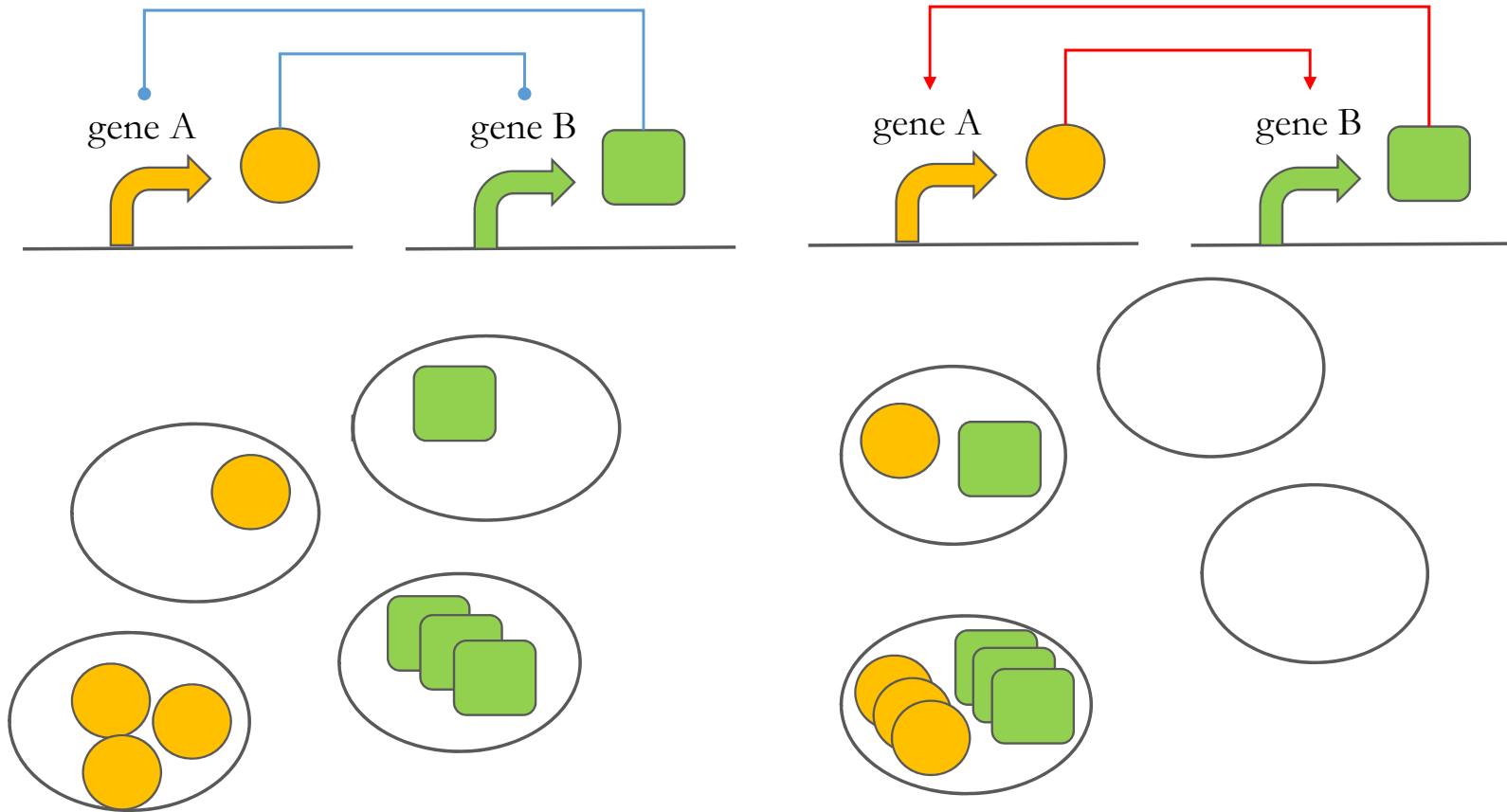


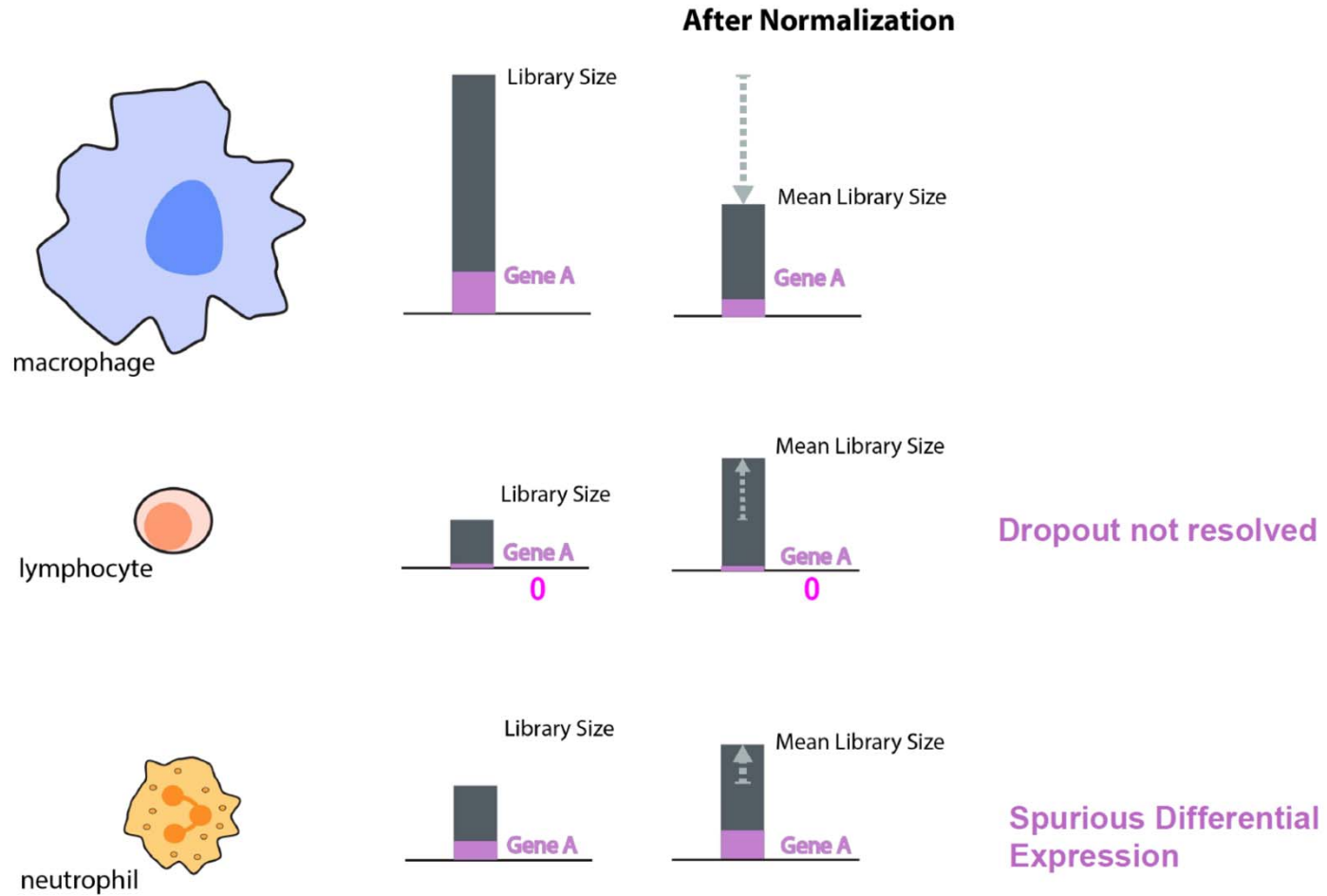
2019年度
第3回バイオインフォマティクス実習

先端医科学研究センター バイオインフォマティクス解析室
中林潤

scRNA-seq vs bulk RNA-seq



Problem with Global Normalization



libraryのコピー

- 課題配布フォルダ → Bioinfojishuフォルダ内の“R”フォルダを自分のZフォルダにコピー
- Rを起動

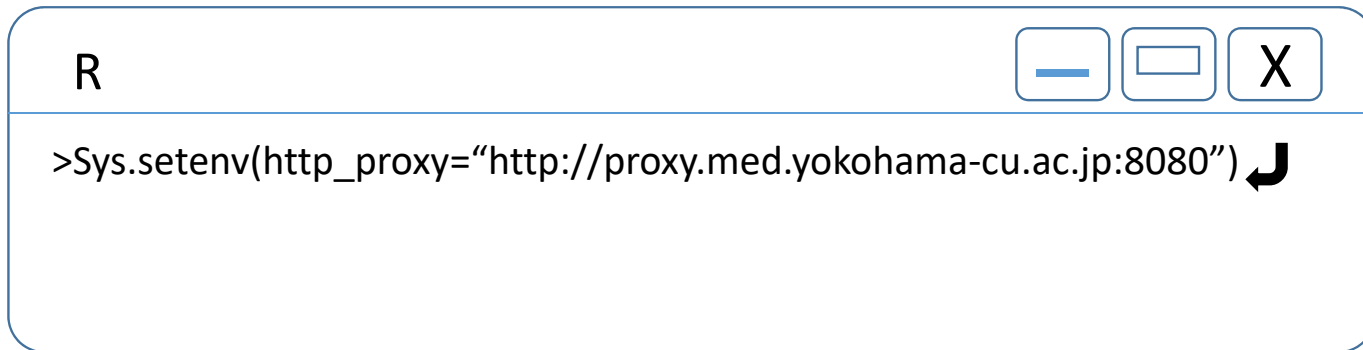
Seurat package

- RのscRNA-seqデータ解析用パッケージ

<https://satijalab.org/seurat/>

- データの正規化、細胞のクラスタ化、可視化などが行えます。

proxyの設定 作業フォルダの変更



The image shows a terminal window titled 'R'. The window has standard window control buttons (minimize, maximize, close) in the top right corner. The terminal content shows the command `>Sys.setenv(http_proxy="http://proxy.med.yokohama-cu.ac.jp:8080")` followed by a carriage return symbol (↵).

Rのファイルメニューの“作業フォルダの変更”で、作業フォルダを各自のデスクトップに変更します。

Seurat packageのロード

```
R
```

```
>library(Seurat) ↵
```

```
>library(dplyr) ↵
```

デモデータファイルのコピーと作業ディレクトリの変更

- 課題配布フォルダから `filtered_gene_bc_matrices` を各自のデスクトップにコピー
