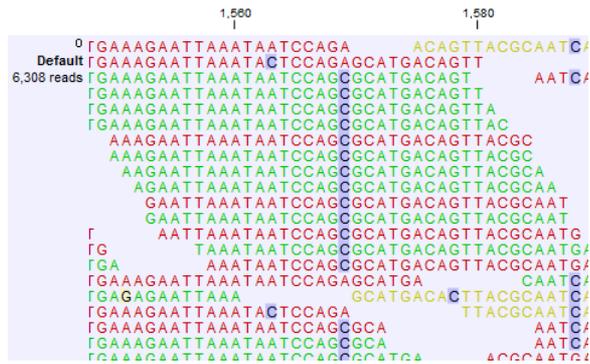


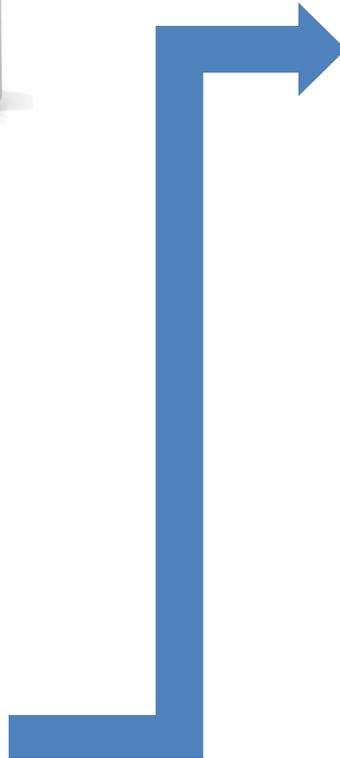
NGSデータ解析入門Webセミナー： 変異解析編



✓ シーケンス



✓ マッピング

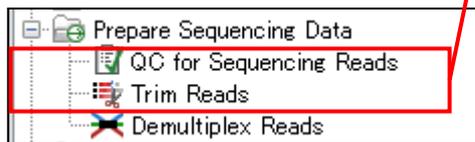
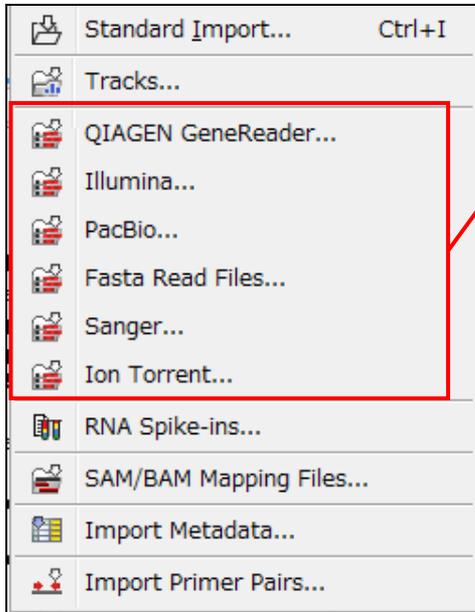


Chromoso...	Region	Type	Reference	Allele
20	76962	SNV	T	C
20	126311..126312	Deletion	CC	-
20	138125	SNV	G	T
20	139409	SNV	G	A
20	139456	SNV	G	A
20	139576	SNV	C	A
20	210061	SNV	G	A
20	256727	SNV	T	A
20	257733	SNV	A	G
20	257812	SNV	T	C
20	271135~271136	Insertion	-	C

✓ 変異検出



✓ データの精査・解釈



シーケンスデータのインポート

- NGS data import

クオリティチェック

- QC for Sequencing Reads
- Trim Reads

参照ゲノム配列へのマッピング・再アライメント

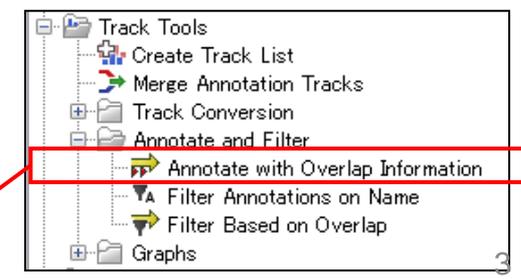
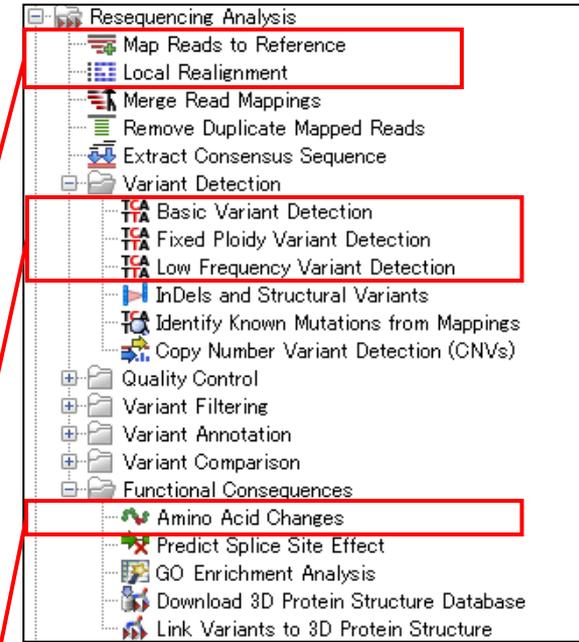
- Map Reads to Reference
- Local Realignment

変異検出

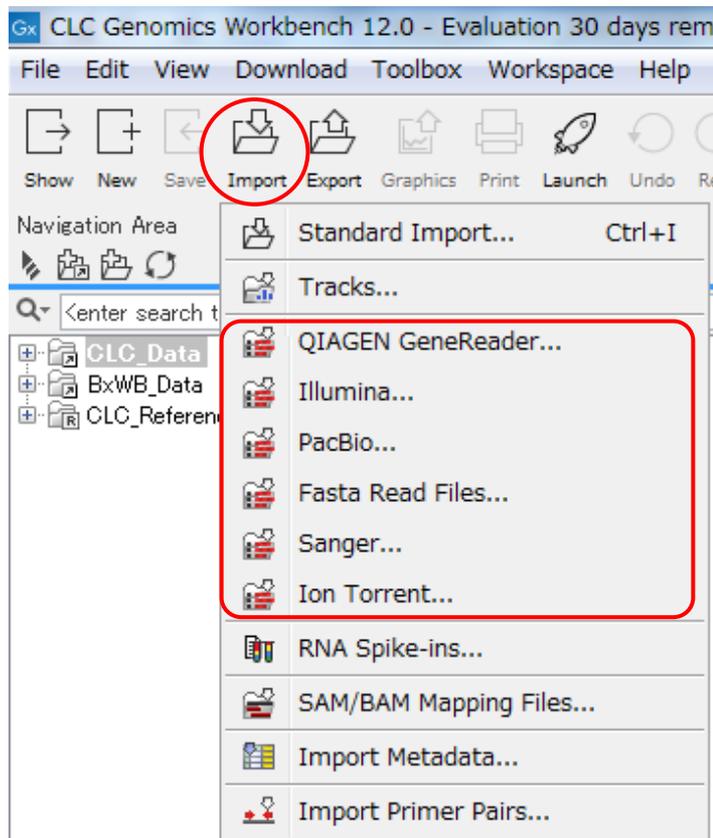
- Basic Variant Detection
- Fixed Ploidy Variant Detection
- Low Frequency Variant Detection

遺伝子名・アミノ酸配列置換情報の付加

- Amino Acid Changes
- Annotate with Overlap Information



- ✓ CLC Genomics Workbenchでは、シーケンサー機種やファイルフォーマットに合わせたインポートメニューを利用可能
- ✓ ToolbarのImportアイコンから表示されるインポーターから選択して、インポートを実行



プラットフォーム	ファイル形式
QIAGEN	.fastq .fq
Illumina	.txt .fastq .fq .qseq
PacBio	.bas.h5/ .bax.h5 .fastq .fq .fasta .fa .fna
Ion Torrent	.sam .bam .sff .fastq .fq

The screenshot shows the CLC Genomics Workbench 12.0 interface. The 'Standard Import...' menu is open, and the 'Illumina...' option is selected. A dialog box titled 'Illumina High-Throughput Sequencing Import' is displayed, showing the file selection process. The dialog box includes a file list with two files selected: 'Miseq_Data_110721_PF_R1.fastq.gz' and 'Miseq_Data_110721_PF_R2.fastq.gz'. The dialog box also shows settings for paired-end reads, including 'Paired-end (forward-reverse)' selected, 'Minimum distance' set to 100, and 'Maximum distance' set to 500. The 'Quality scores' are set to 'NCBI/Sanger or Illumina Pipeline 1.8 and later'.

