

# microRNAの発現解析

発現定量・アノテーション付与・エンリッチメント解析

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



## 次世代シーケンス解析用ソフトウェア

### ◆ リシーケンシング解析

- ・ リファレンスゲノムへのマッピング
- ・ 変異検出

### ◆ トランスクリプトミクス解析

- ・ RNA-seq解析
- ・ small RNA解析

### ◆ エピゲノミクス解析

- ・ ChIP-seq解析
- ・ バイサルファイトシーケンス解析

### ◆ De Novo シーケンス解析

- ・ De Novo Assembly
- ・ BLAST解析

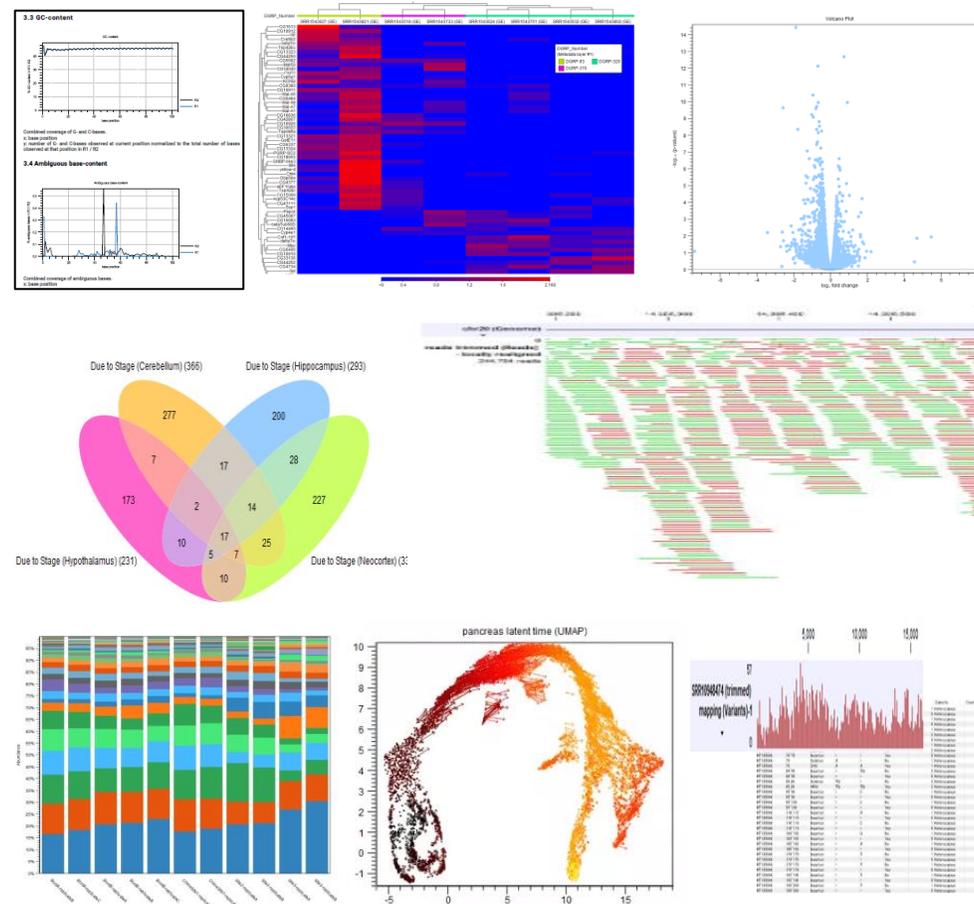
### ◆ 菌叢解析

### ◆ ゲノムフィニッシング解析

### ◆ シングルセル解析

### ◆ 超高速変異解析

} Premium版 限定機能





## 次世代シーケンス解析用ソフトウェア

### ◆ リシーケンシング解析

- ・ リファレンスゲノムへのマッピング
- ・ 変異検出

### ◆ トランスクリプトミクス解析

- ・ RNA-seq解析
- ・ **small RNA解析**

### ◆ エピゲノミクス解析

- ・ ChIP-seq解析
- ・ バイサルファイトシーケンス解析

### ◆ De Novo シーケンス解析

- ・ De Novo Assembly
- ・ BLAST解析

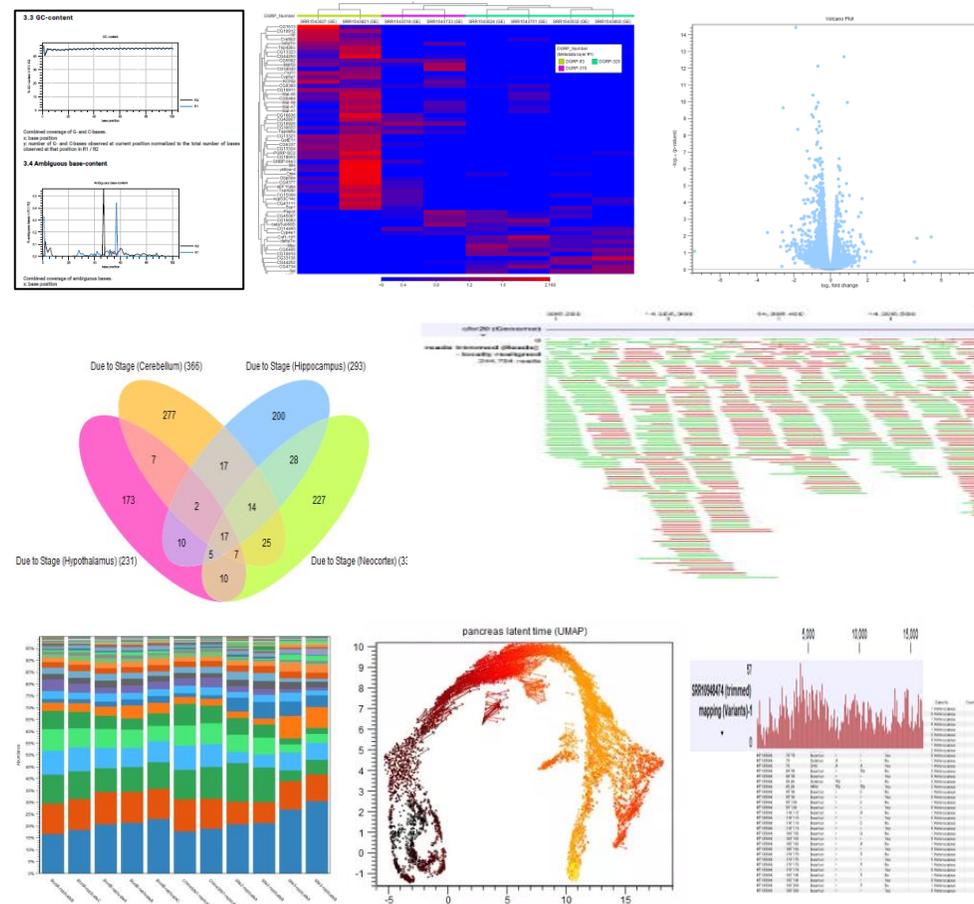
### ◆ 菌叢解析

### ◆ ゲノムフィニッシング解析

### ◆ シングルセル解析

### ◆ 超高速変異解析

} Premium版 限定機能



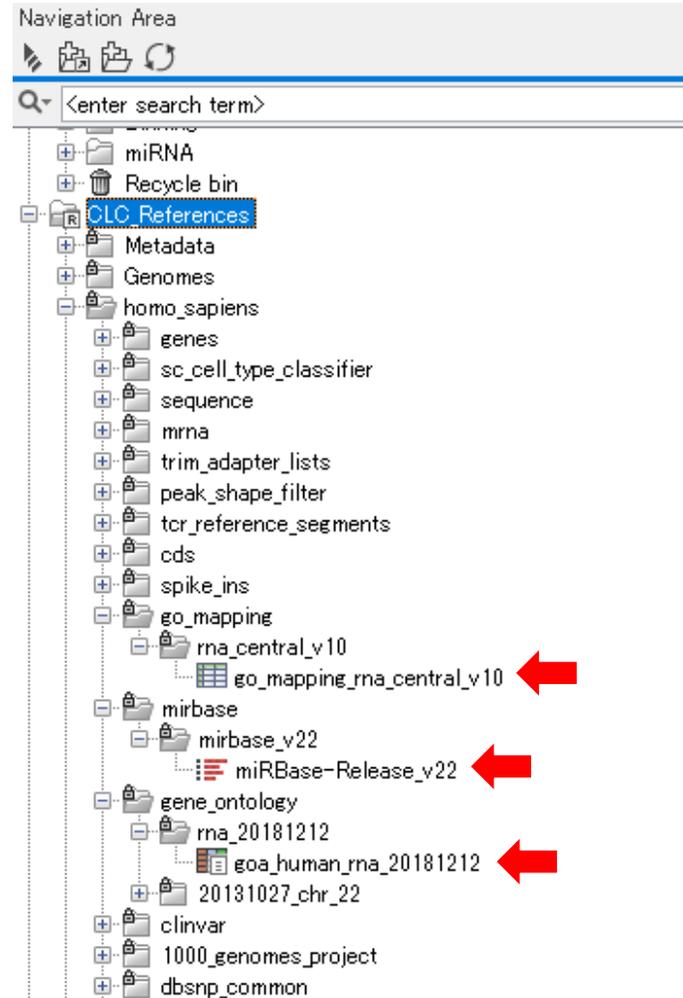
1. microRNAの発現定量、エンリッチメント解析
2. isomiRベースの解析
3. 新規microRNAの検出

## データベース、アノテーションファイルのダウンロード

The screenshot shows the CLC Genomics Workbench 23.04 interface. The 'Manage Reference Data' window is open, displaying a list of reference data sets. The 'QIAGEN Sets' tab is selected, and the 'QIaseq Small RNA' set is highlighted. The details for this set are shown on the right, including a table of reference data included.

