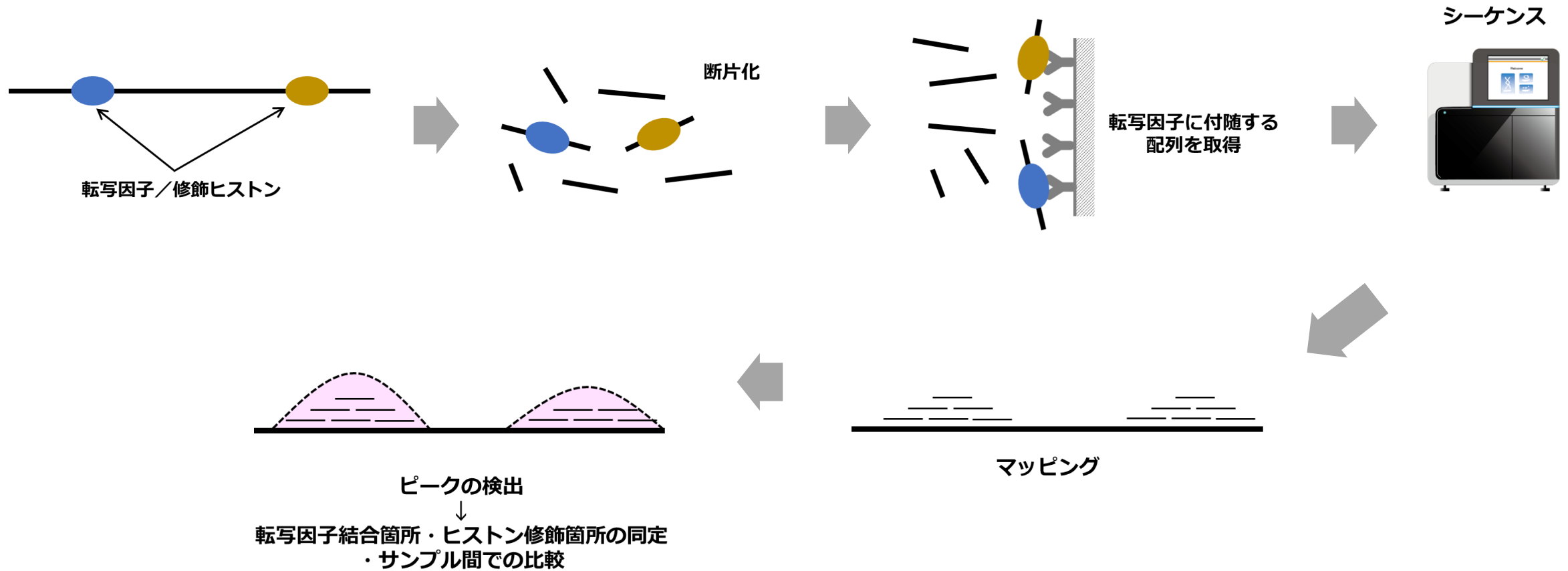


# CLC Genomics Workbenchを用いた ChIP-seq解析

フィルジェン株式会社 バイオインフォマティクス部  
( support@filgen.jp )