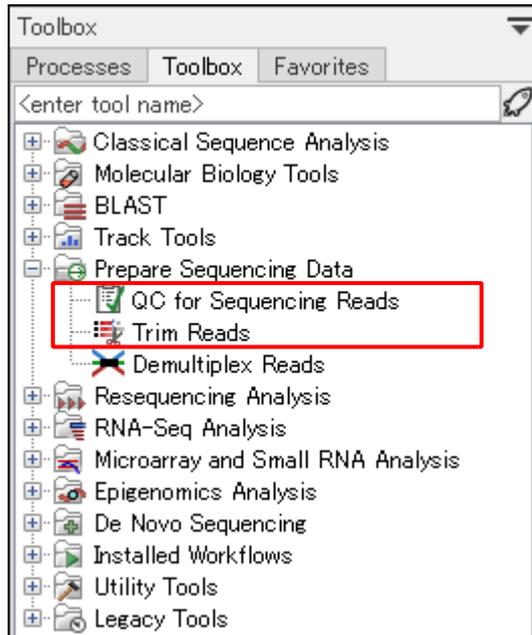
- ✓ シークエンサー機種などに合わせてメニューを選択し、シーケンスタデータファイルを選択
- ✓ ペアエンドシーケンスタデータのインポートにも対応

High-Throughput Sequencing Import

The screenshot displays the CLC Genomics Workbench 12.0 interface. The main window shows a list of sequence files, with 's_G1_L001_R1_001.fastq (paired)' selected and highlighted in red. The central panel displays a detailed view of the selected file, showing sequence reads with their corresponding quality scores. The reads are displayed in a grid format, with the sequence and quality scores for each read. The quality scores are represented by a bar chart below the sequence. The right panel shows the 'Sequence List Settings' dialog, which is used to configure the display of sequence data. The settings include 'Sequence layout' (No spacing), 'Numbers on sequences' (checked), 'Numbers on plus strand' (checked), 'Lock labels' (checked), and 'Show annotations' (checked). The 'Position' is set to 'Next to sequen...', 'Offset' is 'Little offset', and 'Label' is 'Stacked'. The 'Restriction sites', 'Motifs', 'Residue coloring', 'Nucleotide info', 'Find', and 'Text format' options are also visible.

- ✓ シークエンスデータがインポートされ、各種解析に使用できるようになる
- ✓ 各リードの塩基配列やクオリティスコアを確認できる

- ✓ インポートしたシーケンスデータに対して、クオリティチェックレポートの作成や、低クオリティリードの除去などを行う
- ✓ その他、重複リードの除去や、マルチプレックスシーケンス時のサンプルバーコードのソートなどの、各種データ前処理用ツールなども利用が可能

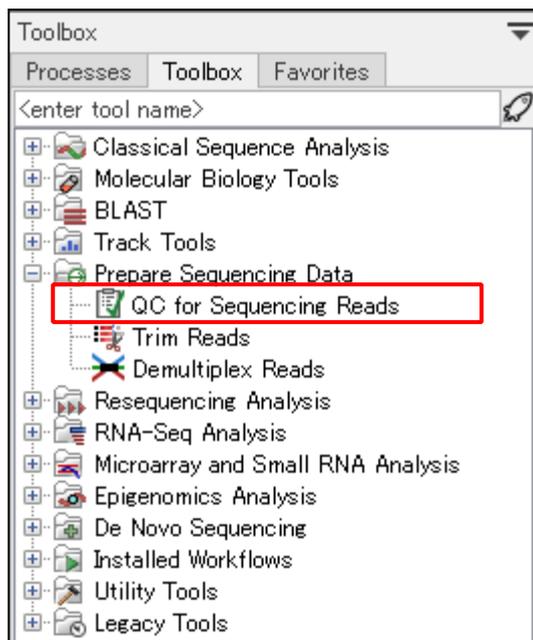


QC for Sequencing Reads

- インポートしたシーケンスデータのクオリティやPCR Duplicate の状況などを確認するためのレポートを作成

Trim Reads

- アダプターの除去、クオリティスコアによる除去、長さを指定した除去などを選択・組み合わせて、リードのトリミングを実行

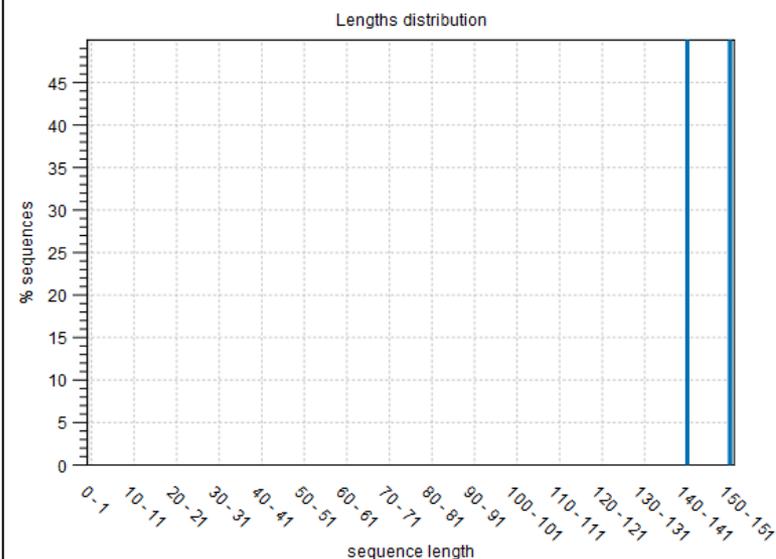


1 Summary

Creation date:	Sat Aug 11 21:37:50 JST 2018
Generated by:	Ozawa
Software:	CLC Genomics Workbench 11.0.1
Based upon:	1 data set
s_G1_L001_R1_001.fastq (paired):	16,386 sequences in pairs
Total sequences in data set	16,386 sequences
Total nucleotides in data set	2,384,163 nucleotides

2 Per-sequence analysis

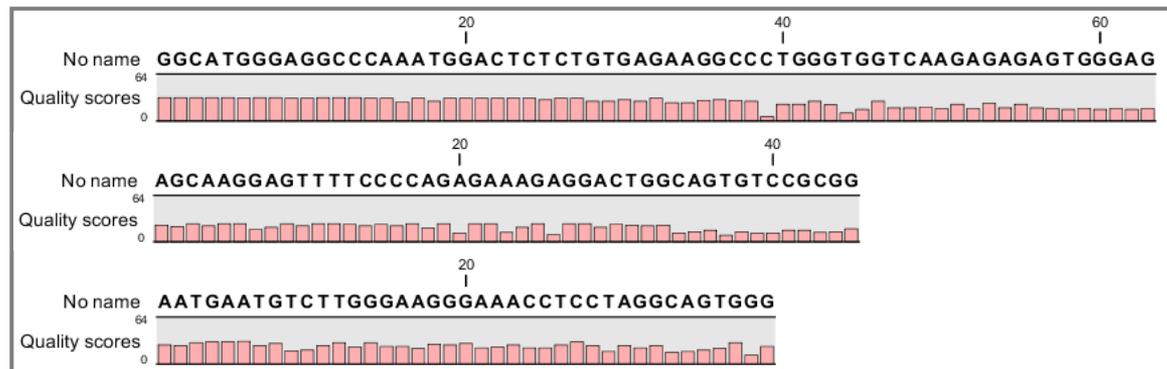
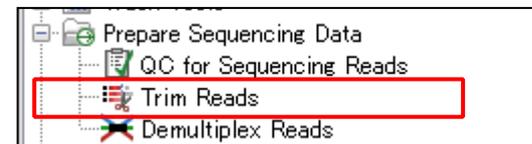
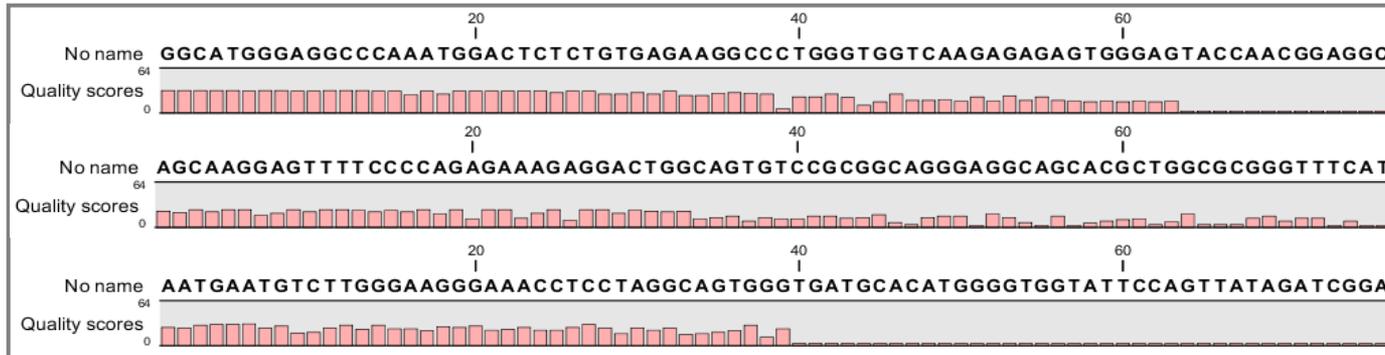
2.1 Lengths distribution