Workflow role	Version	Download Size	On Disk Size
<input checked="" type="checkbox"/> mirbase	mirbase_v22	6.1 MB	7.9 MB
<input checked="" type="checkbox"/> spike_ins	qiaseq_mirna_v1	2 KB	3 KB
<input checked="" type="checkbox"/> trim_adapter_lists	qiaseq_mirna_v1	1008 B	2 KB
<input checked="" type="checkbox"/> gene_ontology	ma_20181212	528 KB	704 KB
<input checked="" type="checkbox"/> ecmapping	rna_central_v10	889 KB	1.1 MB

Reference > QIAGEN Sets > QIaseq Small RNAからダウンロード可能です。



ダウンロードしたデータは、CLC Referencesというロケーションに保存されます。

**1. microRNAの発現定量、エンリッチメント解析**

2. isomiRベースの解析

3. 新規microRNAの検出

NGSから出力されたデータのインポート

アダプターのトリミング、UMI情報付与 (オプション)

 Trim Reads  
 Create UMI Reads for miRNA

miRNAの定量

 Quantify miRNA

- ・ヒートマップの作成
- ・PCAプロットの作成

RNACentral IDによるアノテーション付け

 Annotate with RNACentral Accession Numbers

発現量の比較解析

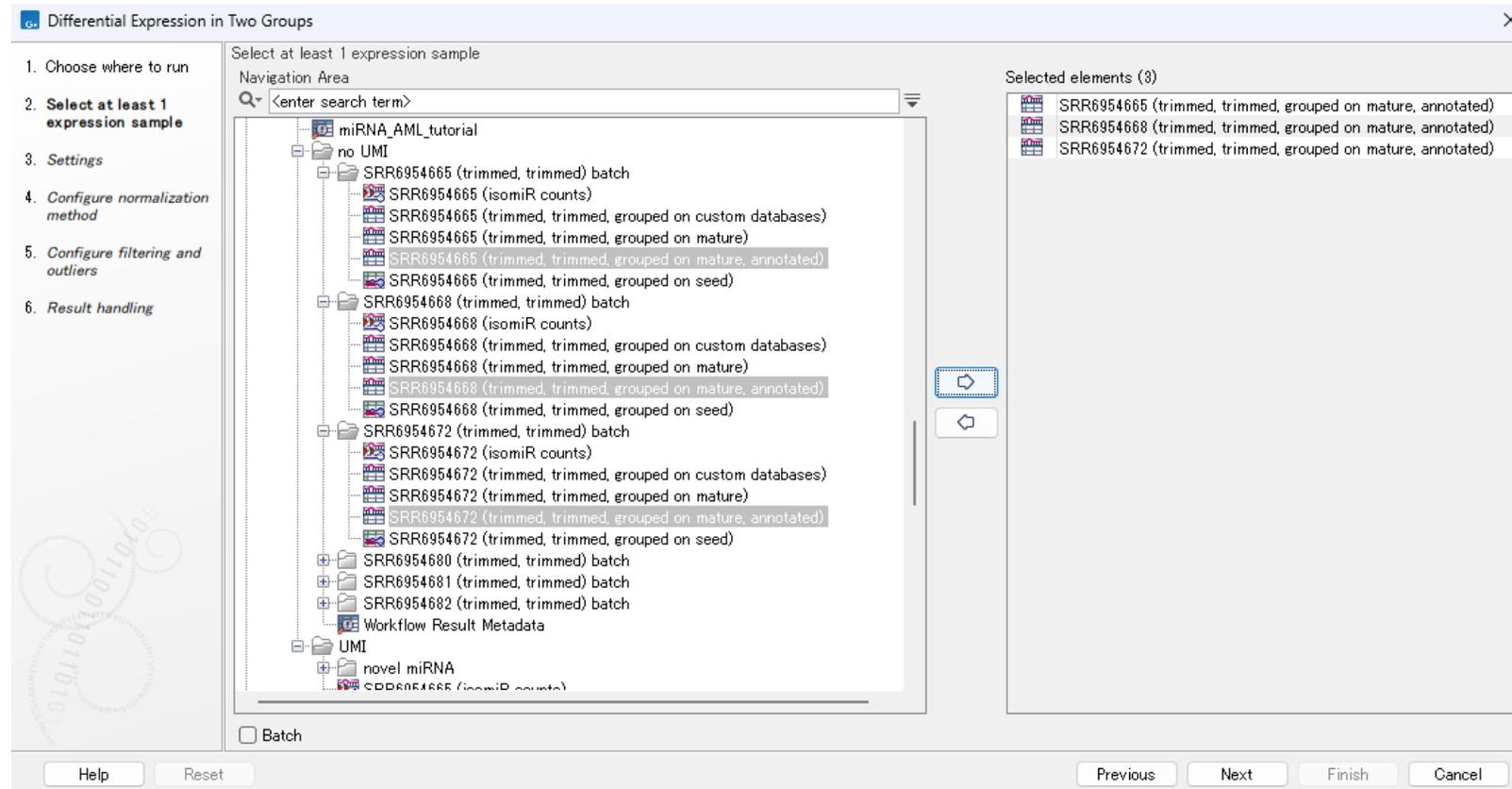
 Differential Expression in Two Groups  
 Differential Expression for RNA-Seq

