

次世代シーケンサーデータの解析手法 第18回遺伝子発現データのクラスタリング ウェブ資料

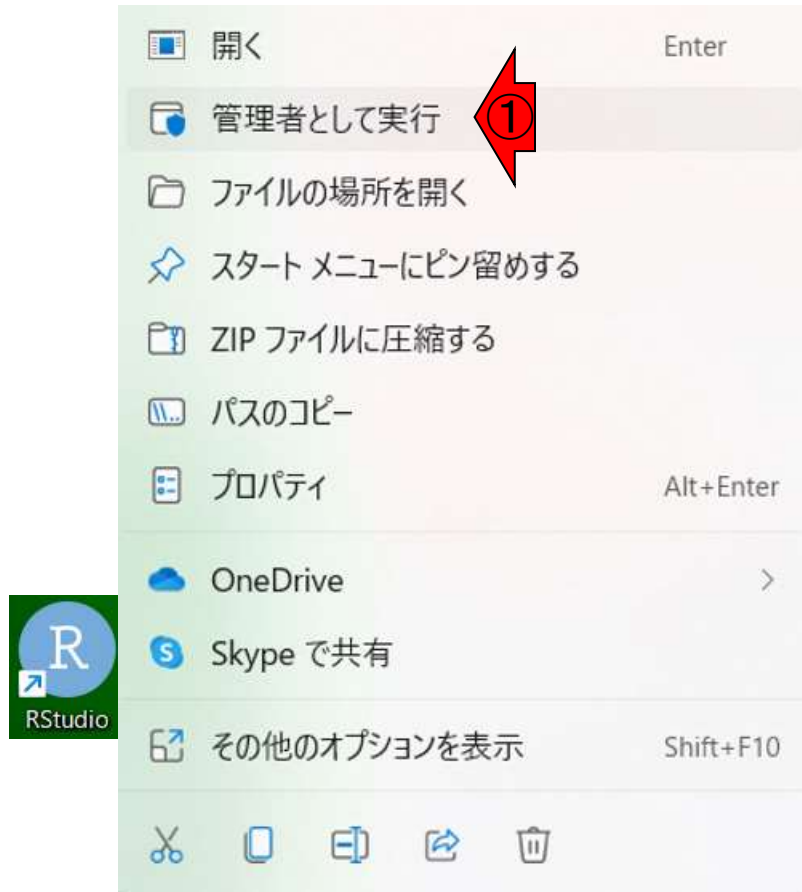
東京大学・大学院農学生命科学研究科
牧野磨音、清水謙多郎、門田幸二

Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W01 : RStudio起動0

RStudio起動時に「①管理者として実行」するのを忘れないようにしてください。理由は、W05でパッケージのインストールを行う際に不具合が生じやすいためです。



W01: RStudio起動1

RStudio起動時に「①管理者として実行」するのを忘れないようにしてください。理由は、W05でパッケージのインストールを行う際に不具合が生じやすいためです。起動後の状態。

The screenshot shows the RStudio interface. The console pane on the left displays the R version information and license details. The Environment pane on the right shows that the environment is empty. The Files pane at the bottom right shows the file explorer view of the user's home directory.

```
R version 4.2.0 (2022-04-22 ucrt) -- "Vigorous Calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

Environment: Global Environment

Environment is empty

Files: Home

Name	Size	Modified
その他		
Zoom		
R		
Officeのカスタム テンプレート		
My Videos		
My Pictures		
My Music		
MEGA X		
html		
desktop.ini	402 B	Dec 3, 2021
biostat		
2022		
2021		
2020		

W01: RStudio起動2

RStudio起動時に「①管理者として実行」するのを忘れないようにしてください。理由は、W05でパッケージのインストールを行う際に不具合が生じやすいためです。起動後の状態。②Filesタブで見えている赤枠内はWindowsのヒトはドキュメントフォルダ直下の情報が見えています。

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
R 4.2.0 · ~/
R version 4.2.0 (2022-04-22 ucrt) -- "vigorous calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
>
```