Distribution of sequence lengths. In cases of untrimmed Illumina or SOLiD reads it will just contain a single peak.
x: sequence length in base-pairs
y: number of sequences featuring a particular length normalized to the total number of sequences

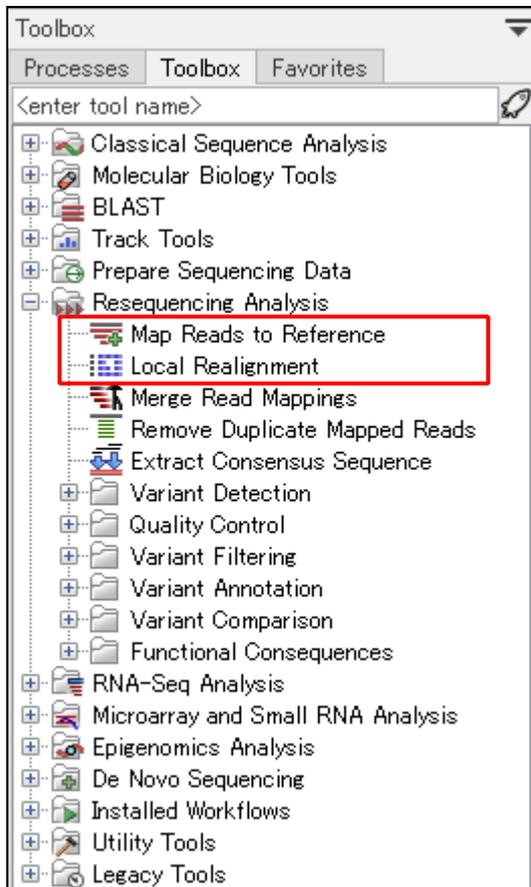
- ✓ QC for Sequencing Readsでは、シーケンスデータのクオリティ情報をまとめたレポートが作成される
- ✓ GC含量やクオリティスコア分布などのグラフデータや数値データを確認が可能

Trim Reads



- ✓ Trim Readsの使用により、各リードの低クオリティ部分がカットされる
- ✓ その他、アダプター配列の除去なども可能

- ✓ 変異の検出前に、シーケンスデータの各リードが、ゲノム上のどの部分を読んだものか、さらにどの部分で塩基が違っているかを調べるために、参照ゲノム配列へのリードのマッピングを行う
- ✓ サイズの大きいInsertion / Deletionを含んだリードは、マッピングが正確に行えない場合もあるので、必要に応じて再アライメントも行う



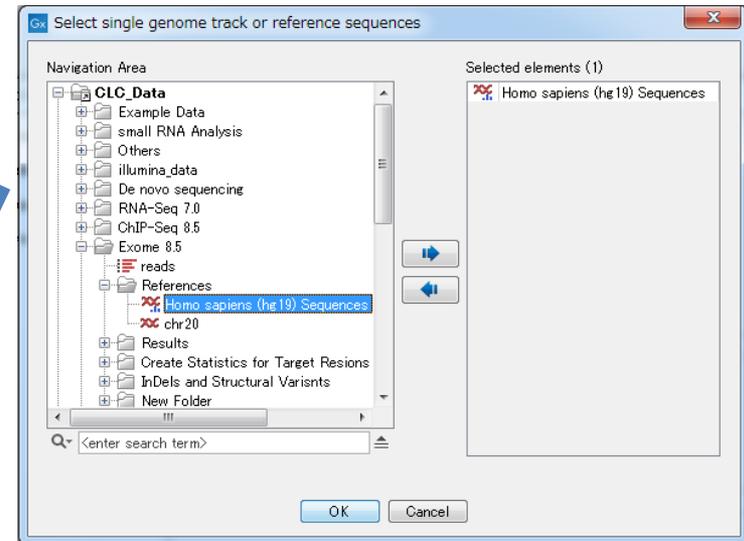
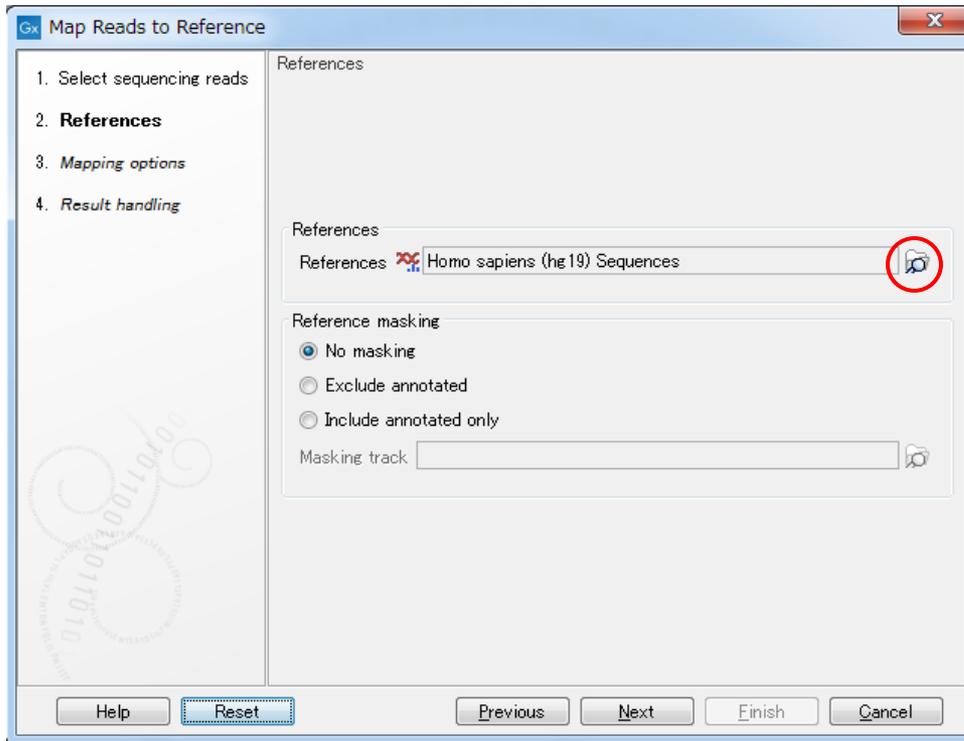
Map Reads to Reference