- ・ベン図の作成
- ・発現量テーブルの作成

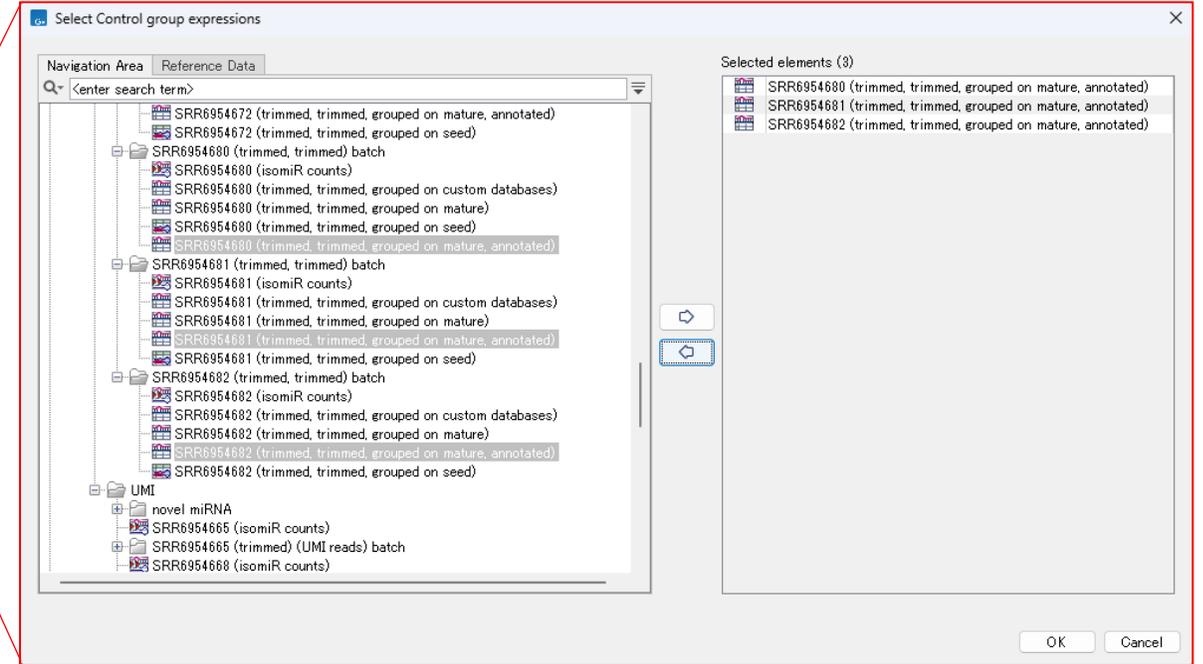
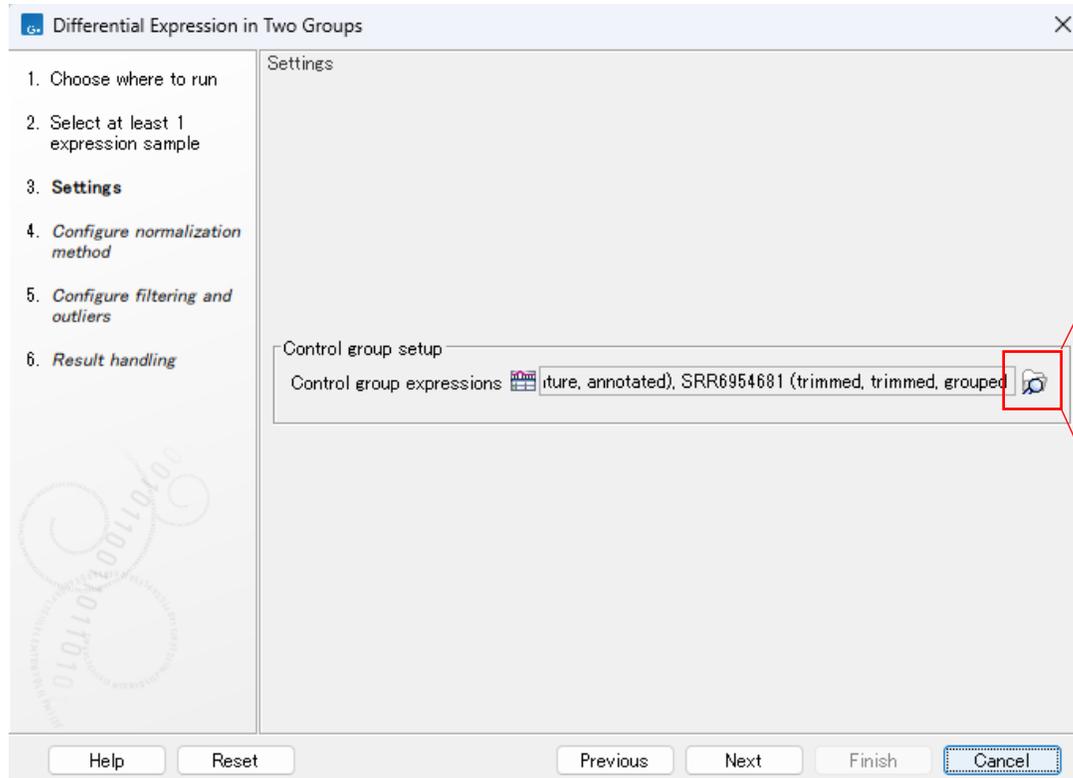
エンリッチメント解析

 Gene Set Test

-  Differential Expression in Two Groups **2群間の比較**
-  Differential Expression for RNA-Seq **多群間の比較** (各サンプルとグループを対応づけるメタデータを用意する必要がある)



1群目に含まれるデータをすべて指定 (今回はn=3)



次の画面で、2群目に含まれるデータをすべて指定（今回はn=3）

1. microRNAの発現定量、エンリッチメント解析

**2. isomiRベースの解析**

3. 新規microRNAの検出

NGSから出力されたデータのインポート

アダプターのトリミング、UMI情報付与 (オプション)

 Trim Reads  
 Create UMI Reads for miRNA

miRNAの定量

 Quantify miRNA

- ・ヒートマップの作成
- ・PCAプロットの作成

IsomiRの定量

 Extract IsomiR Counts

isomiRカウントの比較解析

 Differential Expression in Two Groups  
 Differential Expression for RNA-Seq

- ・ベン図の作成
- ・発現量テーブルの作成

## Extract IsomiR Counts

miRNA samples

Navigation Area

Q > <Enter search term>

novel miRNA

- SRR6954665 (trimmed) (UMI reads) batch
  - SRR6954665 (trimmed) (UMI reads, grouped on custom databases)
  - SRR6954665 (trimmed) (UMI reads, grouped on mature)
  - SRR6954665 (trimmed) (UMI reads, grouped on mature, annotated)
- SRR6954668 (trimmed) (UMI reads) batch
  - SRR6954668 (trimmed) (UMI reads, grouped on custom databases)
  - SRR6954668 (trimmed) (UMI reads, grouped on mature)
  - SRR6954668 (trimmed) (UMI reads, grouped on mature, annotated)
- SRR6954672 (trimmed) (UMI reads) batch
  - SRR6954672 (trimmed) (UMI reads, grouped on custom databases)
  - SRR6954672 (trimmed) (UMI reads, grouped on mature)
  - SRR6954672 (trimmed) (UMI reads, grouped on mature, annotated)
- SRR6954680 (trimmed) (UMI reads) batch
  - SRR6954680 (trimmed) (UMI reads, grouped on custom databases)
  - SRR6954680 (trimmed) (UMI reads, grouped on mature)
  - SRR6954680 (trimmed) (UMI reads, grouped on mature, annotated)
- SRR6954681 (trimmed) (UMI reads) batch
  - SRR6954681 (trimmed) (UMI reads, grouped on custom databases)
  - SRR6954681 (trimmed) (UMI reads, grouped on mature)
  - SRR6954681 (trimmed) (UMI reads, grouped on mature, annotated)
- SRR6954682 (trimmed) (UMI reads) batch
  - SRR6954682 (trimmed) (UMI reads, grouped on custom databases)
  - SRR6954682 (trimmed) (UMI reads, grouped on mature)
  - SRR6954682 (trimmed) (UMI reads, grouped on mature, annotated)