## シーケンスデータのインポート

Import

## クオリティチェック

QC for Sequencing Reads  
Trim Reads

## リファレンスへのマッピング

Map Reads to Reference

## ピークの検出

Transcription Factor ChIP-Seq  
Histone ChIP-Seq

## 遺伝子情報の付加

Annotate with Nearby Information

## シーケンスデータのインポート

Import

## クオリティチェック

QC for Sequencing Reads  
Trim Reads

## リファレンスへのマッピング

Map Reads to Reference

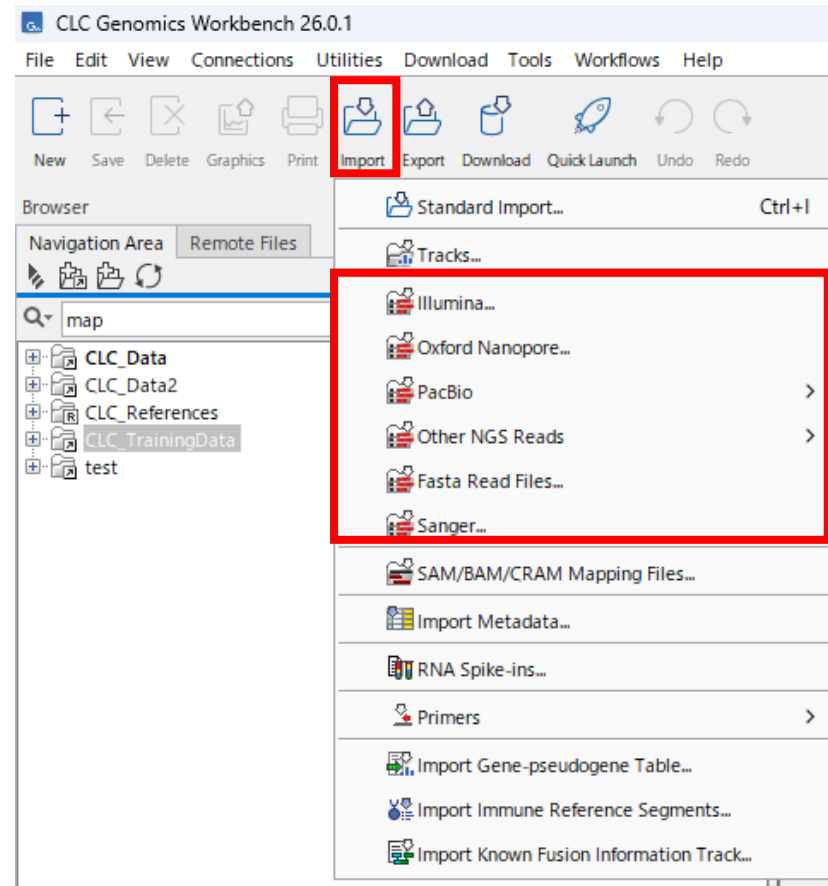
## ピークの検出

Transcription Factor ChIP-Seq  
Histone ChIP-Seq

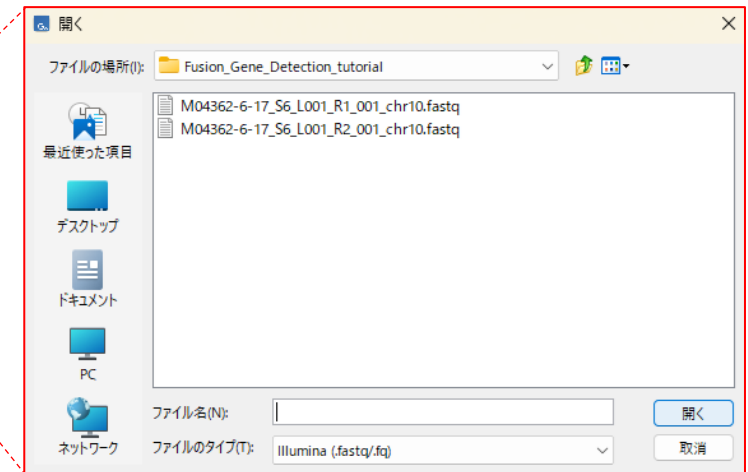
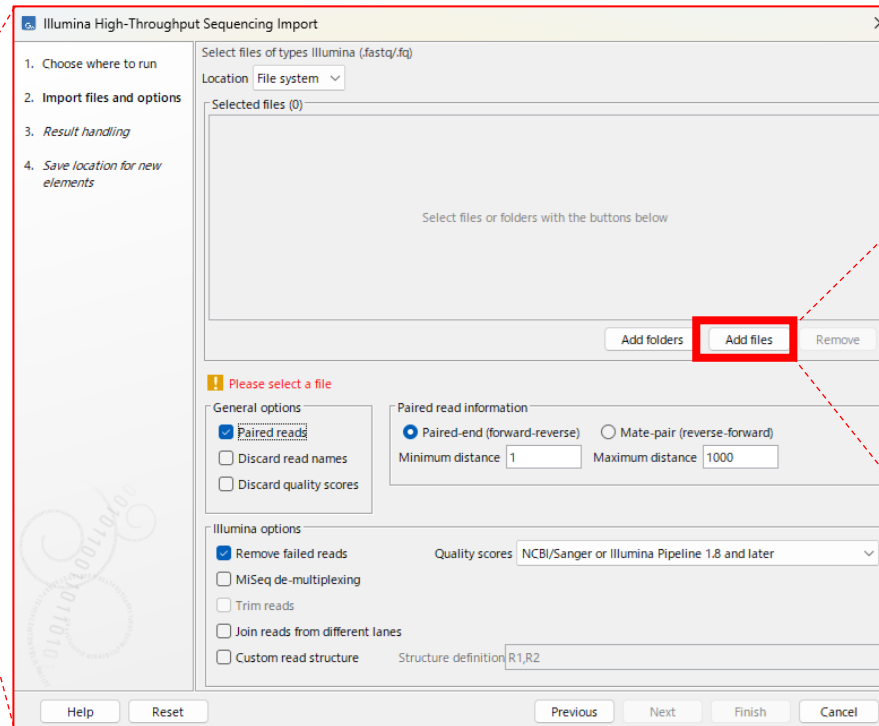
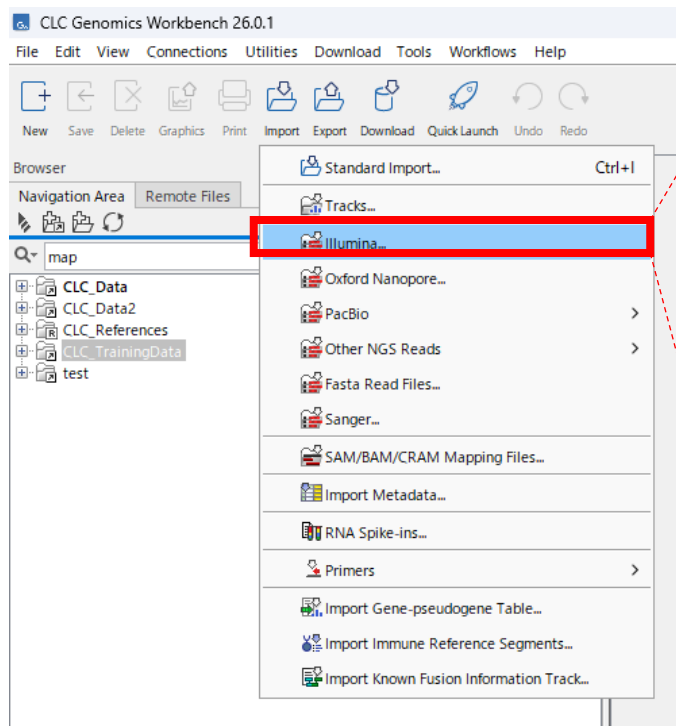
## 遺伝子情報の付加

Annotate with Nearby Information

- ✓ CLC Genomics Workbenchでは、シークエンサー機種や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



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



- ✓ 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

The screenshot displays the CLC Genomics Workbench 26.0.1 interface. The main window shows a list of sequence reads with their corresponding base sequences and quality scores. The reads are organized into pairs, with the top read of each pair highlighted in blue. The quality scores are represented by a bar chart below each sequence. The interface includes a navigation area on the left, a toolbox at the bottom left, and a settings panel on the right. The settings panel is currently open to the 'Sequence List Settings' tab, showing options for sequence layout, spacing, and alignment. The 'Sequence List Settings' panel includes options for 'Spacing' (No spacing), 'Double stranded' (unchecked), 'Numbers on sequences' (checked), 'Relative to' (set to 1), 'Numbers on plus strand' (checked), 'Lock numbers' (unchecked), 'Hide labels' (unchecked), 'Lock labels' (checked), 'Align labels' (set to Right), 'Sequence label' (Name), 'Show selection boxes' (unchecked), 'Annotation layout' (Annotation types), 'Show annotations' (checked), 'Position' (Next to sequence), 'Offset' (Little offset), 'Label' (Stacked), 'Show arrows' (checked), and 'Use gradients' (checked). The 'Restriction sites', 'Motifs', 'Residue coloring', 'Nucleotide info', 'Positional stats', 'Find', and 'Text format' options are all unchecked.