- 任意の参照ゲノム配列に対して、シーケンスデータのマッピングを行う

Local Realignment

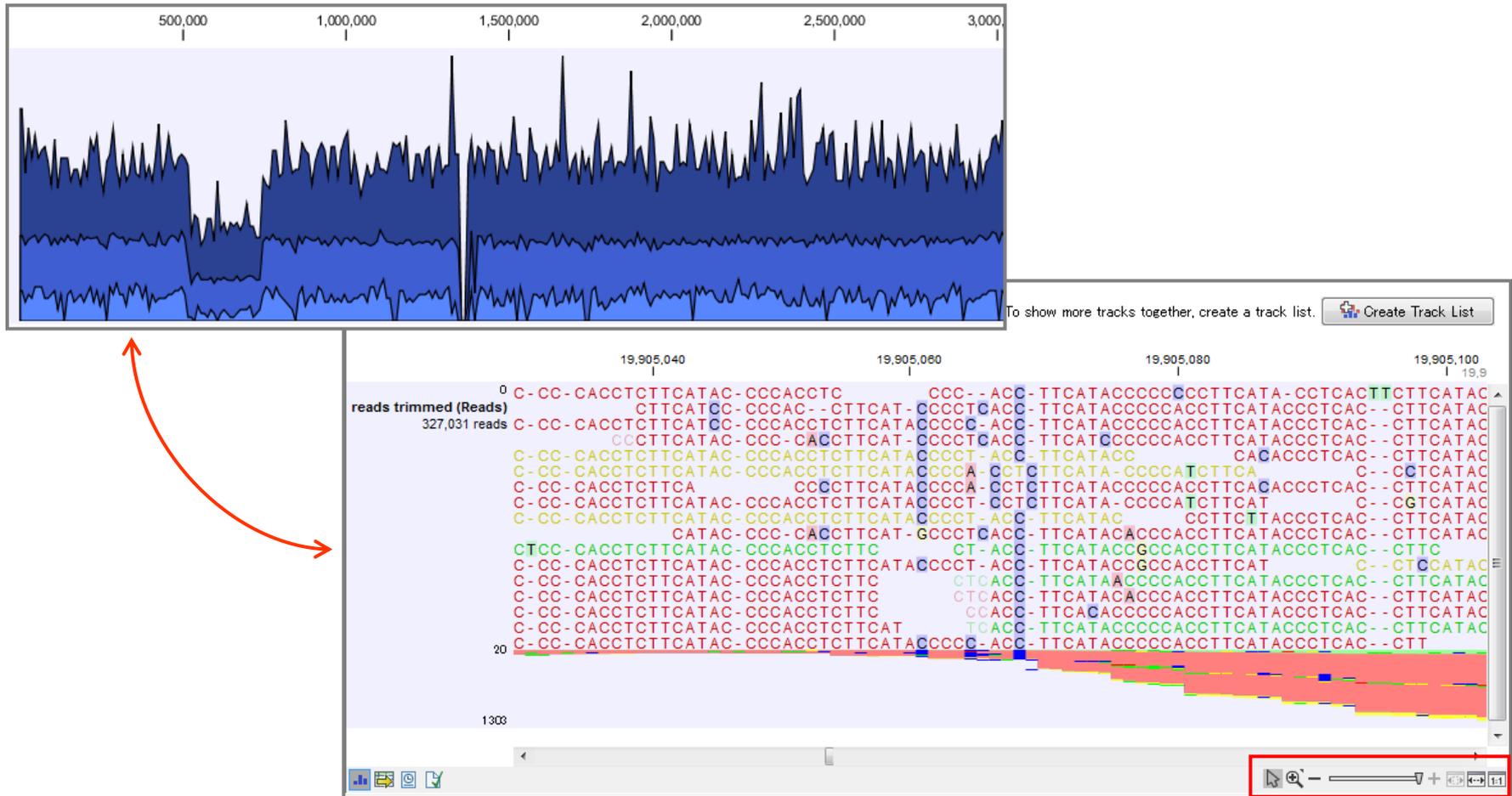
- すでにマッピングを実行したデータを使用し、リードの再アライメントを行う

Map Reads to Reference



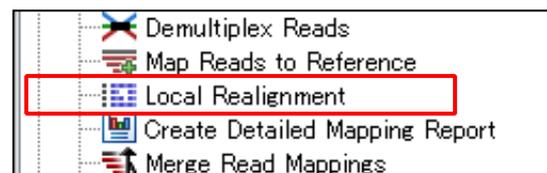
- ✓ Map Reads to Referenceでは、実行時のオプションパラメータで、任意の参照ゲノム配列データを選択が可能
- ✓ ヒト、マウス、ラットなどのモデル生物の参照ゲノム配列データは、ソフトウェア標準搭載のダウンロードツールから取得でき、その他NCBIに登録されている参照ゲノム配列データや、ユーザーカスタム作成の配列データを使用することも可能

Map Reads to Reference



- ✓ Map Reads to Referenceを実行すると、ゲノム配列（染色体）ごとに、マッピングされたリード配列がどの領域に多いのかを示したカバレッジグラフが表示される
- ✓ カバレッジグラフを拡大していくと、各リードの塩基配列も表示され、変異部位なども確認できる

Local Realignment前

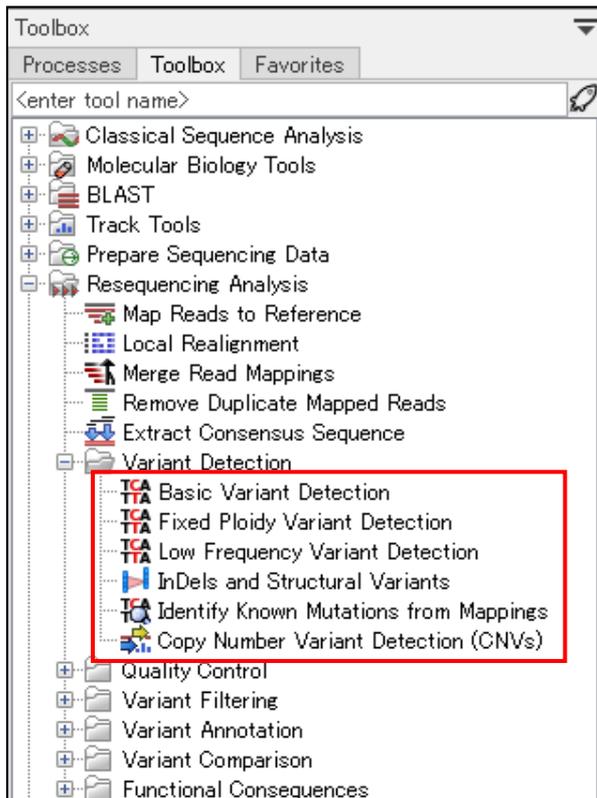


Local Realignment後



- ✓ Local Realignmentの使用により、すでにマッピングされた各リードに対して、同じ領域にマッピングされた周辺のリードに基づき、再アライメントが実行される
- ✓ サイズの大きいInsertion / Deletionの検出に対して有効

- ✓ マッピング・再アライメント後のデータより変異の検出を行うためのツールは6種類あり、それぞれ検出可能な変異の種類やアルゴリズムが異なっているが、そのうち3種類がSNVまたは小サイズの Insertion / Deletionの検出に用いられる
- ✓ 変異検出時の感度や特異度、さらにサンプルの種類などに合わせて、どのツールを用いるのかを判断する必要がある



Basic Variant Detection:

- 特殊な統計モデルを使用せずに、SNV, Small InDelを検出する
- パラメータの設定を調整することで、検出可能な変異に制限を設けずに解析が可能

Fixed Ploidy Variant Detection:

- 確率モデルを用いてSNV, Small Indelを検出する
- パラメータで指定したPloidy（倍数体）の値に基づいて変異の検出を行う
- カバレッジ中に低頻度（15%以下）で存在する変異は検出できない

Low Frequency Variant Detection:

- 確率モデルを用いてSNV, Small Indelを検出する
- カバレッジ中に低頻度で存在する変異の検出が可能

InDels and Structural Variants :