Workflow Result Metadata

OLC\_References

Metadata

Selected elements (6)

- SRR6954665 (trimmed) (UMI reads, grouped on mature)
- SRR6954668 (trimmed) (UMI reads, grouped on mature)
- SRR6954672 (trimmed) (UMI reads, grouped on mature)
- SRR6954680 (trimmed) (UMI reads, grouped on mature)
- SRR6954681 (trimmed) (UMI reads, grouped on mature)
- SRR6954682 (trimmed) (UMI reads, grouped on mature)

Batch

Help Reset Previous Next Finish Cancel

実験群に関わらず、全サンプルを指定する  
(※Batchにはチェックをいれない)

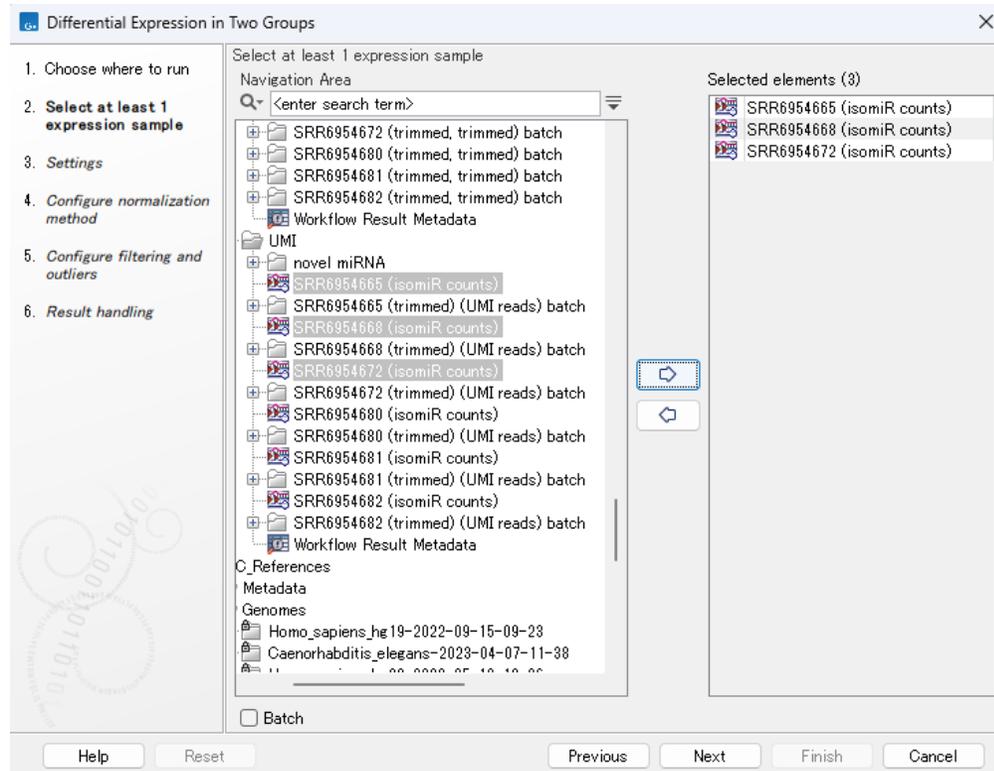


SRR6954668 (isomiR counts)