## シーケンスデータのインポート

Import

## クオリティチェック

QC for Sequencing Reads  
Trim Reads

## リファレンスへのマッピング

Map Reads to Reference

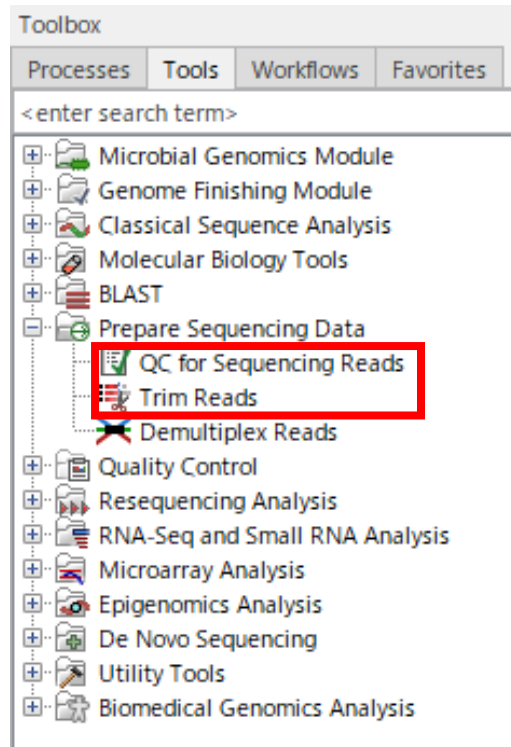
## ピークの検出

Transcription Factor ChIP-Seq  
Histone ChIP-Seq

## 遺伝子情報の付加

Annotate with Nearby Information

- ✓ 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



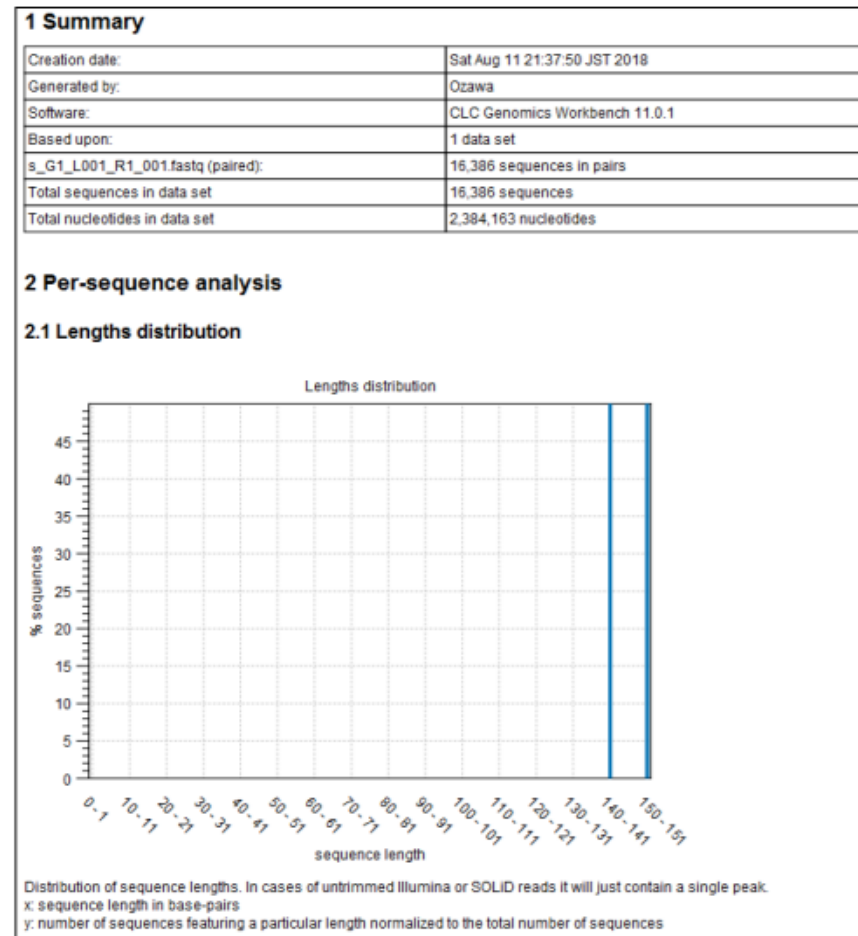
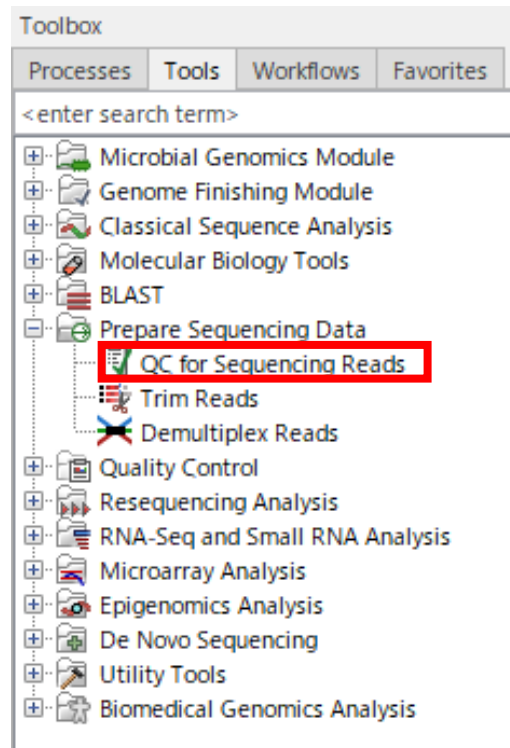
## QC for Sequencing Reads