peripheral blood mononuclear cells (pbmc) 2700 cells
- Rのファイルメニューの作業ディレクトリの変更でデスクトップを選択

Seurat objectの作成

R



```
>pbmc.data <- Read10X("filtered_gene_bc_matrices/hg19") ↵
```

```
>pbmc<-CreateSeuratObject(counts=pbmc.data,project="pbmc3k",min.cells=3, min.features = 200) ↵
```

mitochondria geneの抽出

R



```
>pbmc[["percent.mt"]]<-PercentageFeatureSet(pbmc, pattern="^MT-") ↵
```

```
>VlnPlot(pbmc,features=c("nFeature_RNA","nCount_RNA","percent.mt"),ncol=3) ↵
```

細胞のフィルター

R



```
>pbmc <- subset(pbmc, subset=nFeature_RNA > 200 & nFeature_RNA < 2500 & percent.mt < 5) ↵
```

データの正規化

```
R _ □ X  
>pbmc <- NormalizeData(pbmc) ↵
```

トータルカウントで補正

発現変動遺伝子の抽出

R



```
>pbmc <- FindVariableFeatures(pbmc, selection.method="vst", nfeatures=2000) ↵
```

```
>top10<-head(VariableFeatures(pbmc),10) ↵
```

```
>VariableFeaturePlot(pbmc) ↵
```

```
>LabelPoints(plot=VariableFeaturePlot(pbmc), points=top10, repel=TRUE) ↵
```

データのスケールリング

```
R _ □ X  
>all.genes<-rownames(pbmc) ↵  