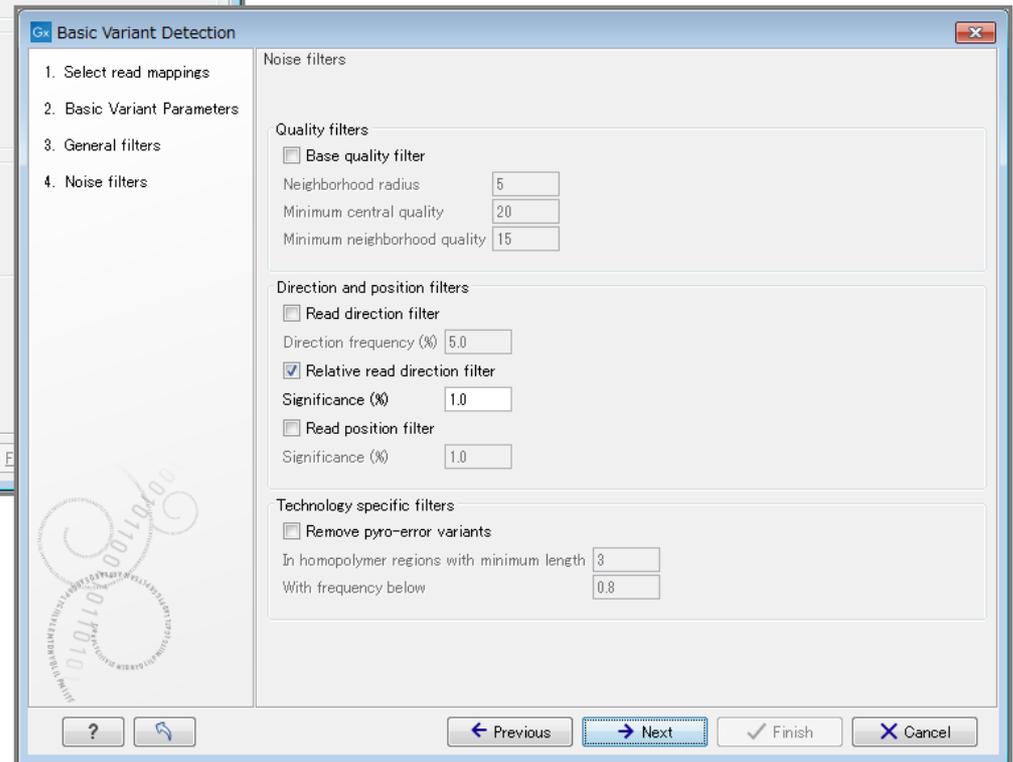
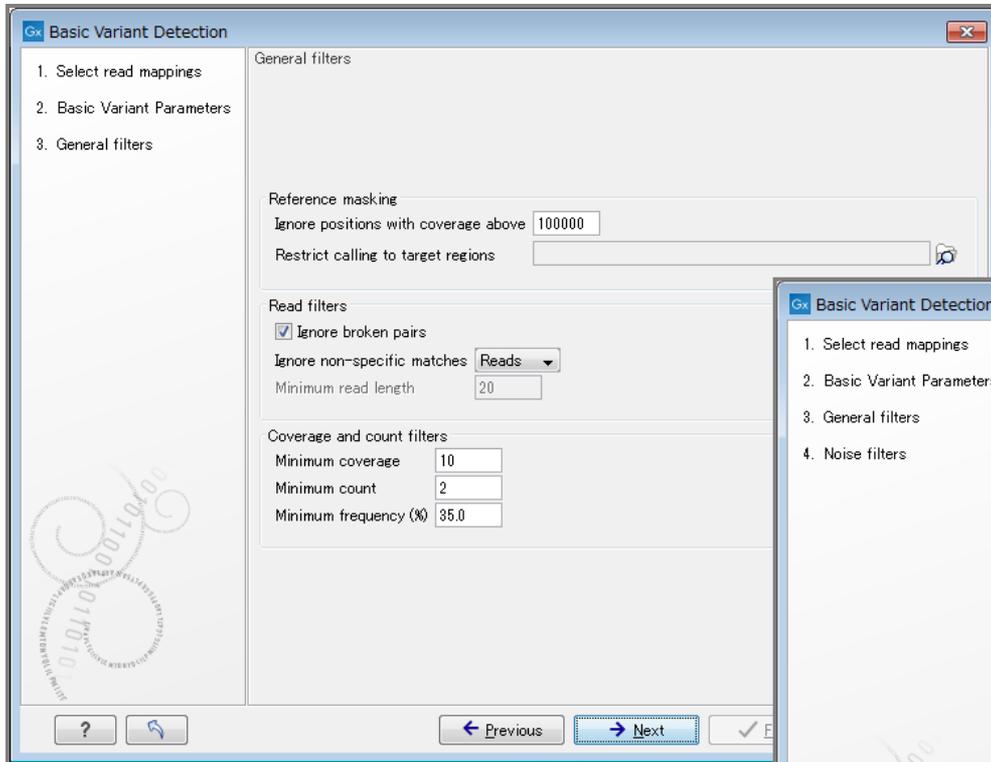
- Large InDelと染色体上の構造変化を検出する

Identify Known Mutations from Mappings:

- サンプルデータより、任意の変異データリストに含まれる変異を検出する
- 変異データリストに含まれない変異は検出できない

Copy Number Variant Detection (CNVs):

- コントロールサンプルとの比較により、CNVを検出する
- アノテーショントラックにより、ゲノム上のターゲット領域を指定する必要がある



- ✓ 3ツールともオプションパラメータで、変異をもつ最低リード数や頻度、リードのクオリティなどを設定が可能
- ✓ 遺伝子パネル解析の場合、変異検出の対象とする、ゲノム上のターゲット領域などを指定する

- Chromosome: 変異の検出された染色体番号
- Region: 変異のポジション
- Type: 変異の種類(SNV, Insertion, Deletionなど)
- Reference: リファレンスの塩基配列
- Allele: 検出された塩基配列

- Zygosity: 変異の接合性(HeteroかHomoか)
- Count: マップされたリードのうち、変異を有するリードの数
- Coverage: マップされたリード数
- Frequency: 変異の頻度

Rows: 19,104 Table view: Homo sapiens

Chromosome	Region	Type	Reference	Allele	Reference a...	Length	Zygosity	Count	Coverage	Frequency	Probability
1	887577	SNV	G	A	No	1	Heterozygous	25	44	56.82	1.00
1	898323	SNV	T	C	No	1	Homozygous	63	63	100.00	1.00
1	900505	SNV	G	C	No	1	Heterozygous	14	27	51.85	1.00
1	909309	SNV	T	C	No	1	Heterozygous	7	15	46.67	1.00
1	911916	SNV	C	T	No	1	Heterozygous	20	50	40.00	1.00
1	916662	SNV	A	C	No	1	Heterozygous	31	70	44.29	1.00
1	1007432	SNV	G	A	No	1	Heterozygous	11	18	61.11	1.00
1	1110987~1110988	Insertion	-	C	No	1	Heterozygous	22	51	43.14	1.00

- ✓ 検出された変異はテーブル形式で保存され、Excelなどにファイル出力が可能
- ✓ 変異テーブルには、変異の存在する位置や変異の種類、検出されたアレル、リード数などの情報が含まれる
- ✓ 変異テーブル上では、各テーブル項目に基づきフィルターをかけることができ、信頼性の高い変異のみを抽出することも可能

遺伝子名・アミノ酸配列置換情報の付加

- ✓ 変異テーブルには、変異部位に存在する遺伝子名や、変異によるアミノ酸置換などの情報を追加することが可能
- ✓ 変異データから、生物学的な解釈などを行う場合に、これらの情報が必要となる

 Amino Acid Changes 使用ツール
 Annotate with Overlap Information



変異データ					遺伝子名データ									
Probability	Forward re...	Reverse re...	Forward/re...	Average qu...	Exact match	Overlap	Coding regi...	Amino acid...	Other varia...	Non-synon...	chr20 (Gen...	Gene Cards	ENSEMBL	dbSNP
0.99	1	3	0.25	24.75	Homo sapie...	Homo sapie...			No	-				1022621
1.00	61	65	0.48	19.69	Homo sapie...	Homo sapie...	ENST00000...		No	No	SPTLC3	SPTLC3	ENSG00000171431433	
1.00	33	28	0.46	20.33					No	-	SPTLC3	SPTLC3	ENSG0000017...	
1.00	34	15	0.31	20.69	Homo sapie...	Homo sapie...	ENST00000...		No	-	SPTLC3	SPTLC3	ENSG00000176078899	
1.00	34	64	0.35	21.07	Homo sapie...	Homo sapie...	ENST00000...		No	-	SPTLC3	SPTLC3	ENSG0000017919604	
1.00	40	32	0.44	21.26	Homo sapie...	Homo sapie...	ENST00000...		No	-	SPTLC3	SPTLC3	ENSG0000017934335	
1.00	41	38	0.48	29.35	Homo sapie...	Homo sapie...	ENST00000...	ENSP00000...	No	Yes	SPTLC3	SPTLC3	ENSG0000017243887	
1.00	33	43	0.43	29.36	Homo sapie...	Homo sapie...	ENST00000...		No	No	SPTLC3	SPTLC3	ENSG0000017243888	
1.00	5	4	0.44	20.33					No	-	SPTLC3	SPTLC3	ENSG0000017...	
1.00	2	6	0.25	17.12	Homo sapie...	Homo sapie...	ENST00000...		No	-	SPTLC3	SPTLC3	ENSG00000172236127	
1.00	2	36	0.05	27.13	Homo sapie...	Homo sapie...	ENST00000...		No	-	SPTLC3	SPTLC3	ENSG0000017243874	
1.00	4	25	0.14	28.72		Homo sapie...			No	-	SPTLC3	SPTLC3	ENSG0000017...	
1.00	10	40	0.20	28.56			ENST00000...		No	-	SPTLC3	SPTLC3	ENSG0000017...	
1.00	6	3	0.33	23.00			ENST00000...		No	-	SPTLC3	SPTLC3	ENSG0000017...	
1.00	4	1	0.20	31.60					No	-	SPTLC3	SPTLC3	ENSG0000017...	
1.00	7	10	0.41	24.47			ENST00000...		No	-	SPTLC3	SPTLC3	ENSG0000017...	
1.00	3	3	0.50	27.17		Homo sapie...			No	-	TASP1, GAP...	GAPDHP2	ENSG0000008...	
1.00	2	2	0.50	21.00	Homo sapie...	Homo sapie...	ENST00000...		No	-	TASP1, GAP...	GAPDHP2	ENSG00000086105083	
1.00	11	1	0.08	27.67		Homo sapie...			No	-	TASP1	TASP1	ENSG0000008...	
1.00	10	1	0.09	27.45	Homo sapie...	Homo sapie...	ENST00000...		No	-	TASP1	TASP1	ENSG00000086033751	
1.00	11	2	0.15	23.08			ENST00000...		No	-	TASP1	TASP1	ENSG0000008...	
1.00	10	36	0.22	19.76	Homo sapie...	Homo sapie...	ENST00000...		No	-	TASP1	TASP1	ENSG00000082070307	
1.00	9	31	0.22	19.75		Homo sapie...			No	-	TASP1	TASP1	ENSG0000008...	
1.00	6	11	0.35	22.06		Homo sapie...			No	No	ESF1	ESF1	ENSG0000008...	
1.00	3	4	0.43	23.00	Homo sapie...	Homo sapie...	ENST00000...		No	No	ESF1	ESF1	ENSG00000086110019	
1.00	20	4	0.17	21.83	Homo sapie...	Homo sapie...	ENST00000...		No	-	ESF1	ESF1	ENSG00000084239710	
1.00	24	17	0.41	21.39	Homo sapie...	Homo sapie...	ENST00000...	ENSP00000...	No	Yes	ESF1	ESF1	ENSG00000083180370	
1.00	1	7	0.12	20.12			ENST00000...		No	-	NDUFAF5	NDUFAF5	ENSG0000010...	