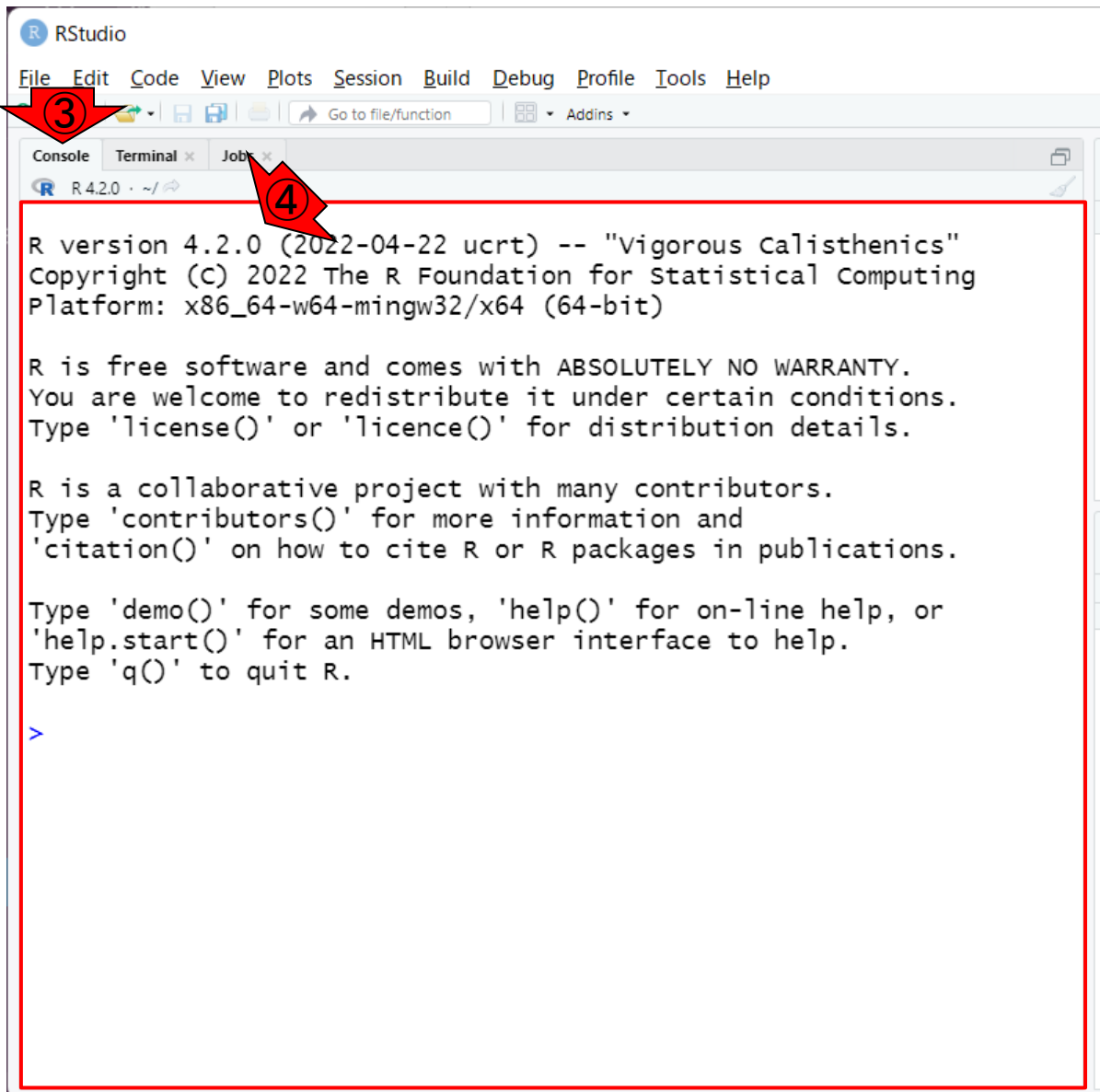


Files Plots Packages Help Viewer

Home

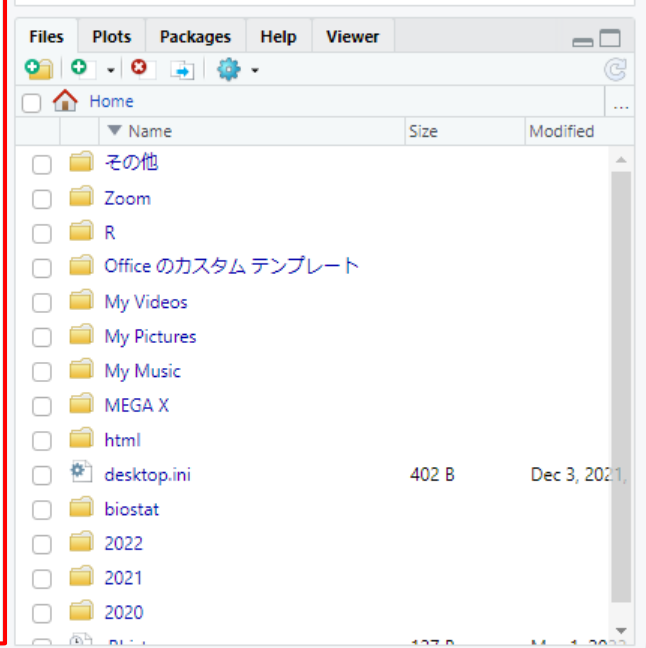
	Name	Size	Modified
<input type="checkbox"/>	その他		
<input type="checkbox"/>	Zoom		
<input type="checkbox"/>	R		
<input type="checkbox"/>	Officeのカスタム テンプレート		
<input type="checkbox"/>	My Videos		
<input type="checkbox"/>	My Pictures		
<input type="checkbox"/>	My Music		
<input type="checkbox"/>	MEGA X		
<input type="checkbox"/>	html		
<input type="checkbox"/>	desktop.ini	402 B	Dec 3, 2021
<input type="checkbox"/>	biostat		
<input type="checkbox"/>	2022		
<input type="checkbox"/>	2021		
<input type="checkbox"/>	2020		

W01: RStudio起動3



```
RStudio  
File Edit Code View Plots Session Build Debug Profile Tools Help  
Go to file/function Addins  
Console Terminal x Job x  
R 4.2.0 · ~/>  
R version 4.2.0 (2022-04-22 ucrt) -- "Vigorous calisthenics"  
Copyright (C) 2022 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
>
```

RStudio起動時に「①管理者として実行」するのを忘れないようにしてください。理由は、W05でパッケージのインストールを行う際に不具合が生じやすいためです。起動後の状態。②Filesタブで見えている赤枠内はWindowsのヒトはドキュメントフォルダ直下の情報が見えています。③Consoleタブで見えている赤枠内がConsole画面と呼ばれるものです。④今回の内容はR ver. 4.2.0で行っています。



Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W02: ファイルの保存1

(Rで)塩基配列解析のサブ

(last modified 2022/05/01, since 2018)



@Agribio_utokyoさんをフォロー

、[アクリバイオ](#)、[\(Rで\)塩基配列解析のメインページ](#)、[生命情報解析研究室](#)

What's new?