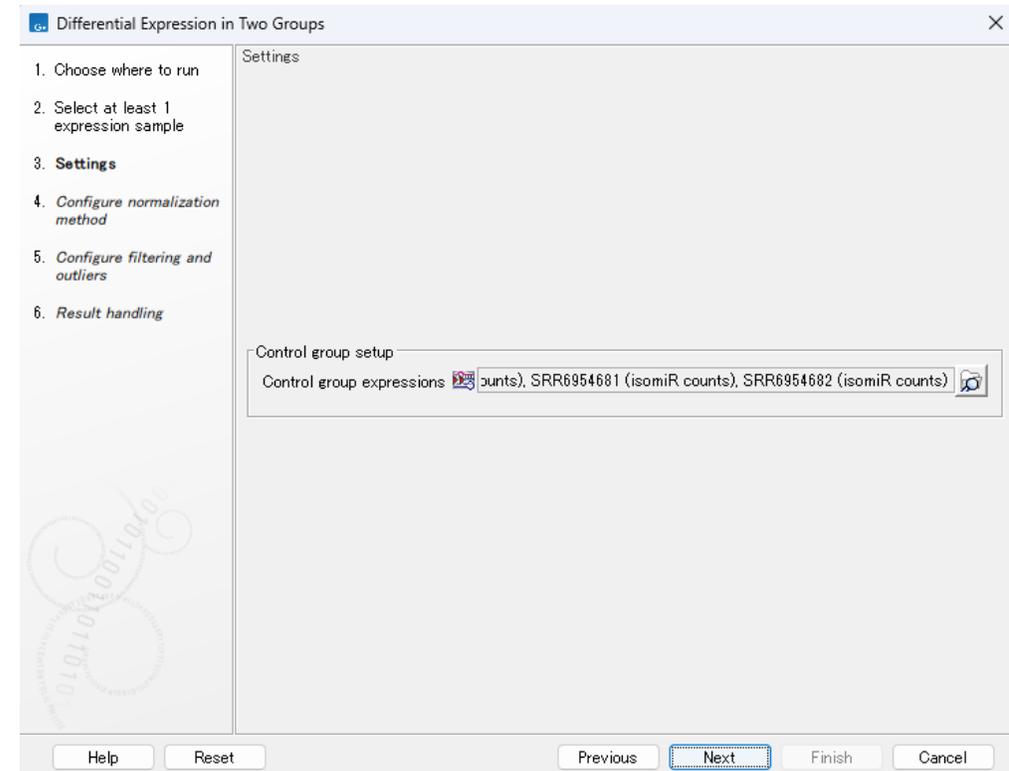
Rows: 30,426 Filter to Selection\*\*\*

Name	Sequence	Count	Ambiguous
hsa-miR-4484.AAs.a	AAAAAAGGGGGGAGAAGCCCC		1 No
hsa-miR-548h-3p.cs.T18G.ca	AAAAACCGCAATGACTTTTG		0 No
hsa-miR-548h-3p.cs.ca	AAAAACCGCAATTACTTTTG		0 No
hsa-miR-548h-3p.cs.A14G.ca	AAAAACCGCAATTGCTTTTG		0 No
hsa-miR-548h-3p.cs.A10C.ca	AAAAACCGCCATTACTTTTG		0 No
hsa-miR-548ac.cs.g	AAAAACCGGCAATTACTTTT		0 No
hsa-miR-451a.AAs	AAAAACCGTTACCATTACTGAGTT		0 No
hsa-miR-451a.AAs.T	AAAAACCGTTACCATTACTGAGTTT		0 No
hsa-miR-548e-3p.ca	AAAAACTGAGACTACTTTTG		0 No
hsa-miR-548e-3p	AAAAACTGAGACTACTTTTGCA		0 No
hsa-miR-548aq-3p.cs.A10C.c	AAAAACTGCCATTACTTTTG		0 No
hsa-miR-548e-3p.A9G.a	AAAAACTGGGACTACTTTTG		0 No
hsa-miR-320b.G5A.a	AAAAACTGGGTTGAGAGGGCA		0 No
hsa-miR-320a-3p.G5A.AA	AAAAACTGGGTTGAGAGGGCGAAA		0 No
hsa-miR-320a-3p.G5A.A22T	AAAAACTGGGTTGAGAGGGCGGT		0 No
hsa-miR-548bc.c	AAAAACTGTGATTACTTTTG		0 No
hsa-miR-548bc.A	AAAAACTGTGATTACTTTTGCA		0 No
hsa-miR-4484.As.A20G	AAAAAGGCGGGAGAAGCCCCG		0 No
hsa-miR-4484.As.A20G.G	AAAAAGGCGGGAGAAGCCCCGG		1 No
hsa-miR-12186.gs.cc	AAAAAGTCATGGAGG		0 No
hsa-miR-12186.gs	AAAAAGTCATGGAGGCC		0 No
hsa-miR-12186.gs.AT	AAAAAGTCATGGAGGCCAT		0 No
hsa-miR-451a.As.T7C.TA	AAAAACCGCTACCATTACTGAGTTTA		0 No
hsa-miR-1537-3p.t	AAAACCGTCTAGTTACAGTTG		0 No
hsa-miR-1537-3p	AAAACCGTCTAGTTACAGTTGT		0 No
hsa-miR-1537-3p.C16T	AAAACCGTCTAGTTATAGTTGT		0 No
hsa-miR-451a.As.T17C.tt	AAAACCGTTACCATTACCGAG		0 No
hsa-miR-451a.As.tt	AAAACCGTTACCATTACTGAG		0 No
hsa-miR-451a.As.t	AAAACCGTTACCATTACTGAGT		0 No
hsa-miR-451a.As	AAAACCGTTACCATTACTGAGTT		0 No
hsa-miR-451a.As.T	AAAACCGTTACCATTACTGAGTTT		0 No
hsa-miR-451a.As.C11T	AAAACCGTTACTATTACTGAGTT		0 No
hsa-miR-451a.As.C11T.T	AAAACCGTTACTATTACTGAGTTT		0 No

- Differential Expression in Two Groups **2群間の比較**
- Differential Expression for RNA-Seq **多群間の比較** (各サンプルとグループを対応づけるメタデータを用意する必要がある)



1群目のisomiRカウントデータをすべて指定  
(今回はn=3)



2群目のisomiRカウントデータをすべて指定  
(今回はn=3)

以降のパラメータ設定は、mature miRNAの発現量比較と同様

 SRR6954668 (isomiR counts)



