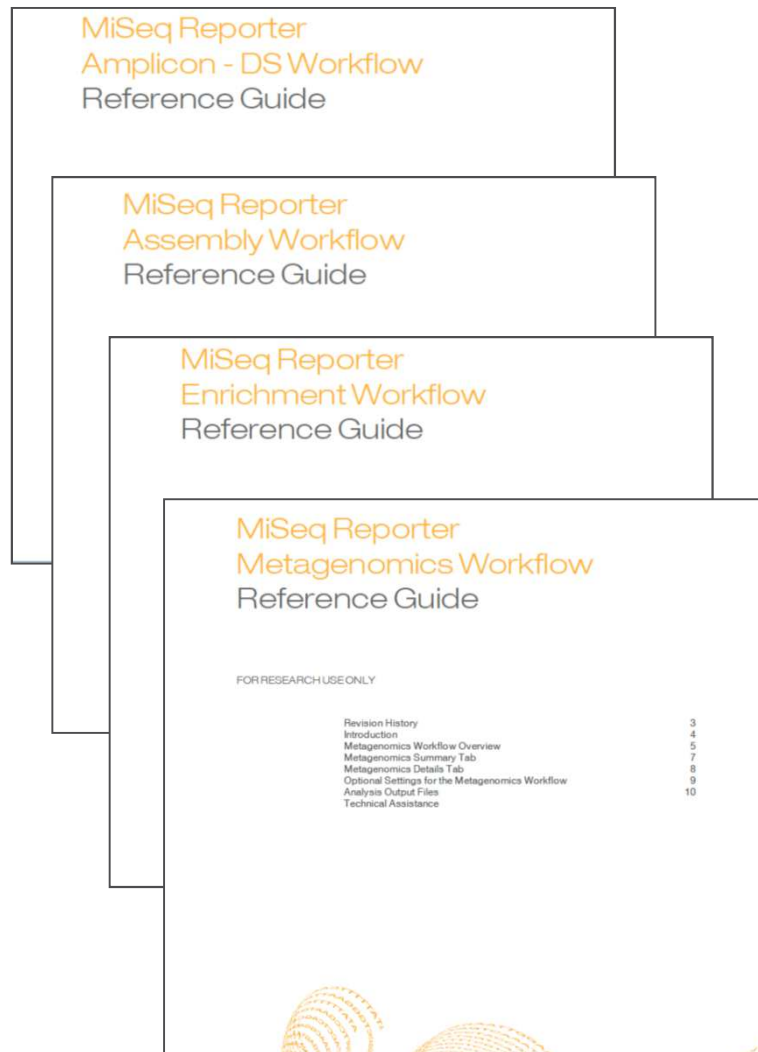
* 従来通りgenus までを望まれる場合は（speceis までと比較し早く解析が終わります）、別途サンプルシートの [Settings] の項目でご指定いただく必要がございます。**TaxonomyFile, gg_13_5_genus_32bp.dat** を入力下さい。

* 本バージョンにてgenus レベルまでの分類で解析した結果は、以前のバージョンのMSR の結果と近くはありますが、厳密に同じではありません。

MSR v2.3のパラメータ設定について ～IEMとバージョンの整合

- ▶ 現在最新のIEM であるIEMv1.6は、MSR v2.3リリース以前にリリースされたものであるため、MSR v2.3から追加になった機能をIEMから設定できることはありません。
- ▶ Notepadなどでサンプルシートを開き手動での編集が必要となります。

MSR ドキュメントが各ワークフロー毎に分割されました



MiSeq Reporter

Latest Updates

Third-Party Analysis Software and Utilities Tech Note 05/01/2013
Understanding Illumina Quality Scores 12/14/2012
Somatic Variant Caller Technical Note 11/13/2012