>pbmc <- ScaleData(pbmc, features=all.genes) ↵
```

細胞周期などの変動を除去

主成分分析 (PCA)

```
R
```

```
>pbmc <- RunPCA(pbmc, features = VariableFeatures(object=pbmc)) ↵
```

次元縮約法の一つ
データの分散が最大となる方向を第一主成分とする。

次元数を決定

R



```
>pbmc <- JackStraw(pbmc, num.replicate=100) ↵
```

```
>pbmc <- ScoreJackStraw(pbmc, dims=1:20) ↵
```

```
>ElbowPlot(pbmc) ↵
```


クラスター分析

R



```
>pbmc <- FindNeighbors(pbmc, dims=1:10) ↵
```

```
>pbmc<-FindClusters(pbmc, resolution=0.5) ↵
```

tSNEで次元縮約、可視化

R



```
>pbmc <- RunTSNE(pbmc, dims = 1:10) ↵
```

```
>DimPlot(pbmc, reduction="tsne") ↵
```

次元縮約法の一つ

高次元空間中で近傍にある点は、低次元空間でも、近傍となる確率が高くなるように次元を縮約する。

signature geneの抽出

R



```
>cluster1.markers <- FindMarkers(pbmc, ident.1=1, min.pct=0.25) ↵
```

```
>write.table(cluster1.markers, "cluster1.markers.txt", quote = F, sep = "¥t") ↵
```

```
>head(cluster1.markers, n=5) ↵
```

```
>VlnPlot(pbmc, features=c("IL32")) ↵
```

```
>FeaturePlot(pbmc, features=c("IL32")) ↵
```

宿題

- クラスタ-3の特徴的遺伝子を抽出し、GO解析を行って機能を推定してください。