 PCA for RNA-Seq

 Create Heat Map for RNA-Seq

 Create Expression Browser

 Create Venn Diagram for RNA-Seq

mature miRNAの発現量データの代わりに、isomiRのカウントデータを使用することで、セクション1のようなプロットの作成が可能

1. microRNAの発現定量、エンリッチメント解析

2. isomiRベースの解析

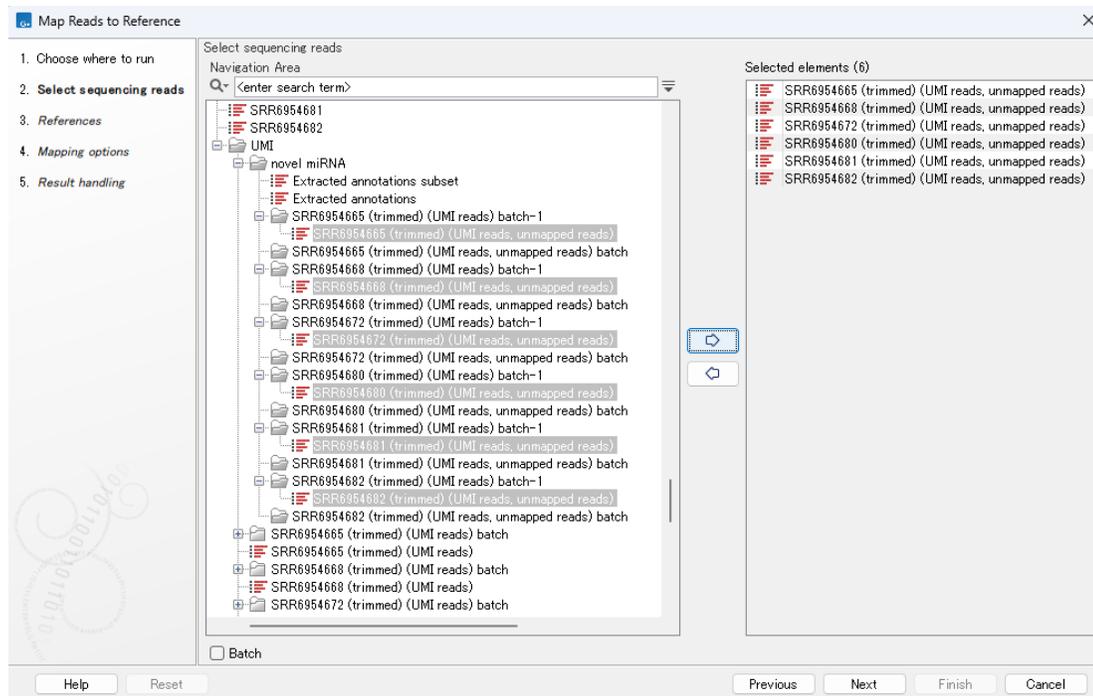
**3. 新規microRNAの検出**



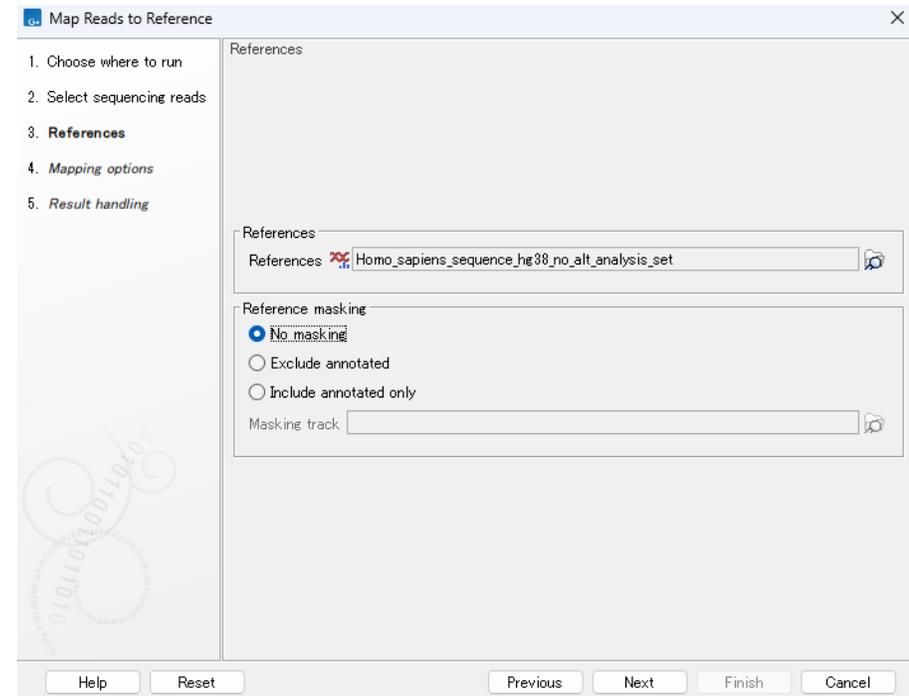
# unmapped readの再マッピング

データベース上の配列にマッピングされなかったリードを、ゲノム配列上にマッピングする。

## Map Reads to Reference



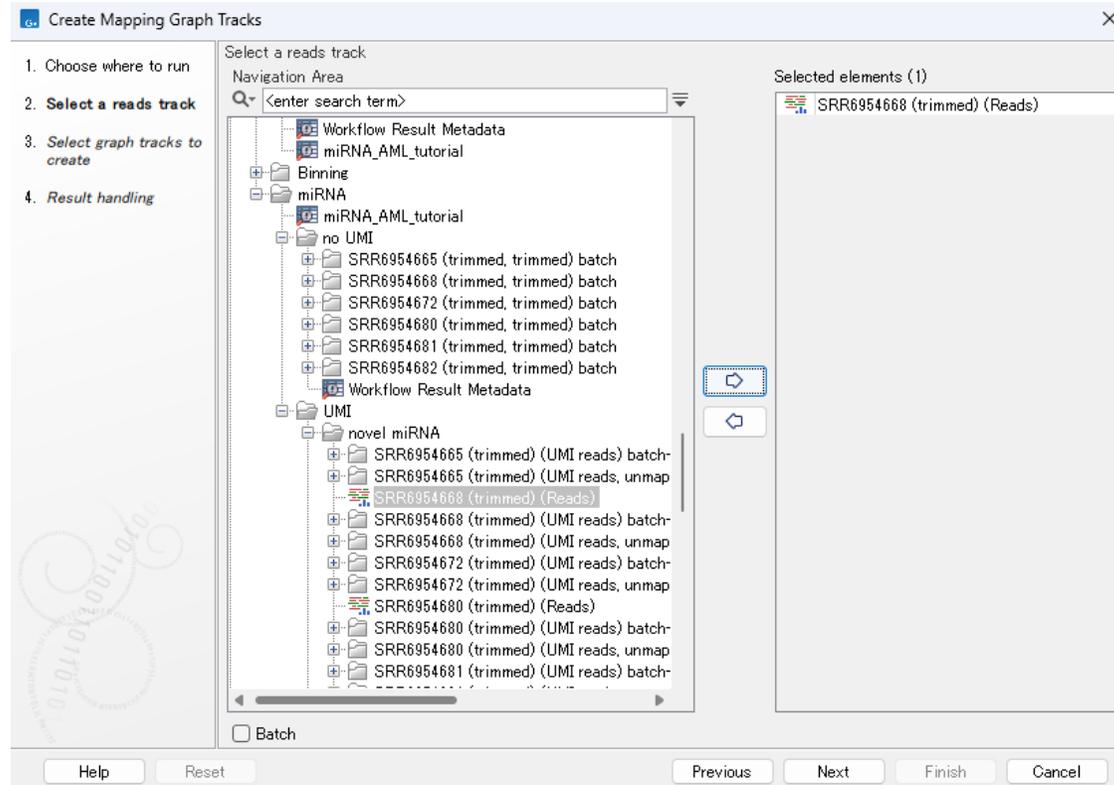
Quantify miRNAツールで出力されたunmapped readデータを指定



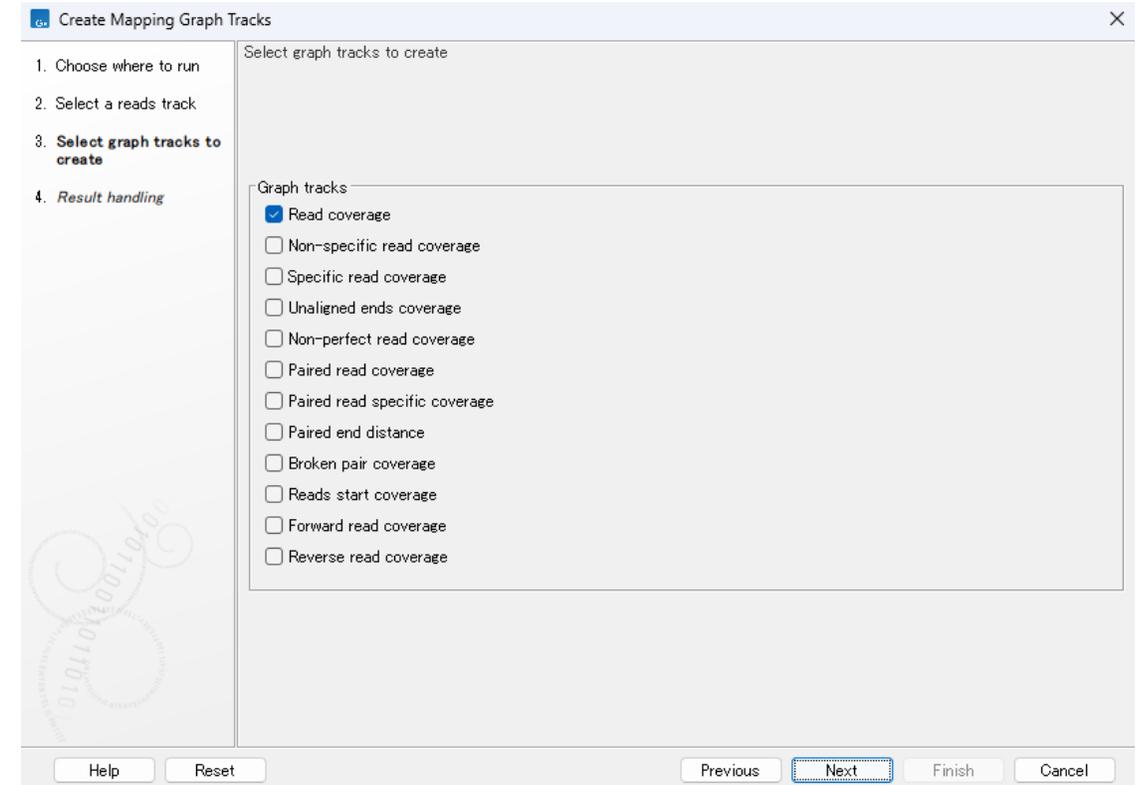
次に、ヒトゲノムデータを指定



## Create Mapping Graph



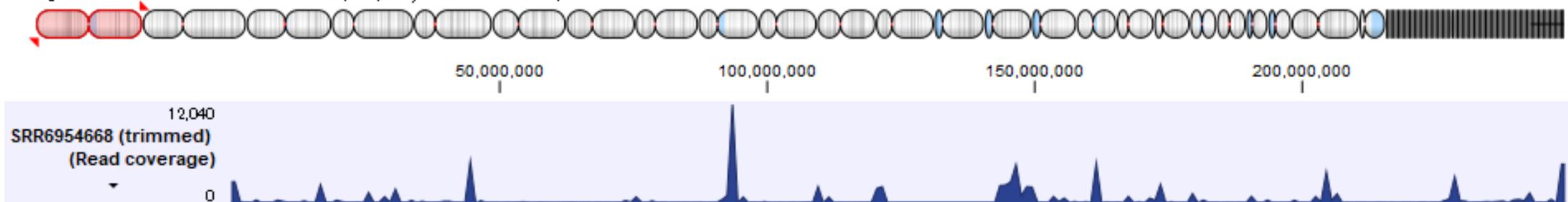
マッピングデータを指定



“Read coverage”を選択

## SRR6954668 (trimmed) (Read coverage)

Navigation overview: All chromosomes (195, only first 100 shown)