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

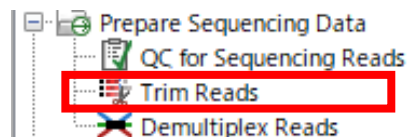
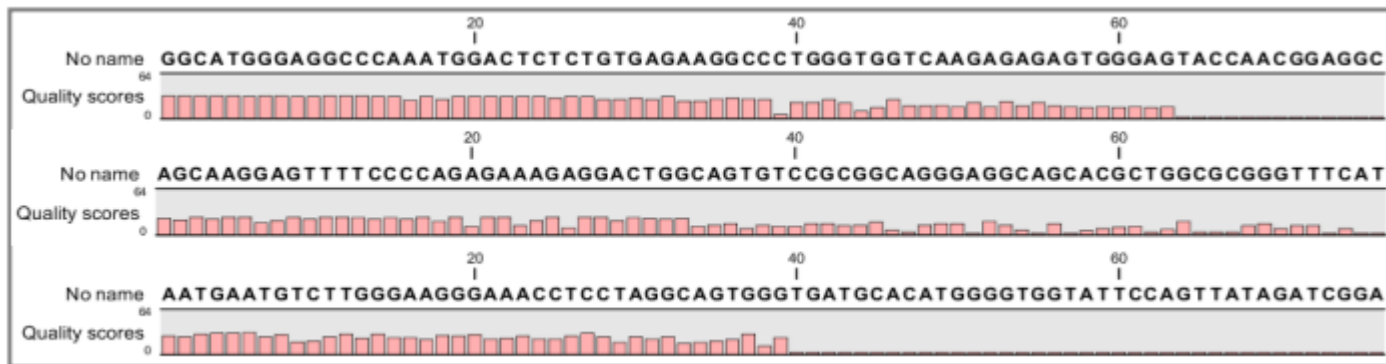
## Trim Reads

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

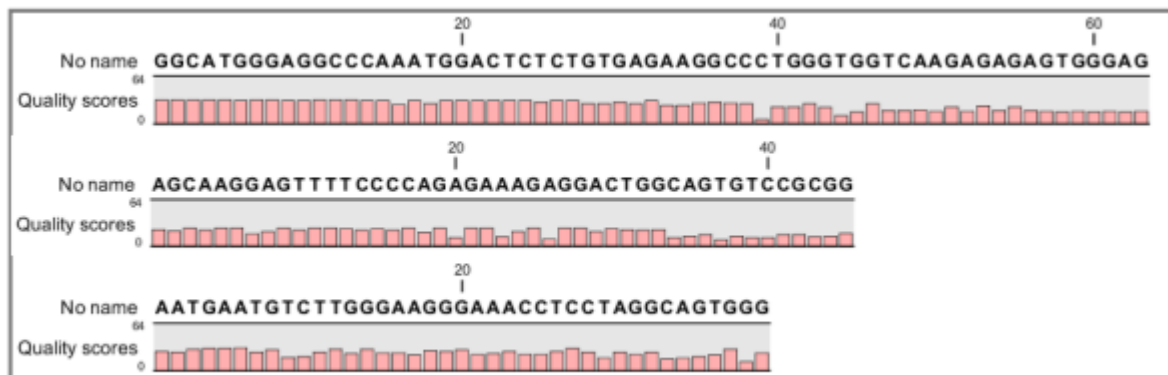
- ✓ QC for Sequencing Readsでは、シークエンスデータのクオリティ情報をまとめたレポートが作成される
- ✓ GC含量やクオリティスコア分布などのグラフデータや数値データを確認が可能