- 「書籍 | 日本乳酸菌学会誌 | [第17回バイオ](#)」の第13回に原稿はまだ公開されていませんが、その間に修正がなされていたので修正しました。(2022/05/01)
- [東京大学・大学院農学生命科学研究科・応用微生物学研究室](#)にて、2022年5月7日(土)と5月28日(土)に「書籍 | 日本乳酸菌学会誌 | [第15回RNA-seq解析 \(その1\)](#)」のリンク切れを修正しました。

(Rで)塩基配列解析のサブ

iu.a.u-tokyo.ac.jp/~kadota/r_seq2.html

- 書籍 | 日本乳酸菌学会誌 | [第5回アセンブル、マッピング、そしてQC](#) (last modified 2017/06/25)
- 書籍 | 日本乳酸菌学会誌 | [第6回ゲノムアセンブリ](#) (last modified 2017/06/21)
- 書籍 | 日本乳酸菌学会誌 | [第7回ロングリードアセンブリ](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第8回アセンブリ後の解析](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第9回ゲノムアノテーションとその可視化、DDBJへの登録](#) (last modified 2017/03/13)
- 書籍 | 日本乳酸菌学会誌 | [第10回DDBJへの塩基配列の登録\(後編\)](#) (last modified 2017/06/28)
- 書籍 | 日本乳酸菌学会誌 | [第11回総合データ解析環境Galaxy](#) (last modified 2017/11/13)
- 書籍 | 日本乳酸菌学会誌 | [第12回Galaxy: ヒストリーとワークフロー](#) (last modified 2018/07/04)
- 書籍 | 日本乳酸菌学会誌 | [第13回RNA-seq解析 \(その1\)](#) (last modified 2022/05/01) **NEW**
- 書籍 | 日本乳酸菌学会誌 | [第14回RNA-seq解析 \(その2\)](#) (last modified 2022/05/01) **NEW**
- 書籍 | 日本乳酸菌学会誌 | [第15回RNA-seq解析 \(その3\)](#) (last modified 2022/05/01) **NEW**
- 書籍 | 日本乳酸菌学会誌 | [第16回なぜ次から次へと新規手法が開発されるのか?](#) (last modified 2021/08/16)
- 書籍 | 日本乳酸菌学会誌 | [第17回バイオインフォマティクス教育の今後](#) (last modified 2022/05/01) **NEW**
- 書籍 | 日本乳酸菌学会誌 | [第18回遺伝子発現データのクラスタリング](#) (last modified 2022/05/01) **NEW**



はじめに

[トップページへ](#)

W02: ファイルの保存2

①のページの、②第18回のリンク先へ。
③のところに飛ぶので、今回用いる④と⑤のファイルを、⑥デスクトップ上に作成したhogeフォルダ内に保存(されているという前提で話を進めていきます)。

書籍 | 日本乳酸菌学会誌 | 第18回遺伝子発現データのクラスタリング NEW **③**

日本乳酸菌学会誌の第18回分です。

- [原稿PDF](#)
- [ウェブ資料PDF](#)

はじめに

- [RStudio](#)
- [R](#)
- [連載第1回の原稿PDF](#)
- [連載第17回の原稿PDF](#)(リンク先はまだありま
- [アグリバイオインフォマティクス教育研究プ](#)
- [JSLAB18.xlsx](#) **④**
第15回のW12-1のxlsxファイル(Kallisto_9sa
ータ以外の部分)を使ったものです。
- [JSLAB18.R](#) **⑤**

hoge

新規作成

C:\Users\kadota\Desktop\hoge **⑥**

名前	更新日時	種類	サイズ
JSLAB18.xlsx	2022/05/02 20:53	Microsoft Excel ワ...	187 KB
JSLAB18.R	2022/05/02 20:53	R ファイル	5 KB

2 個の項目

Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W03: 作業ディレクトリ...1

RStudioの画面を再掲。①Filesタブで見えている赤枠内は、Windowsのヒトはドキュメントフォルダ直下の情報が見えています。これを先ほど保存した「デスクトップ上のhogeフォルダの中身が見える状態」にします。②をクリックすると…