カバレッジグラフが作成される

## Identify Graph Threshold Areas

Identify Graph Threshold Areas

1. Choose where to run

Select graph track

Navigation Area

Q- <enter search term>

- SRR6954681 (trimmed, trimmed) batch
- SRR6954682 (trimmed, trimmed) batch
- Workflow Result Metadata
- UMI
  - novel miRNA
  - SRR6954665 (trimmed) (UMI reads) batch-1
  - SRR6954665 (trimmed) (UMI reads, unmapped re
  - SRR6954668 (trimmed) (Read coverage)**
  - SRR6954668 (trimmed) (UMI reads) batch-1
  - SRR6954668 (trimmed) (UMI reads, unmapped re
  - SRR6954672 (trimmed) (UMI reads) batch-1
  - SRR6954672 (trimmed) (UMI reads, unmapped re
  - SRR6954680 (trimmed) (Read coverage)
  - SRR6954680 (trimmed) (UMI reads) batch-1
  - SRR6954680 (trimmed) (UMI reads, unmapped re
  - SRR6954681 (trimmed) (UMI reads) batch-1
  - SRR6954681 (trimmed) (UMI reads, unmapped re
  - SRR6954682 (trimmed) (UMI reads) batch-1
  - SRR6954682 (trimmed) (UMI reads, unmapped re
- SRR6954665 (trimmed) (UMI reads) batch
- SRR6954668 (trimmed) (UMI reads) batch
- SRR6954672 (trimmed) (UMI reads) batch
- SRR6954680 (trimmed) (UMI reads) batch
- SRR6954681 (trimmed) (UMI reads) batch
- SRR6954682 (trimmed) (UMI reads) batch

Selected elements (1)

- SRR6954668 (trimmed) (Read coverage)

Batch

Help Reset Previous Next Finish Cancel

カバレッジグラフを指定

Identify Graph Threshold Areas

1. Choose where to run

2. Select graph track

3. Parameters

4. Result handling

Parameters

Threshold parameters

Window size 1

Use lower threshold

Lower threshold 100

Use upper threshold

Upper threshold 1000.0

Region parameters

No restriction

Exclude annotated

Include annotated only

Region track

Help Reset Previous Next Finish Cancel

Use lower thresholdにチェックを入れ、値を指定  
ここで指定した値以下のカバレッジの箇所は考慮されない

# 高カバレッジ領域の抽出

Rows: 176 Table view: Genome Filter to Selection\*\*\* Filter

Chromosome	Region	Name	Origin graph	Window size	Lower threshold	Upper threshold	Average value...
1	630832..630862	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,320.97
1	630904..630928	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,350.60
1	16514122..16514137	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,198.31
1	16666785..16666800	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,133.19
1	16740664..16740679	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,168.06
1	16896128..16896143	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,136.75
1	44778428..44778455	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	4,596.11
1	93847631..93847657	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	11,657.59
1	143699457..143699470	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,066.21
1	143729407..143729422	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,115.50
1	144560666..144560681	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,166.69
1	145157157..145157175	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,497.21
1	145465617..145465632	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,171.88
1	146035692..146035723	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	4,390.50
1	146376955..146376970	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,159.62
1	161443317..161443330	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,030.71
1	161447268..161447299	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	4,459.81
1	161454648..161454679	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	4,506.88
1	161462059..161462090	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	4,558.25
1	161465482..161465492	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,020.27
1	161469439..161469470	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	4,487.75
1	173864152..173864178	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,212.11
1	204506527..204506545	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	3,529.00
1	204507030..204507053	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,313.83
1	204507084..204507102	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	3,496.89
1	228558285..228558303	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,996.68
1	228610267..228610285	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	3,027.53
1	228612508..228612526	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,983.63
1	228614749..228614767	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,929.63
1	228616990..228617008	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	3,033.00
1	228619231..228619249	threshold	SRR6954668 (trimmed) (Read coverage)	1	2,000.00	(none)	2,947.37

➡ SRR6954668 (trimmed) (threshold)

ゲノム中から該当領域が抽出され、アノテーションデータとして出力される

## Extract Annotated Regions

The screenshot shows the 'Extract Annotated Regions' dialog box. On the left, a sidebar lists steps: 1. Choose where to run, 2. Select annotated sequences OR a track of annotations, variants or statistical comparisons, 3. Set parameters, and 4. Result handling. The main area is titled 'Select annotated sequences OR a track of annotations, variants or statistical comparisons'. It features a 'Navigation Area' with a search bar containing '<enter search term>'. Below the search bar is a tree view of a file system. The 'novel miRNA' folder is expanded, showing several sub-folders. The folder 'SRR6954668 (trimmed) (threshold)' is selected and highlighted in blue. To the right of the tree view is a 'Selected elements (1)' list containing the selected folder. At the bottom of the dialog are buttons for 'Help', 'Reset', 'Previous', 'Next', 'Finish', and 'Cancel'.

高カバレッジ領域のアノテーションデータを指定

The screenshot shows the 'Extract Annotated Regions' dialog box with the 'Set parameters' section active. The sidebar steps are the same as in the previous screenshot. The main area is titled 'Set parameters'. It contains several sections: 'Select a reference sequence track' with a dropdown menu showing 'Homo\_sapiens\_sequence\_hg38\_no\_alt\_analysis\_set'; 'Refine extracted annotations' with a 'Search terms' input field and an 'Annotation types' dropdown menu; 'Flanking' with 'Flanking upstream residues' and 'Flanking downstream residues' input fields, both set to '0'; and 'Naming of result sequences' with four checkboxes: 'Include annotation name' (checked), 'Include annotation type', 'Include annotation region', and 'Include sequence/track name'. At the bottom of the dialog are buttons for 'Help', 'Reset', 'Previous', 'Next', 'Finish', and 'Cancel'.