# リードのQCとトリミング



- ✓ 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
- ✓ アダプター配列の除去なども可能



## シーケンスデータのインポート

Import

## クオリティチェック

QC for Sequencing Reads  
Trim Reads

## リファレンスへのマッピング

Map Reads to Reference

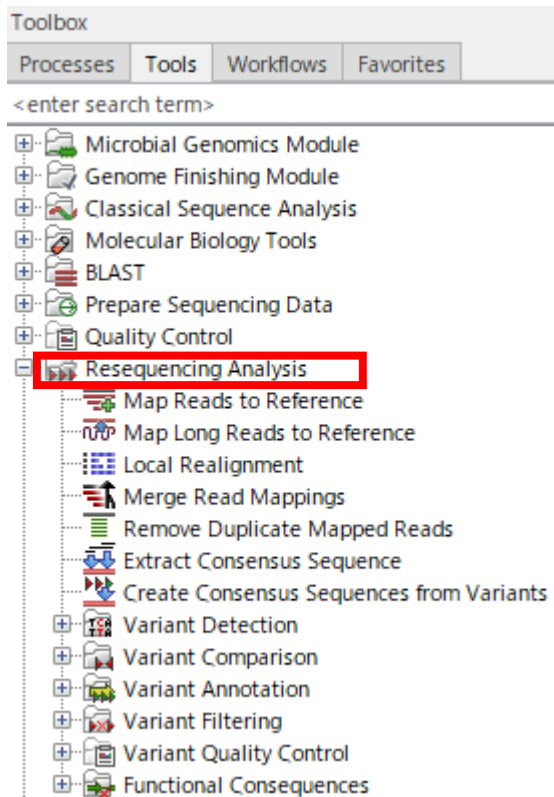
## ピークの検出

Transcription Factor ChIP-Seq  
Histone ChIP-Seq

## 遺伝子情報の付加

Annotate with Nearby Information

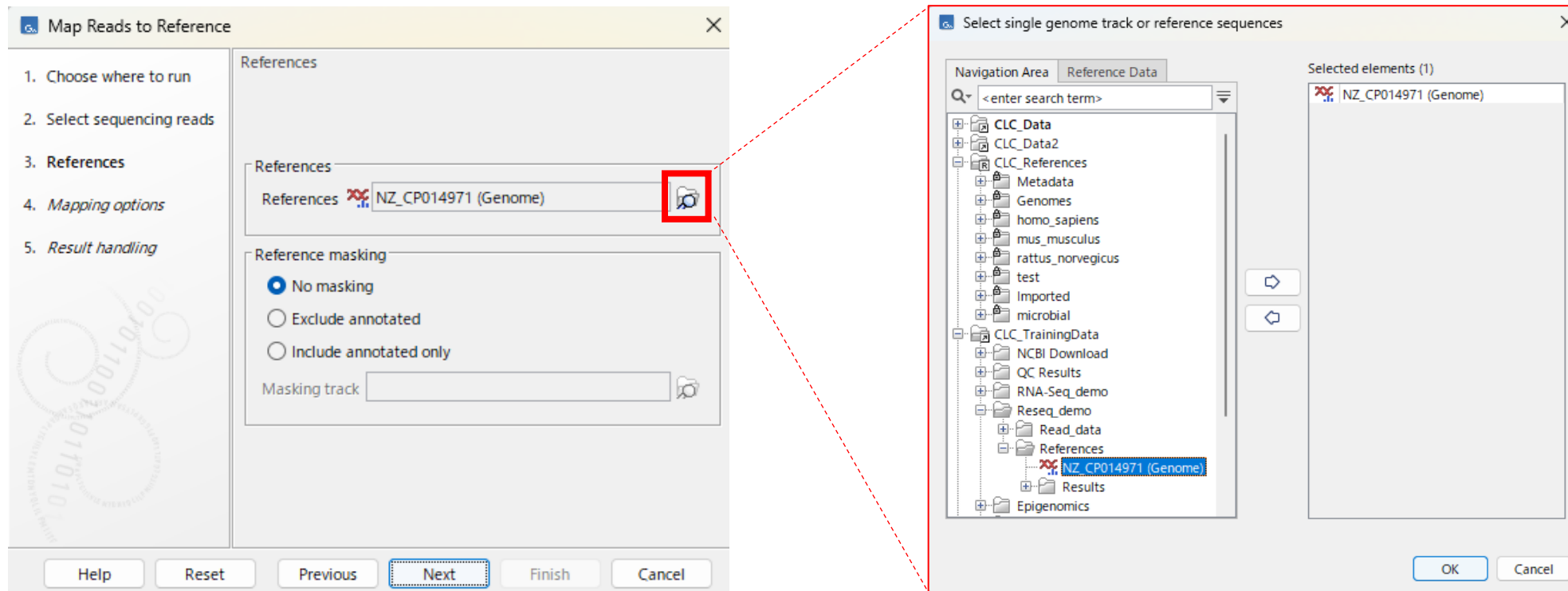
- ✓ ChIP-seqで得られた配列断片が、ゲノム上のどこに由来しているかを調べるために、リードデータのマッピングを行う
- ✓ コントロール (Input) のデータがある場合は、それについてもマッピングを行う