The screenshot shows the RStudio interface. The console window displays the R version 4.2.0 startup message and instructions. The file explorer window is open to the 'Home' directory, showing a list of folders and files. A red box highlights the file explorer content, and a red arrow labeled '1' points to the 'Files' tab. Another red arrow labeled '2' points to the 'View' button in the file explorer toolbar.

```
R version 4.2.0 (2022-04-22 ucrt) -- "Vigorous Calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

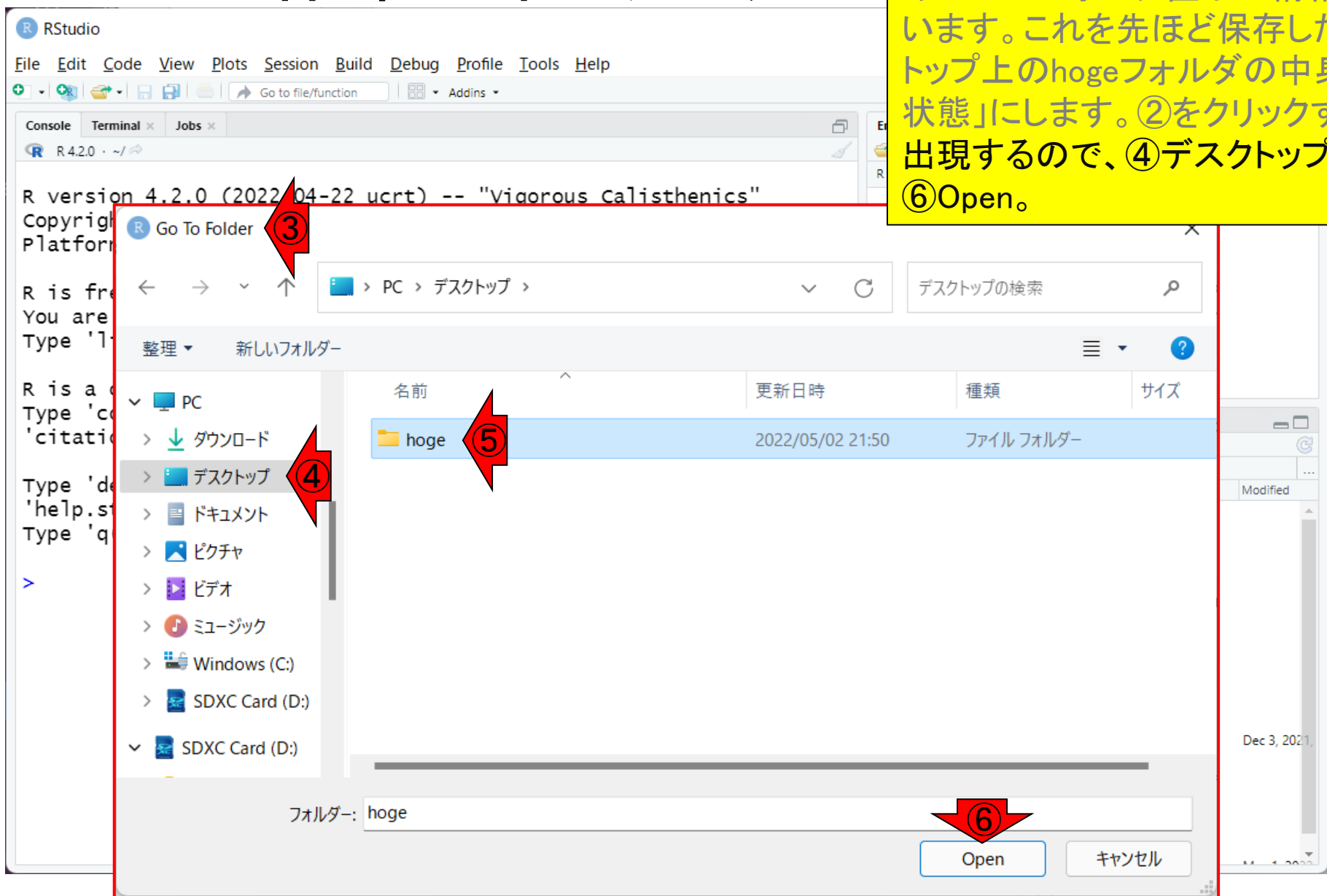
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

Name	Size	Modified
その他		
Zoom		
R		
Officeのカスタム テンプレート		
My Videos		
My Pictures		
My Music		
MEGA X		
html		
desktop.ini	402 B	Dec 3, 2021
biostat		
2022		
2021		
2020		

W03: 作業ディレクトリ...2

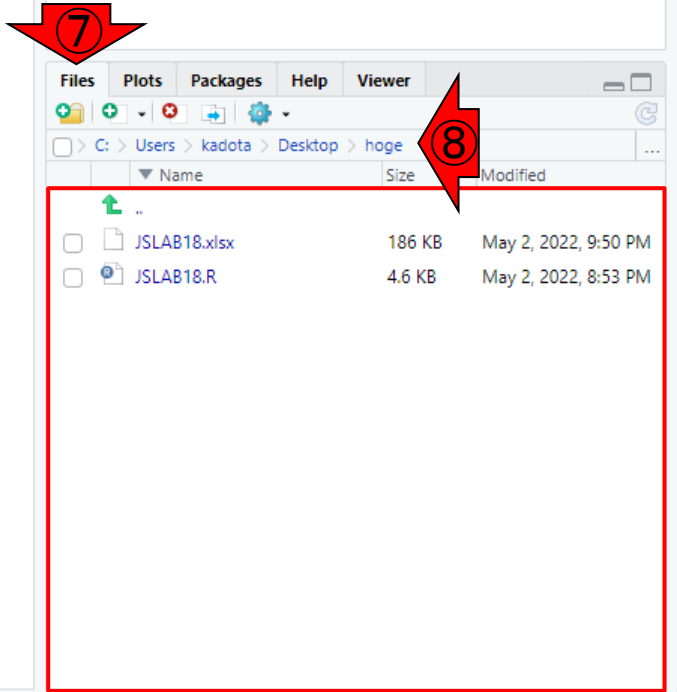
RStudioの画面を再掲。①Filesタブで見えている赤枠内は、Windowsのヒトはドキュメントフォルダ直下の情報が見えています。これを先ほど保存した「デスクトップ上のhogeフォルダの中身が見える状態」にします。②をクリックすると、③が出現するので、④デスクトップ、⑤hoge、⑥Open。



W03: 作業ディレクトリ...3

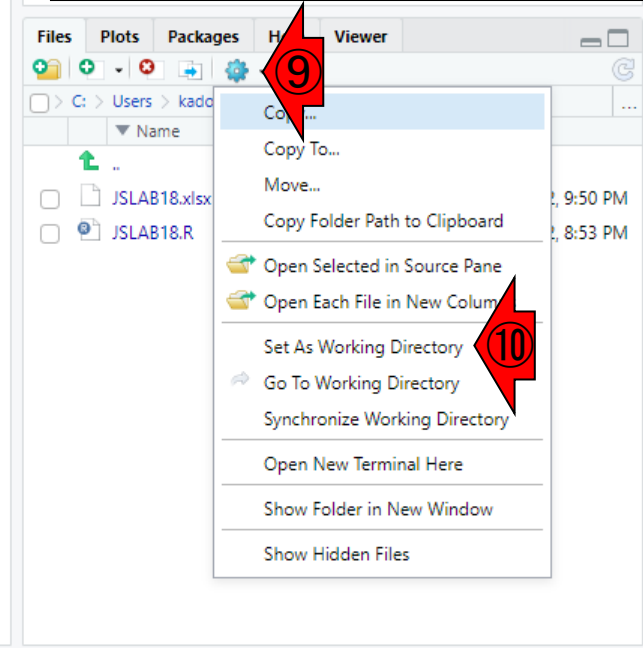
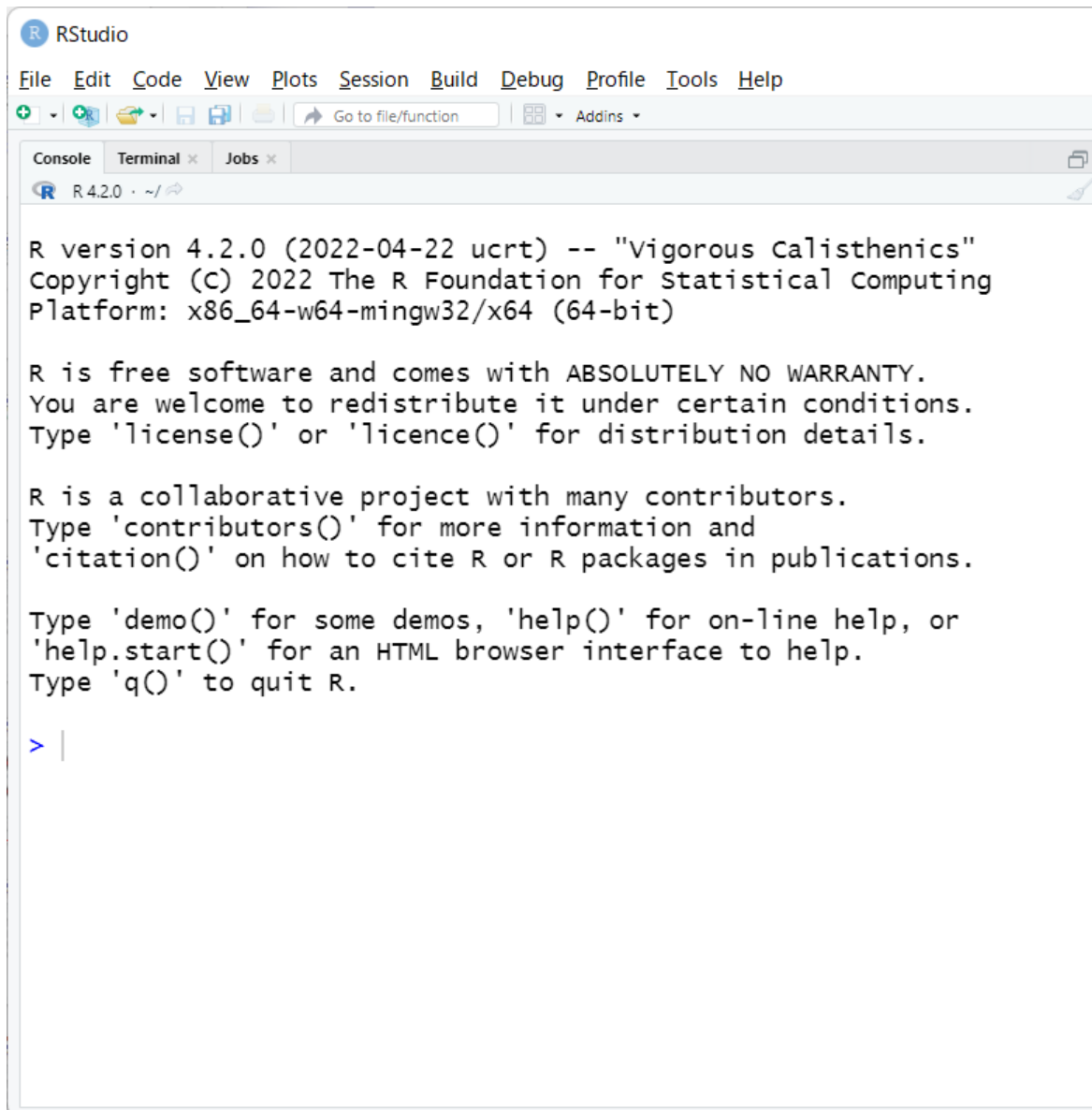
RStudioの画面を再掲。①Filesタブで見えている赤枠内は、Windowsのヒトはドキュメントフォルダ直下の情報が見えています。これを先ほど保存した「デスクトップ上のhogeフォルダの中身が見える状態」にします。②をクリックすると、③が出現するので、④デスクトップ、⑤hoge、⑥Open。こんな感じになって、⑦Filesタブの中身が、⑧デスクトップ上のhogeフォルダのものになればまずはOK。まだこれで終わりではない!