Amino Acid Changes

1. Select variant track

2. **Set parameters**

3. Result handling

Set parameters

Select CDS track
CDS track

Select transcript (mRNA) track
mRNA track

Select sequence track
Sequence track

Filtering and annotation

Filter away synonymous variants

Filter away CDS regions with no variants

Genetic code

Help Reset Previous Next Finish Cancel

Coding region change	Amino acid change	Non-synony...
		-
		-
		-
YP_001729091.1:c.44delA	YP_001729091.1:p.Lys17fs	Yes
		No
YP_001729117.1:c.28delA	YP_001729117.1:p.Ile10fs	Yes
		No
YP_001729130.1:c.592delG	YP_001729130.1:p.Glu198fs	Yes
		No
		-
		-
YP_001729172.1:c.239T>A	YP_001729172.1:p.Leu80Gln	Yes
YP_001729176.1:c.36T>G		No
YP_001729185.1:c.80A>G	YP_001729185.1:p.Lys27Arg	Yes

Annotate with Overlap Information

1. Select a variant track, annotation track, expression track or statistical comparison track

2. **Overlap track**

3. Result handling

Overlap track

Select an annotation track for overlap comparison

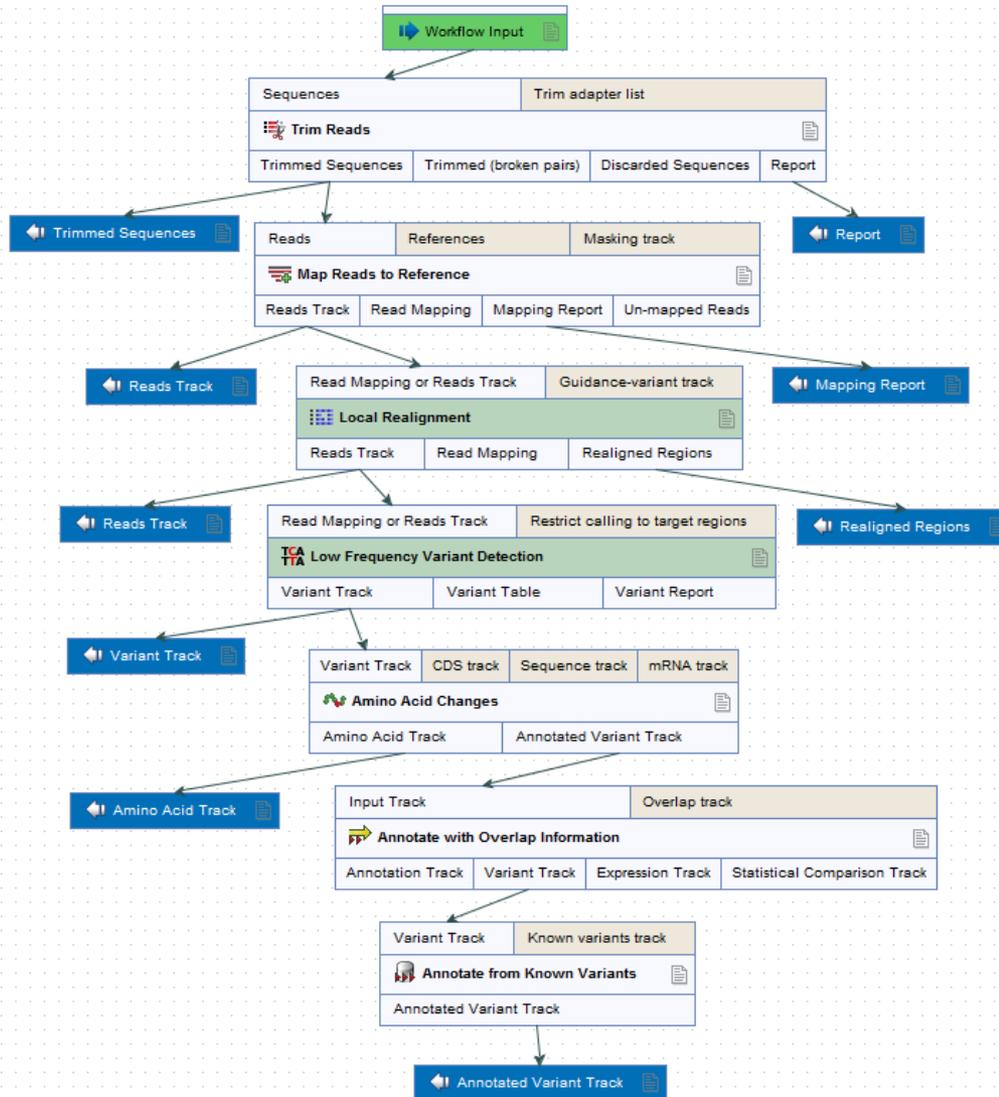
Overlap track

Help Reset Previous Next Finish Cancel

chr20 (Gene)	Gene Cards	ENSEMBL
ESF1	ESF1	ENSG00000089048
ESF1	ESF1	ENSG00000089048
ESF1	ESF1	ENSG00000089048
ESF1	ESF1	ENSG00000089048
ESF1	ESF1	ENSG00000089048
ESF1	ESF1	ENSG00000089048
ESF1	ESF1	ENSG00000089048
ESF1	ESF1	ENSG00000089048
NDUFAF5	NDUFAF5	ENSG00000101247
NDUFAF5	NDUFAF5	ENSG00000101247
NDUFAF5	NDUFAF5	ENSG00000101247
NDUFAF5	NDUFAF5	ENSG00000101247
NDUFAF5	NDUFAF5	ENSG00000101247
NDUFAF5	NDUFAF5	ENSG00000101247
NDUFAF5	NDUFAF5	ENSG00000101247
NDUFAF5	NDUFAF5	ENSG00000101247