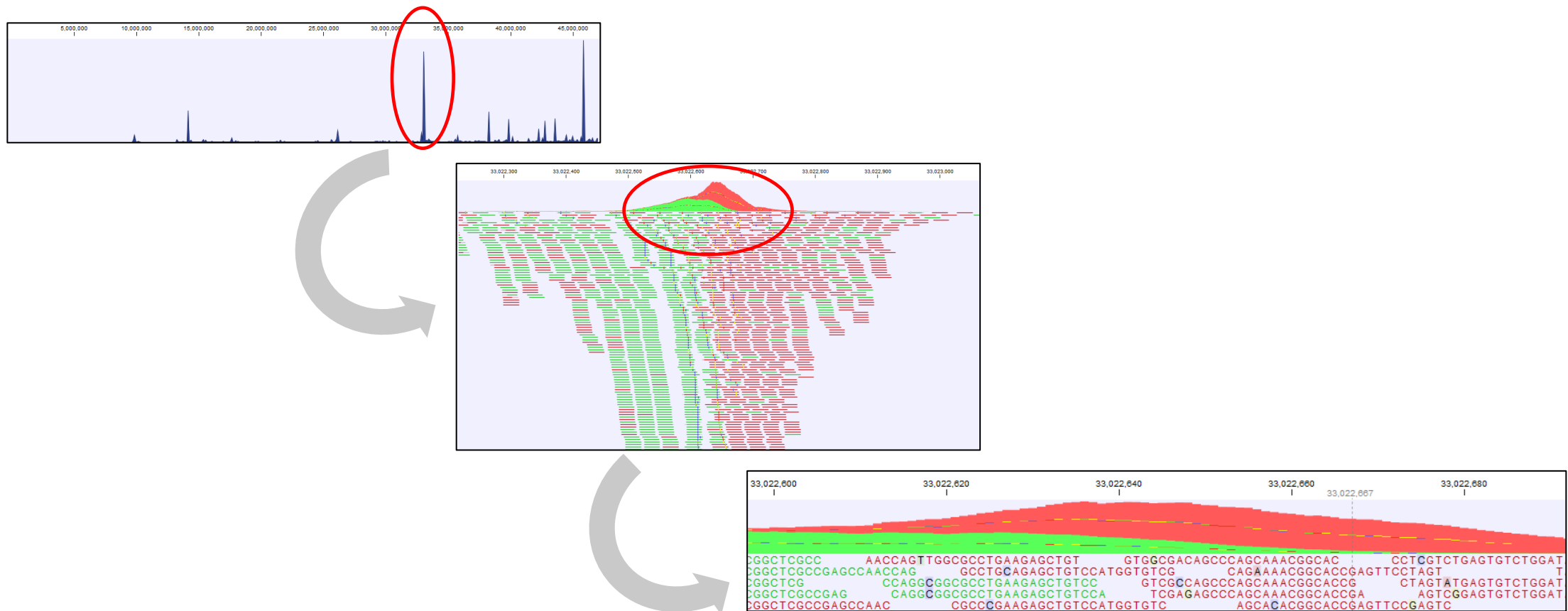
## Map Reads to Reference

任意の参照ゲノム配列に対して、シークエンスデータのマッピングを行う

- ✓ 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が確認できる



## シーケンスデータのインポート

Import

## クオリティチェック

QC for Sequencing Reads  
Trim Reads

## リファレンスへのマッピング

Map Reads to Reference

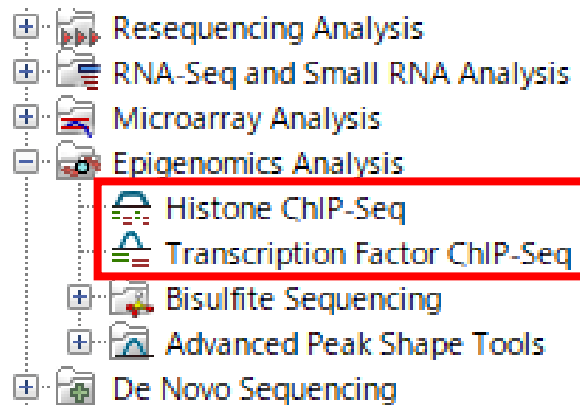
## ピークの検出

Transcription Factor ChIP-Seq  
Histone ChIP-Seq

## 遺伝子情報の付加

Annotate with Nearby Information

- ✓ マッピングデータを基にピークの検出を行い、転写因子結合部位もしくはヒストン修飾部位を予測する
- ✓ 転写因子結合部位を検出する場合とヒストン修飾部位を検出する場合で、それぞれ専用のツールが用意されている