```
RStudio  
File Edit Code View Plots Session Build Debug Profile Tools Help  
Go to file/function Addins  
R 4.2.0 · ~/ |  
R version 4.2.0 (2022-04-22 ucrt) -- "Vigorous calisthenics"  
Copyright (C) 2022 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (64-bit)  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
> |
```



W03: 作業ディレクトリ...4

RStudioの画面を再掲。①Filesタブで見えている赤枠内は、Windowsのヒトはドキュメントフォルダ直下の情報が見えています。これを先ほど保存した「デスクトップ上のhogeフォルダの中身が見える状態」にします。②をクリックすると、③が出現するので、④デスクトップ、⑤hoge、⑥Open。こんな感じになって、⑦Filesタブの中身が、⑧デスクトップ上のhogeフォルダのものになればまずはOK。まだこれで終わりではない!⑨をクリックして、⑩Set As Working Directory。



W03: 作業ディレクトリ...5

①のようなコマンドが自動的に実行されればOK。②の部分はヒトそれぞれ。

RStudio interface showing the terminal and file explorer. The terminal displays the R version and the command `setwd("C:/Users/kadota/Desktop/hoge")` is entered. Red arrows and numbers 1 and 2 highlight the command and its arguments respectively. The file explorer shows the directory `C:/Users/kadota/Desktop/hoge` containing files `JSLAB18.xlsx` and `JSLAB18.R`.