✓ パラメータオプションで、遺伝子アノテーション情報データを指定して実行

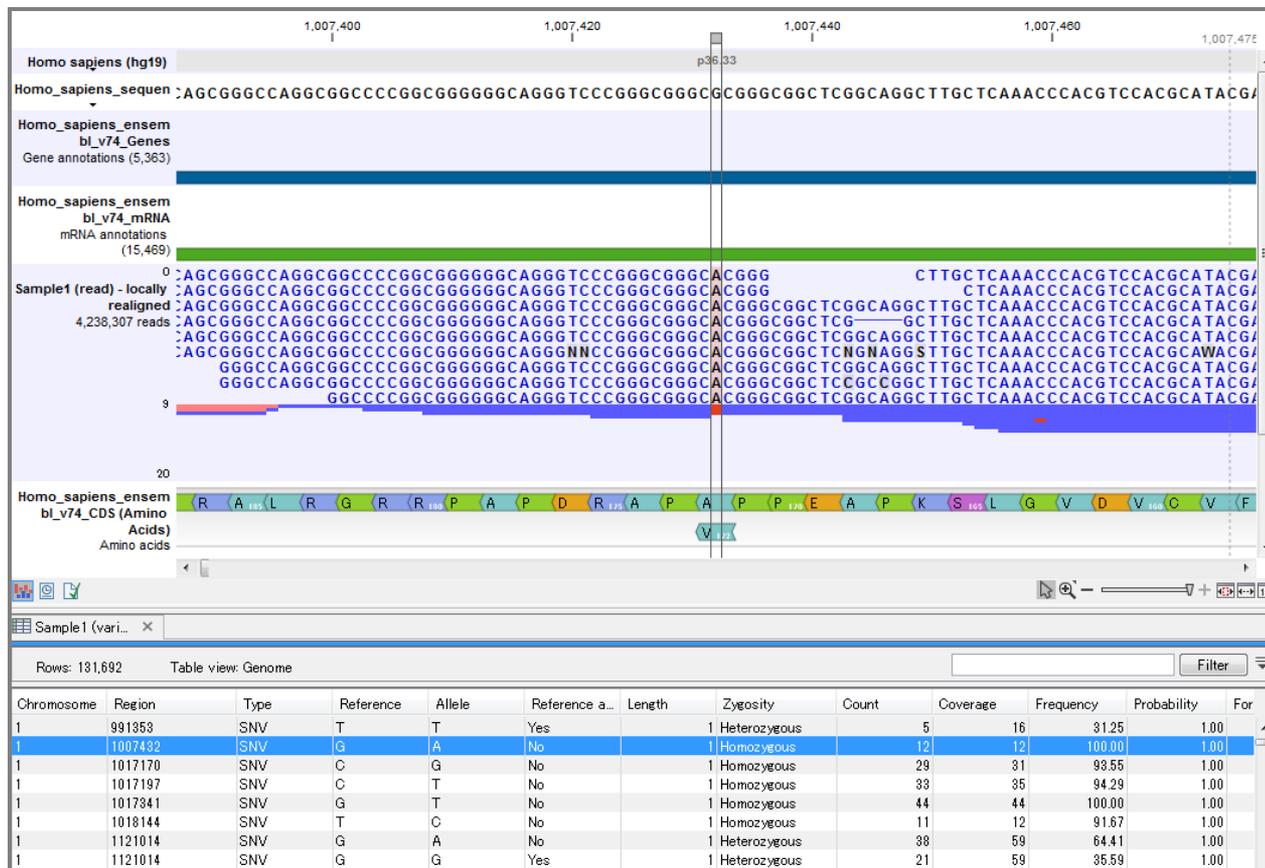
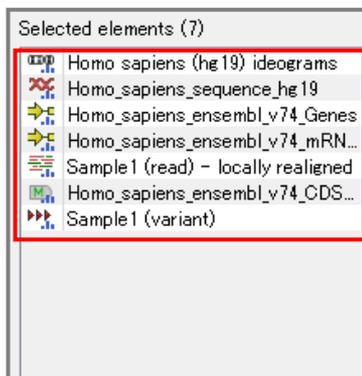
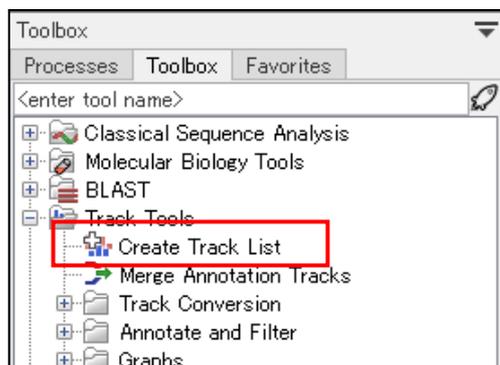
解析ワークフローの作成

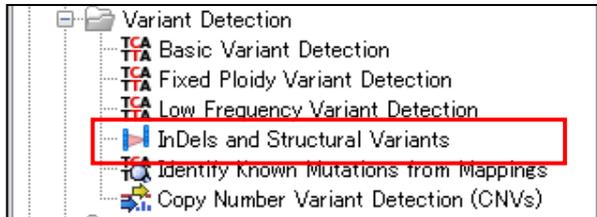


- ✓ 解析の自動化を行うために、各ツールの使用の順番をプログラムしたワークフローを作成することが可能
- ✓ サンプルが複数存在する場合は、全サンプルまとめてバッチ処理を行うことも可能

その他の変異解析用の機能

- ✓ 変異テーブルのみではなく、マッピングされたリードを直接目視で確認し、実際にリードに変異が含まれているか、などの確認のために用いられる
- ✓ マッピングや変異テーブルデータのみではなく、データベースのアノテーション情報なども同時に表示させ、各変異の生物学的な解釈を行うことも可能





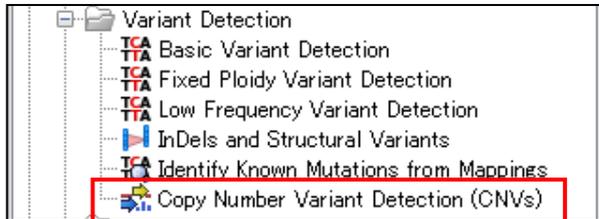
- ✓ 構造変化検出用ツールでは、サイズの大きな挿入・欠失と染色体上の構造異常のデータが出力される
- ✓ 参照ゲノム配列にリードをマッピングした際のUnalign end領域に基づいて検出を行うため、短いリードでは検出できないことがある

InDelsデータ

Chromosome	Region	Type	Reference	Allele	Reference...	Length	Zygoty	Evidence	Repeat	Variant ratio	# Reads	Sequence complexity
1	884100~884101	Insertion	-	CCCCTGGTCCCTTTGGCCCT...	No	41	Heterozygous	Self mapped		0.75	10	0.20
1	1260158..1260168	Deletion	GAGGGCGGGGC	-	No	11	Homozygous	Self mapped		1.00	1	0.45
1	1886501..1886518	Deletion	CGGCTCACACCGGAAGTG	-	No	18	Homozygous	Self mapped	CCGGAAGT...	1.00	1	0.51
1	3545196..3545214	Deletion	TCTGGGAGCTCCTCCCCCT	-	No	19	Homozygous	Self mapped		1.00	1	0.44
1	15439083~15439084	Insertion	-	AGGCTTTTTTTTTCCTCC...	No	27	Heterozygous	Paired breakpoint		0.80	4	0.04
1	16199095..16199099	Deletion	AAAAA	-	No	5	Homozygous	Self mapped	A	1.00	1	0.44
1	16776768..16776779	Deletion	TTGTTTTGTTTT	-	No	12	Homozygous	Self mapped		1.00	1	0.02
1	19071181~19071182	Insertion	-	GGAT	No	4	Heterozygous	Self mapped		0.50	2	0.50
1	20305466~20305467	Insertion	-	CACGCACACACACA	No	14	Homozygous	Self mapped	CA	1.00	1	0.31
1	23955212~23955213	Insertion	-	ATTATTATTATTATTATT	No	21	Homozygous	Self mapped		1.00	1	6.48E-3
1	36032231..36032236	Deletion	CACACA	-	No	6	Homozygous	Self mapped	CA	1.00	2	0.29
1	37967745..37967763	Deletion	GTTGACAGAGAAGATACAT	-	No	19	Homozygous	Self mapped		1.00	1	0.34