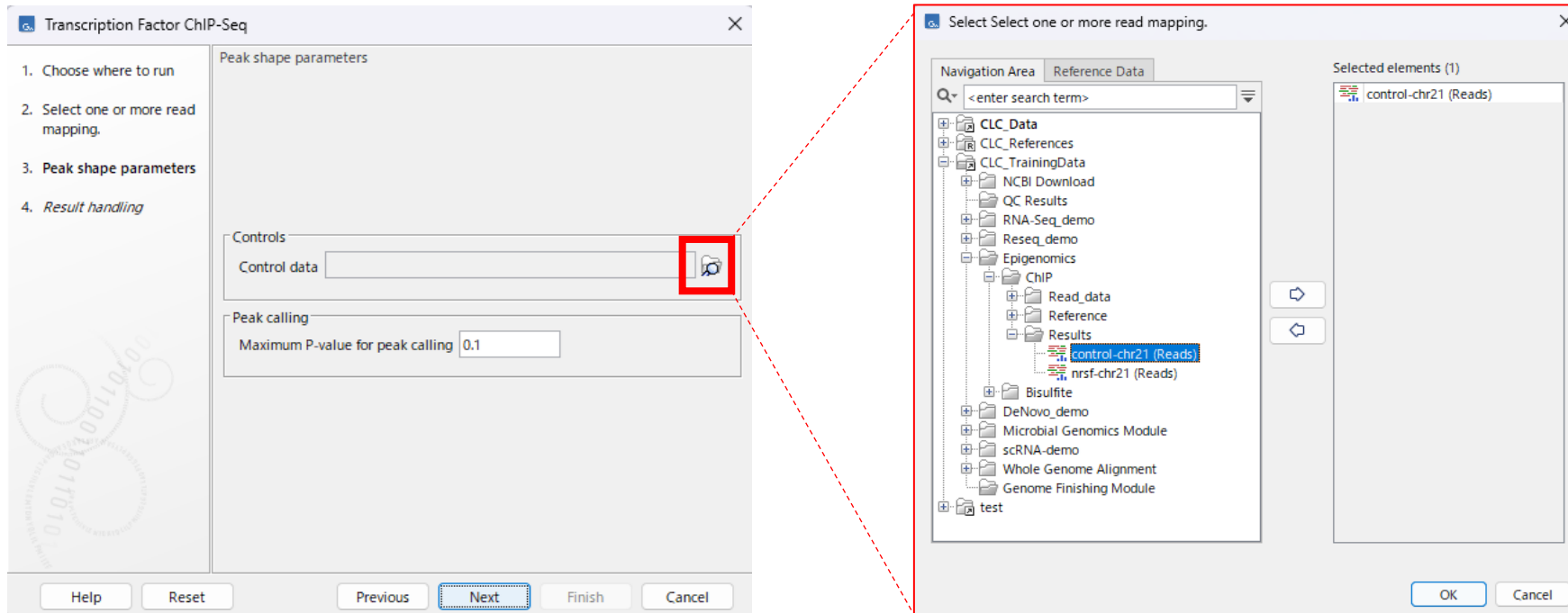
## Histone ChIP-Seq

ヒストン修飾をターゲットとするChIP-seqのピーク検出

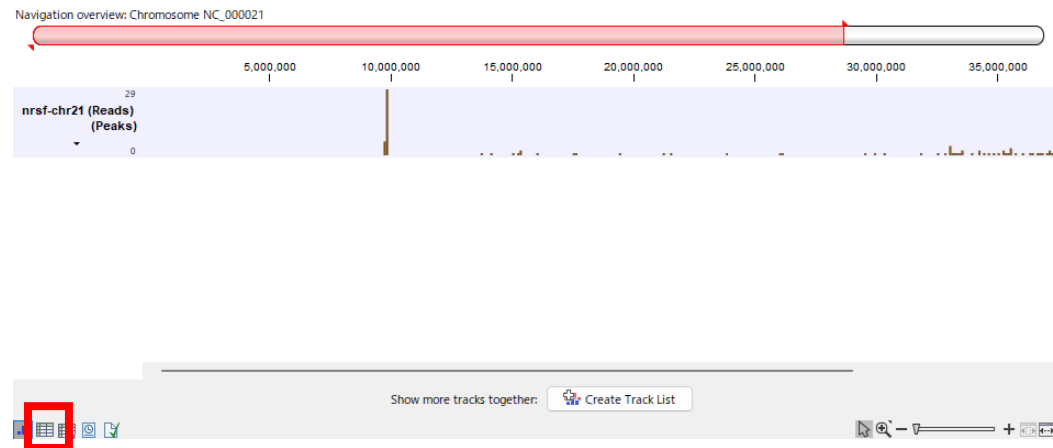
## Transcription Factor ChIP-Seq

転写因子をターゲットとするChIP-Seqのピークの検出

- ✓ コントロール (Input) のマッピングデータがある場合は、それに基づいて、ノイズやPCRによるアーティファクトと考えられるピークを除外できる



- ✓ ゲノム配列（染色体）ごとに、ピークの位置を確認できる
- ✓ テーブルビューに切り替えることで、ピーク位置の詳細等を確認できる



| Chromosome | Region           | Center of peak | Length | P-value |
|------------|------------------|----------------|--------|---------|
| NC_000021  | 9795720..9795925 | 9795829        | 206    | 0.05    |
| NC_000021  | 9796153..9796358 | 9796261        | 206    | 0.07    |
| NC_000021  | 9796465..9796666 | 9796563        | 202    | 0.06    |
| NC_000021  | 9797950..9798170 | 9798073        | 221    | 0.07    |
| NC_000021  | 9799025..9799264 | 9799123        | 240    | 0.04    |
| NC_000021  | 9799818..9800066 | 9799969        | 249    | 0.06    |
| NC_000021  | 9800343..9800544 | 9800441        | 202    | 0.03    |
| NC_000021  | 9801031..9801230 | 9801129        | 200    | 0.07    |
| NC_000021  | 9802071..9802289 | 9802169        | 219    | 0.07    |
| NC_000021  | 9805192..9805387 | 9805290        | 196    | 0.02    |
| NC_000021  | 9805754..9805951 | 9805854        | 198    | 4.69E-3 |
| NC_000021  | 9806925..9807186 | 9807023        | 262    | 0.03    |
| NC_000021  | 9807449..9807738 | 9807547        | 290    | 0.02    |
| NC_000021  | 9808132..9808379 | 9808282        | 248    | 0.05    |

## シーケンスデータのインポート

Import

## クオリティチェック

QC for Sequencing Reads  
Trim Reads

## リファレンスへのマッピング

Map Reads to Reference

## ピークの検出

Transcription Factor ChIP-Seq  
Histone ChIP-Seq

## 遺伝子情報の付加

Annotate with Nearby Information

- ✓ ピークの一覧表には、各ピークの5'側および3'側に存在する遺伝子の情報を付加できる
- ✓ ピーク近傍の遺伝子情報を付加した表により、どの遺伝子が制御されているかを予測可能

## 検出されたピーク

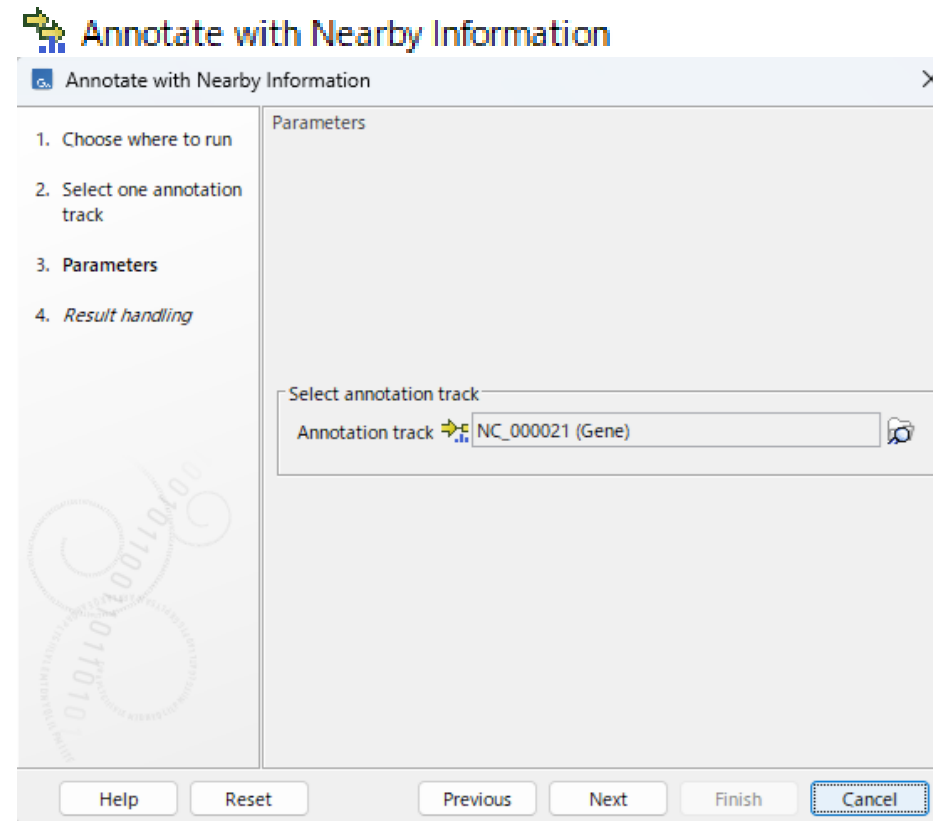
## ピークの上流／下流にある遺伝子

| Chromosome | Region             | Center of peak | Length | P-value  | 5' annotation | 5' distance | 3' annotation | 3' distance |
|------------|--------------------|----------------|--------|----------|---------------|-------------|---------------|-------------|
| NC_000021  | 9862753..9862953   | 9862856        | 201    | 0.04     |               |             | TPTE          | 65658       |
| NC_000021  | 13730198..13730397 | 13730300       | 200    | 0.06     | FGF7P2        | 82445       | LOC642460     | 0           |
| NC_000021  | 14141010..14141215 | 14141118       | 206    | 3.98E-25 | LOC391266     | 19190       | CYP4F3LP      | 0           |
| NC_000021  | 15028559..15028755 | 15028658       | 197    | 0.06     | LOC388813     | 91242       | POLR2CP       | 22924       |
| NC_000021  | 15283069..15283277 | 15283167       | 209    | 0.01     | NRIP1         | 20350       | CYCSP42       | 501436      |
| NC_000021  | 15358434..15358634 | 15358532       | 201    | 0.01     | NRIP1         | 95715       | CYCSP42       | 426079      |
| NC_000021  | 15359477..15359675 | 15359578       | 199    | 1.35E-36 | NRIP1         | 96758       | CYCSP42       | 425038      |
| NC_000021  | 16012930..16013199 | 16013102       | 270    | 0.03     | CYCSP42       | 227957      | C21orf34      | 351513      |
| NC_000021  | 17573952..17574151 | 17574050       | 200    | 1.08E-3  | VDAC2P        | 1184288     | CXADR         | 233049      |
| NC_000021  | 17637362..17637562 | 17637460       | 201    | 5.53E-61 | VDAC2P        | 1247698     | CXADR         | 169638      |
| NC_000021  | 19403978..19404179 | 19404076       | 202    | 5.53E-3  | PPIAL3        | 251512      | SLC6A6P       | 134403      |
| NC_000021  | 21292227..21292439 | 21292325       | 213    | 0.05     | RPS3AP1       | 567183      | NCAM2         | 64          |
| NC_000021  | 21524960..21525160 | 21525063       | 201    | 1.61E-22 | NCAM2         | 0           | MAPK6PS2      | 1203964     |
| NC_000021  | 23817057..23817253 | 23817155       | 197    | 0.01     | TUBAP         | 113543      | VN2R20P       | 387059      |
| NC_000021  | 26095514..26095741 | 26095612       | 228    | 0.01     | GABPA         | 29300       | APP           | 78990       |
| NC_000021  | 26149904..26150103 | 26150002       | 200    | 0.02     | GABPA         | 83690       | APP           | 24628       |
| NC_000021  | 29594185..29594429 | 29594283       | 245    | 0.05     | BACH1         | 0           | C21orf41      | 26891       |