```
R version 4.2.0 (2022-04-22 ucrt) -- "vigorous calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

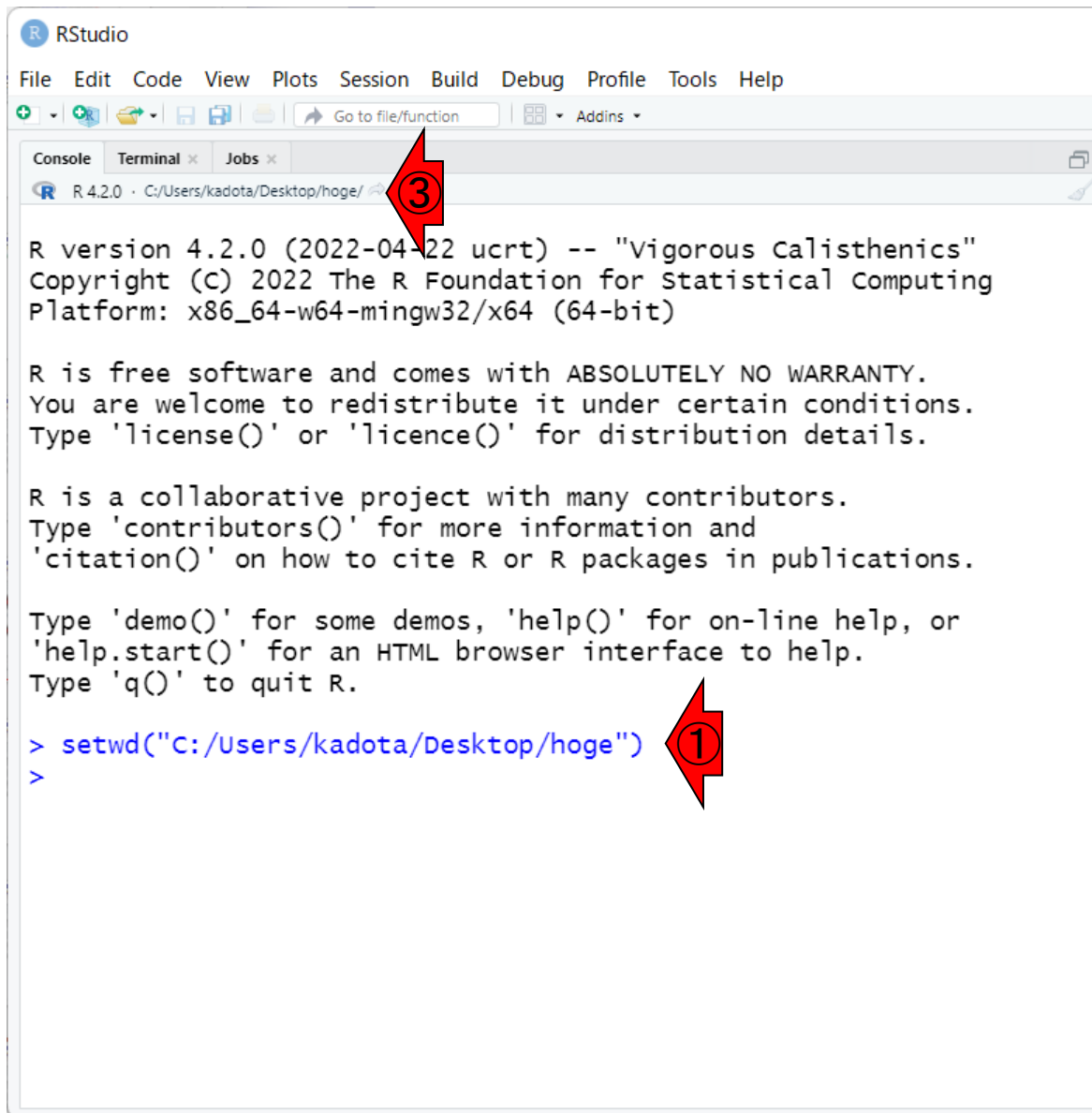
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> setwd("C:/Users/kadota/Desktop/hoge")
>
```

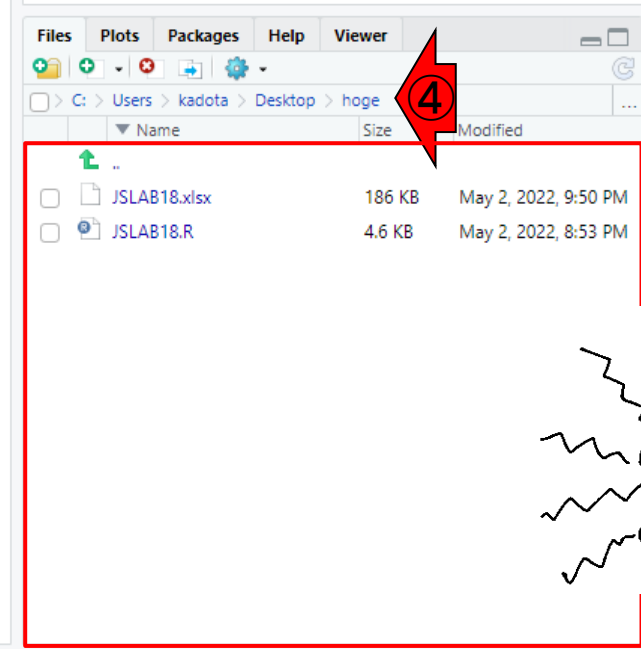
Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.6 KB	May 2, 2022, 8:53 PM

W03: 作業ディレクトリ...6

①のようなコマンドが自動的に実行されればOK。②の部分はヒトそれぞれ。①の実行と同時に③の部分も変わります。③と④は実質的に同じであることがわかりますが、初心者のうちは③と④が同じであることに注意を払うようにしましょう。初心者のやりがちなミスのは、作業ディレクトリの変更ができてないとか、赤枠内に存在しないファイルを入力として指定していることに起因します。