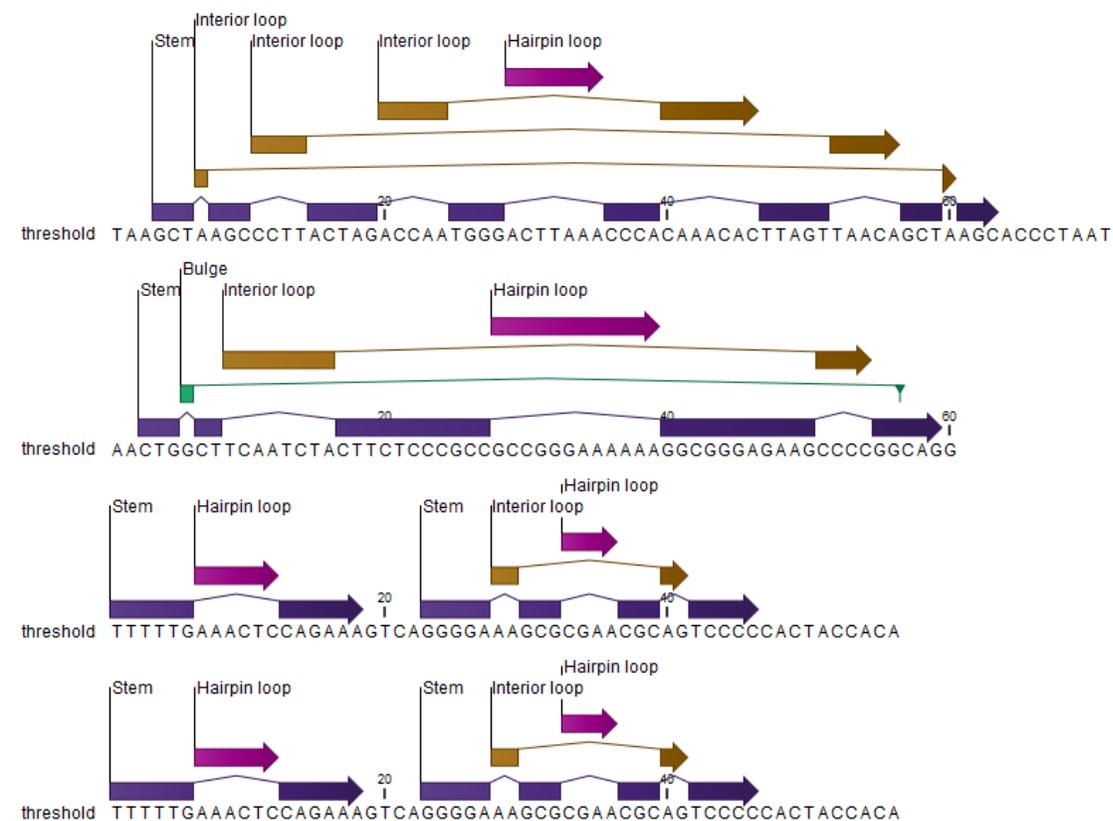
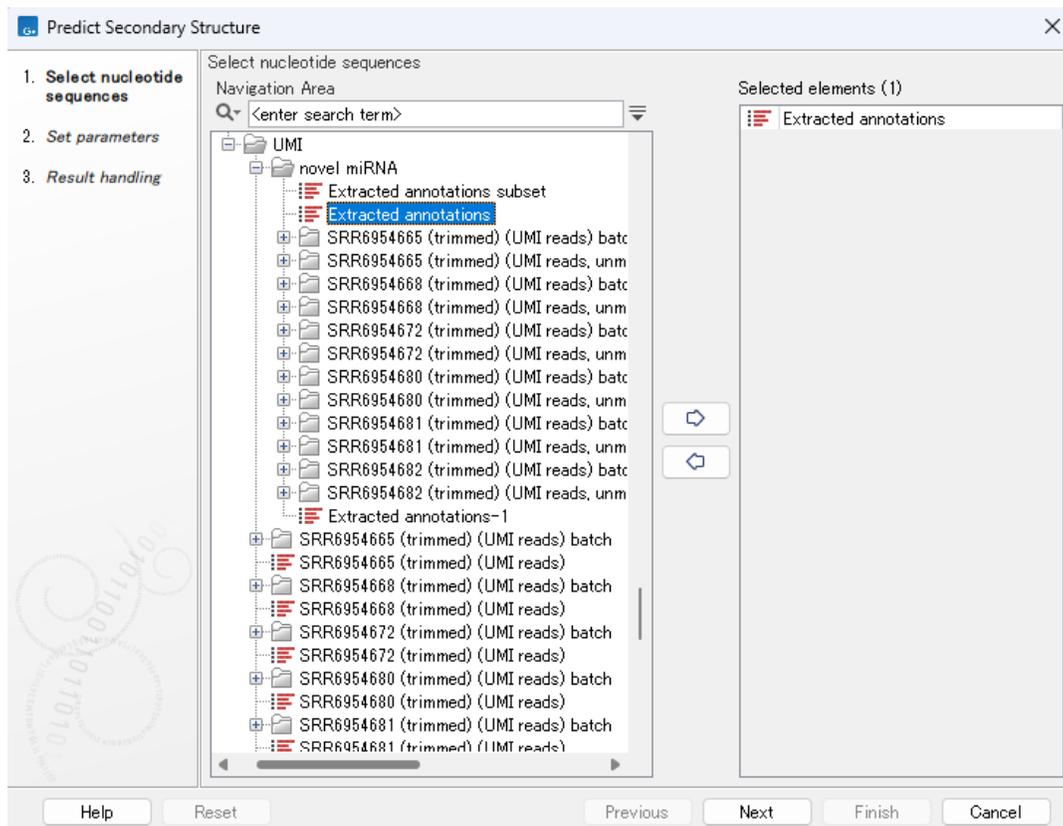
ヒトゲノムデータを指定

## ☰ Extracted annotations

```
threshold TAAGCTAAGCCCTTACTAGACCAATGGGACTTAAACCCACAAACACTTAGTTAACAGCTAAGCACCCCTAAT
threshold AACTGGCTTCAATCTACTTCTCCCGCCGCCGGGAAAAAAGGGCGGAGAAAGCCCCGGCAGG
threshold TTTTGAAGCTCCAGAAAGTCAGGGGAAAGCGGAACGCAGTCCCCCACTACCACA
threshold TTTTGAAGCTCCAGAAAGTCAGGGGAAAGCGGAACGCAGTCCCCCACTACCACA
threshold TGTGGTAGTGGGGGACTGCGTTCGCGCTTCCCTGACTTCTGGAGTTTCAAAA
threshold TGTGGTAGTGGGGGACTGCGTTCGCGCTTCCCTGACTTCTGGAGTTTCTAAA
threshold TGTCCAGTTCTGCTACTGACAGTAAGTGAAGATAAAGTGTGTCTGAGGAGACAGCTGGCTTCATGCTT
threshold AGGCTGAAGGCATTCAAAGTTCCGGGTTTCGAGTCCCGCGGAGTCGTAACGCTTTTTTCCCTCCCC
threshold CTTTTGCACTGGAAAATTCAGGGGAAAGCGGAACGCAGTCCCCCACTACCAC
threshold TTCTTTAAACTACAAAAAATCAGGGGAAAGCGGAACGCAGTCCCCCACTACCACA
threshold TTCTTTAAACTACAAAAAATCAGGGGAAAGCGGAACGCAGTCCCCCACTACCACA
threshold CCGCGCAGGTTCCGCGTGTGGTTTCCGTAGTGTAGTGGTTATCATGTTCCGCTCACACG
```

ゲノム配列中から、高カバレッジ領域の配列が抽出され、リスト化される

## Predict Secondary Structure



抽出された配列に対して、二次構造の予測を行うことで、ヘアピン構造等の有無を確認する

# 新規miRNAの定量・比較

Quantify miRNA

1. Choose where to run
2. Select trimmed reads ( $\leq 55\text{bp}$ )
3. Quality control
4. **Select annotation**
5. Specify reads used for seed counting
6. Result handling

Select annotation

miRBase annotations

miRBase miRBase-Release\_v22

Prioritized species Homo sapiens

Allow length-based isomiRs:

Additional upstream bases

Additional downstream bases

Missing upstream bases

Missing downstream bases

Custom databases

Database files Extracted annotations

Alignment settings

Maximum mismatches

Help Reset **Previous** Next Finish Cancel

Quantify miRNAツールのCustom databaseで、作成したリストを指定することで、このリストに含まれるmiRNAに関する定量データが得られる。

さらに、1章同様、この定量データを用いて発現変動解析を行うことも可能

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: [biosupport@filgen.jp](mailto:biosupport@filgen.jp)