使用ツール

 [Annotate with Nearby Information](#)

- ✓ 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



✓ マッピング結果や、ピークの位置の遺伝子の位置関係を目視で確認できる

Utility Tools

- Mask Low-Complexity Regions
- Extraction
- Filtering
- Renaming
- Reports
- Result Metadata
- Sequences
- Tracks
  - Create Track List**
  - Convert Tracks
  - Graph Tracks
  - Merge Tracks
  - Modify Tracks

Create Track List

1. Select tracks from same genome

Select tracks from same genome

Navigation Area Reference Data

<enter search term>

Epigenomics

- ChIP
- Read\_data
- Reference
  - NC\_000021 (Gene)
  - NC\_000021 (Genome)
- Results
  - control-chr21 (Reads)
  - nrnf-chr21 (Reads)
  - nrnf-chr21 (Reads) (Peaks)
  - nrnf-chr21 (Reads) (Peaks, Ann)
- Bisulfite
- DeNovo\_demo
- Microbial Genomics Module
- scRNA-demo

Selected elements (4)

- NC\_000021 (Genome)
- NC\_000021 (Gene)
- nrnf-chr21 (Reads) (Peaks)
- nrnf-chr21 (Reads)

Help Reset Previous Next Finish Cancel

32,900,000 32,950,000 33,000,000 33,050,000

NC\_000021 (Genome)

NC\_000021 (Gene)  
Gene annotations (352)

nrnf-chr21 (Reads) (Peaks)

nrnf-chr21 (Reads)  
486,301 reads

拡大

33,021,500 33,022,000 33,022,500 33,023,000

NC\_000021 (Genome)

NC\_000021 (Gene)  
Gene annotations (352)

nrnf-chr21 (Reads) (Peaks)  
Peak annotations (146)

nrnf-chr21 (Reads)  
486,301 reads

- ✓ 各ピーク領域の塩基配列を抽出することができる
- ✓ 抽出した配列は、配列リストとして出力し、外部ツールでのモチーフ検索等に使用可能

## Extract Annotated Regions

|      |  |     |     |     |     |
|------|--|-----|-----|-----|-----|
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | AGTAGAGTGTATTGGAGTGGCGGCGAATGGAATGGAATGGTAAGGAATGGAATTGAAAGGAGTGCAGTGGAGCGGAGTGGAGTGGAAATGTAATGGAATGGATGGAAT           |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | AAGTGGAAATGGAGTGGAAATAGAATAGGGTGGAAATGGAATTGAATGTAATGGAGAGGAGTGGAGTGGAAATGGAGTGGAAATGAAATGGAATGGGAAGGAATGGTATTGA       |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | GTGGAATGGAATAGAATGCAATGGAATGGAAAGGTATGGACTCTAATGGAATGAAGTGGGATGGAGTGGACTCGAATAGAATGGACTGGAATGGAACAGACATGAAT            |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | GGAATGGAGTCGAATAGAATGTAATAGAATGGAATTGAATCAAATCAAATGGAATCGAATGGAATTCGGTGAAGTGAAGTGCAGTGGAGTGGAAAGGATGGGAATGC            |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | GGAATGGAACGGAATAGAGTGGAAATAGAACAGAATGGATTGGAATGGAATGGACACGAATGGAATGGAATGGAATGCAATGGAATGCAATGGAATGCAATGGAAAGG           |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | AAGGAATGGTATTGAAAGGAGTGGAGTGGAGTGGAAATGGAGTAGAAAGGAATGCAATGGAATGGAATGGAGTGGAGCGGAGTGGAGTGGAAATGGAGATGAGTGGAGT          |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | TGCAGTGGAAATGGAACAGAAAGGAATGGAATGGAGTGGAGTGAAGTGGAAACAAAATGGAATGGAATGGAATGGAAAGCAGGGACATGTAATGGATTGGATTGGAATG          |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | TGGAATGGAACGGAATGGAATGGAAGCGAATGGAATGGTATGGAATGGAGTGTAGTGGAGTGGAAATGGAGTGGAAACGGAATGGAGTCATGTGGAGTGGAAATGGAATG         |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | GGAATGGAATGGAGTGTAGTGGACTGGAGTTGAGTGGAGTGGAGTAGAGTTGAATGGAAAGAAGTGGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAGTGGAGTGGATAAA |     |     |     |     |
| 60   | 80   | 100 | 120 | 140 | 160 |
| Peak | AATGGAAAGGAAGGAATGGAGTGGAGTGTAGTGGACTGGAGTGGAGTGGAGTGGAGTTGAATGGAGTTGAGTGTAGTGGAGTGGAAATGGAGTGGTTTGGAGTGGAAATGGA       |     |     |     |     |

お問い合わせ先：フィルジエン株式会社

TEL: 052-624-4388 (9:00～17:00)

FAX: 052-624-4389

E-mail: [support@filgen.jp](mailto:support@filgen.jp)