Guides

MiSeq Reporter Amplicon - DS Workflow Reference Guide (15042903 B)
MiSeq Reporter Assembly Workflow Reference Guide (15042313 B)
MiSeq Reporter Enrichment Workflow Reference Guide (15042315 B)
MiSeq Reporter Generate FASTQ Workflow Reference Guide (15042322 B)
MiSeq Reporter Library QC Workflow Reference Guide (15042316 B)
MiSeq Reporter Metagenomics Workflow Reference Guide (15042317 B)
MiSeq Reporter PCR Amplicon Workflow Reference Guide (15042318 B)
MiSeq Reporter Resequencing Workflow Reference Guide (15042319 B)
MiSeq Reporter Small RNA Workflow Reference Guide (15042320 A)
MiSeq Reporter Targeted RNA Workflow Reference Guide (15042321 A)
MiSeq Reporter TruSeq Amplicon Workflow Reference Guide (15042314 B)

http://support.illumina.com/sequencing/sequencing_software/miseg_reporter.ilmn

- 上記URLからワークフローごとのReference Guideを配布している
- 弊社から提供している資料の中では、個々のワークフローについて最も詳細に内容を記載
- データ解析の詳細について知りたい場合に参照

解析内容のトラブルシュート時にまず頂きたいファイル

ランフォルダ直下の

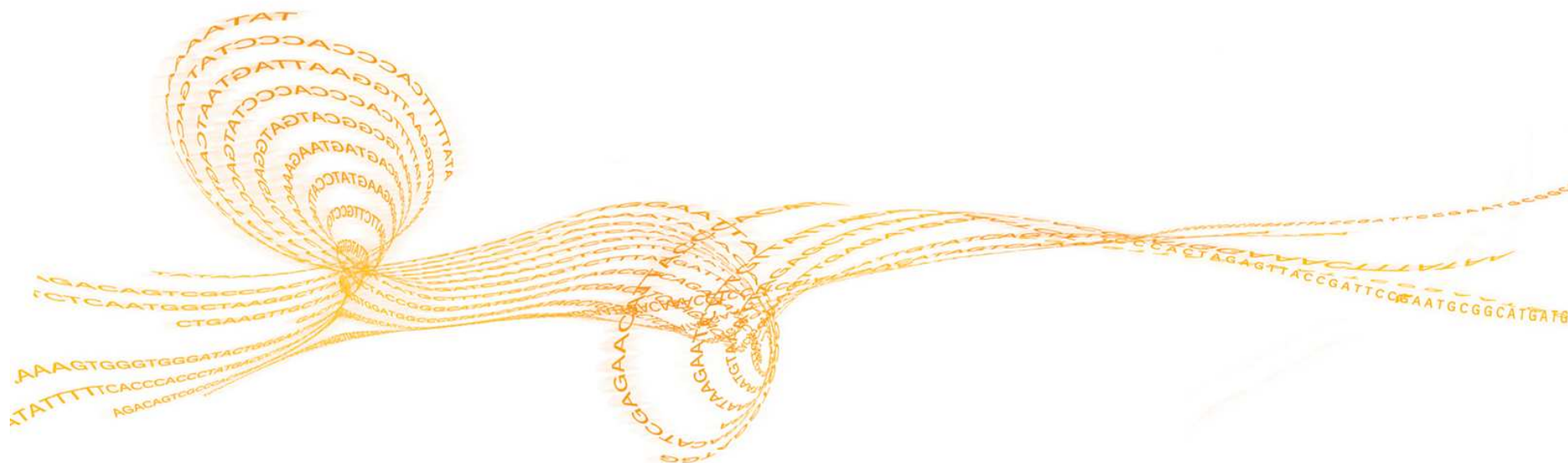
- ▶ AnalysisLog.txt
- ▶ AnalysisError.txt
- ▶ RunParameters.xml
- ▶ RunInfo.xml
- ▶ SampleSheet.csv

*BaseSpaceをご利用の場合はシェアいただければ送付の必要はございません。

*MSR自体が動かない、という場合はまず以下をご確認下さい。

http://www.illumina.co.jp/events/webinar_japan.ilmn?ws=ss

2012/06/15 「MiSeq Reporterをはじめよう」



ご清聴ありがとうございました。

ご質問はtechsupport@illumina.comでも承ります。