Structural Variantsデータ

Chromosome	Region	Name	Evidence	Length	Reference sequence	Variant sequence	Repeat	Variant ratio	# Signatures	Left breakpoints	Right breakpoints
1	join(2002490~2002491,20...	Complex	Can't resolve sequence	0				1.50	2	2002490	2002501
1	15438857..15439055	Replacement	Paired breakpoint	199	GTGTGTGTGTGTGTGTGGGAGCT...	CTGGGGAGCTGACTGAGCCTCCCC...		1.00	2	15438859	15438856
1	17301940..17302139	Replacement	Paired breakpoint	200	GGGAGCCCCGGACAGAACTGGCA...	TGGAGCCCCAGGAGTGGAGGAGGTC...		0.57	2	17301924	17301939
1	47652871..47652881	Replacement	Paired breakpoint	11	CTTAGCCCTCA	ACACA		0.21	2	47652879	47652870
1	92251569..92251668	Insertion	Tandem duplication	17785685		CCTCAGCCTCCCAAAGTGTGAGA...		0.40	2	92251568	110037253
1	94312556..94312588	Replacement	Paired breakpoint	33	CCGCGGCTGGGTCTGGGCGCGGGC...	GCCCGCGCCCGTCCGGGCGCGG...		1.00	2	94312584	94312554
1	109656089..109656288	Replacement	Paired breakpoint	205	AGCCCCACGTGCGGCGCTTGGGA...	GGAGAGAGAGAGAGAGAGAGAGG...		0.80	2	109656073	109656088
1	<110037154..110037253	Insertion	Tandem duplication	17785685		CCTCAGCCTCCCAAAGTGTGAGA...		0.40	2	92251568	110037253
1	152681695..152681894	Replacement	Paired breakpoint	204	TGGTTGCTGCAGCTGTGAGGGAG...	CTGTAGCTCTGGGGCGCTGTTG...		0.94	2	152681680	152681694



- ✓ CNV検出用ツールでは、ゲノム上の3種類の領域ごとに検出したCNVデータが出力される
- ✓ 解析時には、コントロールサンプルのマッピングデータと、遺伝子パネルなどのターゲット領域のアノテーショントラックが必要となる

Region level のCNVデータ

Chromosome	Region	Name	Minimum CNV length	p-value	Fold-change (adjusted)	Consequence	Number of tar...	Targets	Comments
14	73789815..73789934	mRNA AF171938,mRNA AF171939...	120	6.18E-5	-3.85	Loss	1	mRNA AF171938,mRNA AF171939...	Small region
14	101507093..101507212	ref MIR376A1,ref NR_029868...	120	9.05E-4	-3.65	Loss	1	ref MIR376A1,ref NR_029868,miRNA...	Small region
14	104201452..104201571	mRNA AK307149,mRNA BC04763...	120	8.50E-7	-3.32	Loss	1	mRNA AK307149,mRNA BC047639...	Small region

Target level のCNVデータ

Chromoso...	Region	Name	Target nu...	Case cov...	Baseline ...	Length	FDR-corr...	Fold-chan...	Region (joined target...	Regional f...	Regional ...	Comments
14	19377544..19378624	ref OR11H12,ccds CCDS32017...	1	72.67	71.95	1081	0.50	1.01	19377544..73783173	1.01	1.00	
14	19487336..19487453	-	2	4.38	4.16	118	NaN	1.05	19377544..73783173	1.01	1.00	Low coverage target
14	19495360..19495479	-	3	0.50	0.50	120	NaN	-1.00	19377544..73783173	1.01	1.00	Low coverage target
14	19553977..19553976	mRNA BC172559,mRNA BC1564...	4	272.10	242.94	600	0.22	1.12	19377544..73783173	1.01	1.00	
14	19558714..19558833	mRNA BC172559,mRNA BC1564...	5	12.17	8.55	120	NaN	1.42	19377544..73783173	1.01	1.00	Low coverage target
14	19558958..19559197	mRNA BC172559,mRNA BC1564...	6	0.50	0.50	240	NaN	-1.00	19377544..73783173	1.01	1.00	Low coverage target
14	19561987..19562106	mRNA BC172559,mRNA BC1564...	7	0.50	0.50	120	NaN	-1.00	19377544..73783173	1.01	1.00	Low coverage target

Gene level のCNVデータ

Chromosome	Region	Name	Region length	ENSEMBL	CNV region	CNV region length	Consequence	Fold-change...	p-value	Number...	Comments	Targets
14	complement(7378981...	NUMB	120	ENSG00000133961	73789815..73789934	120	Loss	-3.85	6.18E-5	1	Small region	mRNA AF171...
14	101507104..101507206	MIR376A1	103	ENSG00000264882	101507093..101507212	120	Loss	-3.65	9.05E-4	1	Small region	ref MIR376A1...
14	complement(1042014...	PPP1R13B	120	ENSG00000088808	104201452..104201571	120	Loss	-3.32	8.50E-7	1	Small region	mRNA AK307...

- ✓ 専用のダウンロードツールを使い、様々な生物種の参照ゲノムデータをダウンロード
- ✓ 生物種によっては、ゲノム配列や遺伝子アノテーションの他に、変異データベースも取得可能

The screenshot shows the CLC Genomics Workbench 12.0 interface. The 'Manage Reference Data' dialog box is open, displaying the 'Download Genomes' tab. The dialog is divided into several sections:

- Download Genomes (Public Repositories):** A list of species with download icons and descriptions. The selected species is 'Homo sapiens - hg19'.
- Download genome:** Options to 'Download genome sequence' (selected) or 'Use existing genome sequence track'.
- Download size for selection: 5.28 GB**
- Download details table:**

Name	Version	Provider	Size
<input checked="" type="checkbox"/> Sequence	75	Ensembl	828.0 MB
<input checked="" type="checkbox"/> Genome Annotations	75	Ensembl	37.5 MB
<input checked="" type="checkbox"/> Chromosome bands (ideogram)		UCSC	6 KB
<input checked="" type="checkbox"/> Dbsnp (common) variants	138	UCSC	620.7 MB
<input checked="" type="checkbox"/> Dbsnp variants	138	UCSC	1.78 GB
<input checked="" type="checkbox"/> Clinical variants in dbSNP		NCBI	16.8 MB
<input checked="" type="checkbox"/> HapMap Variants	75	Ensembl	513.3 MB
<input checked="" type="checkbox"/> 1000genomes	phase 1	Ensembl	1.53 GB

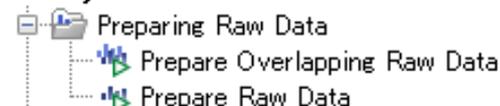
Below the table is a 'Previous downloads' section with a table header: Date, Size, Versions, Folder. A 'Delete' button is present.

The 'References' icon in the top toolbar is highlighted with a red dashed box, and the 'Manage Reference Data' dialog box is also highlighted with a red dashed box.

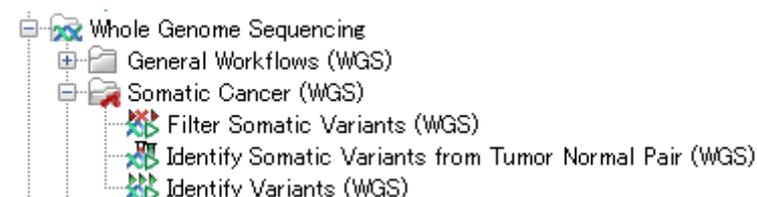
- ✓ Biomedical Genomics Analysisプラグインには、レディーメイドのワークフローが標準で搭載されており、これらワークフローを使用するだけで、各解析を一気に実行することが可能
- ✓ がん・遺伝性疾患用のワークフローがあり、腫瘍／正常ペアサンプル比較や、家族性のトリオサンプル解析を実行することも可能

クオリティチェック用ワークフロー :

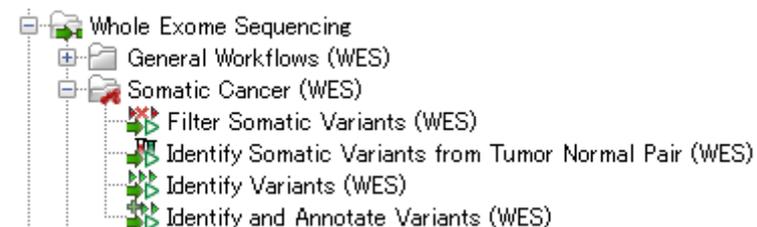
Ready-to-Use Workflows



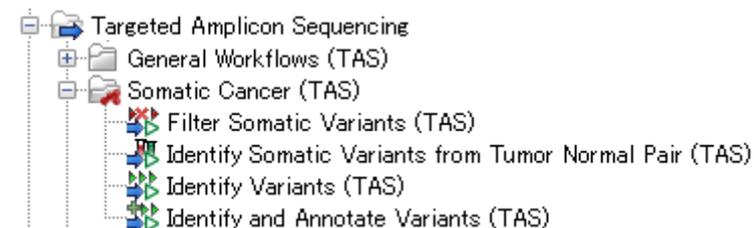
全ゲノムシーケンス解析用ワークフロー :



全エクソームシーケンス解析用ワークフロー :



ターゲットアンプリコンシーケンス解析用ワークフロー :



お問い合わせ先: フィルジェン株式会社

TEL 052-624-4388 (9:00~17:00)

FAX 052-624-4389

E-mail: biosupport@filgen.jp