```
RStudio  
File Edit Code View Plots Session Build Debug Profile Tools Help  
Go to file/function Addins  
R 4.2.0 · C:/Users/kadota/Desktop/hoge/ ③  
R version 4.2.0 (2022-04-22 ucrt) -- "vigorous calisthenics"  
Copyright (C) 2022 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
> setwd("C:/Users/kadota/Desktop/hoge") ①  
>
```



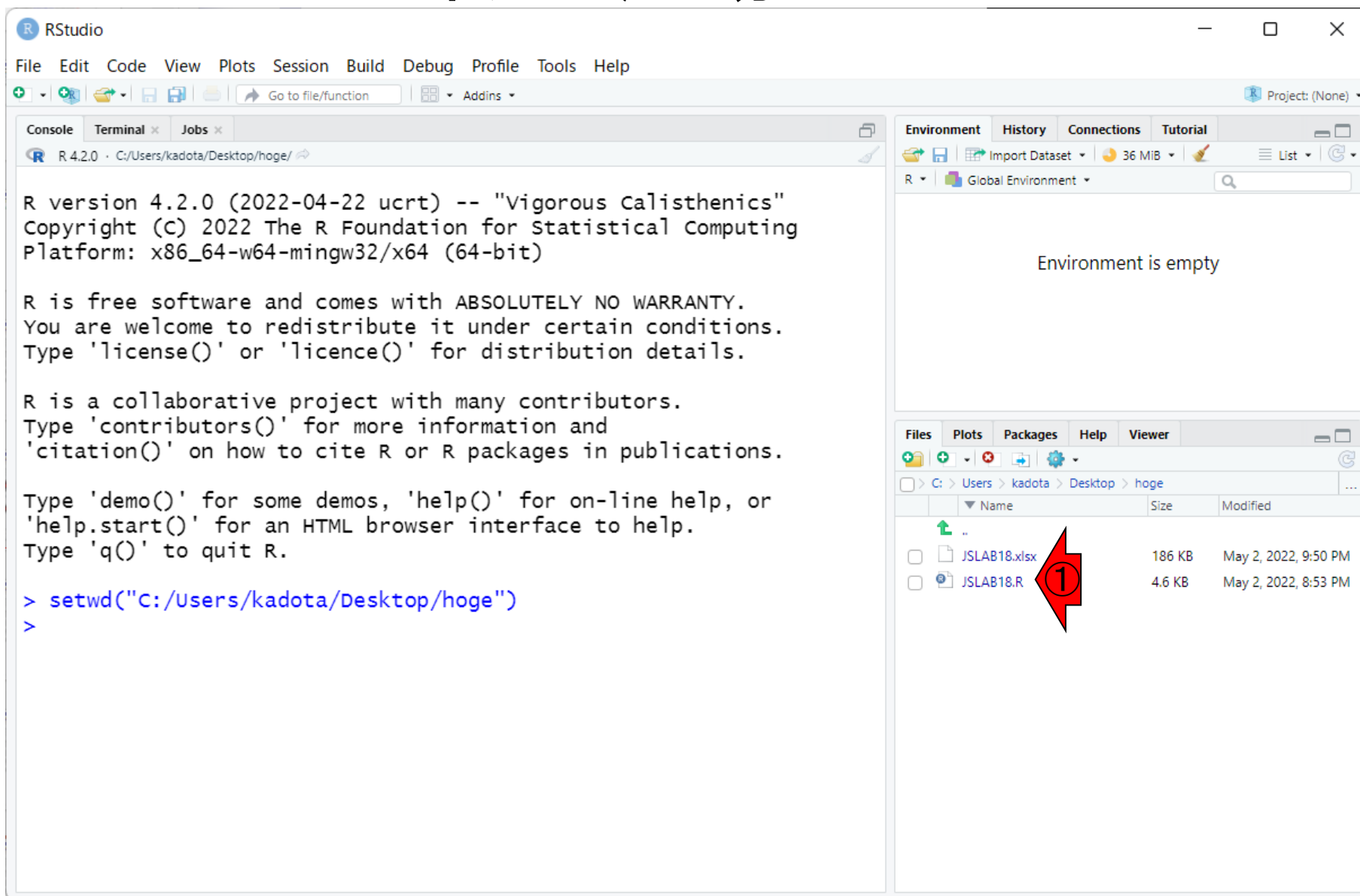
Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.6 KB	May 2, 2022, 8:53 PM



Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- **W04: エディタの起動**
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W04: エディタの起動1



RStudio interface showing the console output and the file explorer.

Console Output:

```
R version 4.2.0 (2022-04-22 ucrt) -- "vigorous calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> setwd("C:/Users/kadota/Desktop/hoge")
>
```

File Explorer:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.6 KB	May 2, 2022, 8:53 PM

A red arrow with the number 1 points to the JSLAB18.R file.

W04: エディタの起動2

①JSLAB18.Rのほうをクリック。こんな感じになります。②JSLAB18.Rの中身が赤枠内で見えているものになります。③がさきほどまで左上に見えていたConsoleタブになります。これが原稿中の図2と同じものです。

The screenshot shows the RStudio interface with the following elements:

- Editor:** A script named 'JSLAB18.R' is open. The code is as follows:

```
1 ## w04:パッケージのインストール ##
2 if (!requireNamespace("BiocManager", quietly=T))
3   install.packages("BiocManager")
4 BiocManager::install("TCC", update=F)
5 BiocManager::install("MBCluster.Seq", update=F)
6 BiocManager::install("openxlsx", update=F)
7 BiocManager::install("ggplot2", update=F)
8 BiocManager::install("ggsci", update=F)
9 BiocManager::install("reshape2", update=F)
10
```

A red box highlights the code from line 2 to line 9.
- Environment:** The Environment pane on the right shows 'Environment is empty'.
- Files:** The Files pane at the bottom shows the directory 'C:\Users\kadota\Desktop\hoge'. It contains two files:

Name	Size	Modified
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	5.2 KB	May 3, 2022, 1:17 PM

A red arrow points to the 'JSLAB18.R' file.
- Console:** The Console pane at the bottom shows the R startup output:

```
R version 4.2.0 (2022-04-22 ucrt) -- "Vigorous Calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
```

A red arrow points to the top of the console pane.

W04: エディタの起動3

ちょっと細かい話になりますが、④のところにメッセージが出ているように、⑤をクリックすれば、CRANという場所から提供されている⑥のパッケージはインストールされます。しかしBioconductorという場所から提供されている⑦TCCパッケージはインストールされません。しかし初心者のうちは、どのパッケージがCRAN or Bioconductorで提供されているかもわからないので、ここでは⑤ではなく④の Don't Show Againをクリックしてください。

The screenshot shows the RStudio interface. The top pane displays an R script with the following code:

```
1 ## w04:パッケージのインストール ##
2 if (!requireNamespace("BiocManager", quietly=F))
3   install.packages("BiocManager")
4 BiocManager::install("TCC", update=F)
5 BiocManager::install("MBCluster.Seq", update=F)
6 BiocManager::install("openxlsx", update=F)
7 BiocManager::install("ggplot2", update=F)
8 BiocManager::install("ggsci", update=F)
9 BiocManager::install("reshape2", update=F)
10
11
```

Red arrows point to specific elements: ⑤ points to the 'Install' button in the message bar; ④ points to the 'Don't Show Again' button; ⑦ points to the 'TCC' package name in the script; ⑥ points to the list of packages being installed. The bottom pane shows the R console output:

```
R version 4.2.0 (2022-04-22 ucrt) -- "vigorous calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
```

The right pane shows a file explorer window with the following table:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	5.2 KB	May 3, 2022, 1:17 PM

W04: エディタの起動4

ちょっと細かい話になりますが、④のところにメッセージが出ているように、⑤をクリックすれば、CRANという場所から提供されている⑥のパッケージはインストールされます。しかしBioconductorという場所から提供されている⑦TCCパッケージはインストールされません。しかし初心者のうちは、どのパッケージがCRAN or Bioconductorで提供されているかもわからないので、ここでは⑤ではなく④のDon't Show Againをクリックしてください。こんな感じになります。

The screenshot shows the RStudio interface. The script editor contains the following code:

```
1 ## w04:パッケージのインストール ##
2 if (!requireNamespace("BiocManager", quietly=T))
3   install.packages("BiocManager")
4 BiocManager::install("TCC", update=F)
5 BiocManager::install("MBCluster.Seq", update=F)
6 BiocManager::install("openxlsx", update=F)
7 BiocManager::install("ggplot2", update=F)
8 BiocManager::install("ggsci", update=F)
9 BiocManager::install("reshape2", update=F)
10
11 ## w05:パッケージのロード ##
12
```

The console output shows the R version and system information:

```
R version 4.2.0 (2022-04-22 ucrt) -- "Vigorous Calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
```

The file explorer on the right shows the directory structure:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	5.2 KB	May 3, 2022, 1:17 PM

Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W05: パッケージのイ...1

Rエディタ内の2-9行目を反転させて、① Run。

The screenshot shows the RStudio interface with a script editor containing the following R code:

```
1 ##### w04:パッケージのインストール #####
2 if (!requireNamespace("BiocManager", quietly = TRUE))
3   install.packages("BiocManager")
4 BiocManager::install("TCC", update=F)
5 BiocManager::install("MBCluster.Seq", update=F)
6 BiocManager::install("openxlsx", update=F)
7 BiocManager::install("ggplot2", update=F)
8 BiocManager::install("ggsci", update=F)
9 BiocManager::install("reshape2", update=F)
10
11
```

A red arrow labeled '①' points to the Run button in the toolbar. A tooltip over the Run button reads: "Run the current line or selection (Ctrl+Enter)".

The Environment pane on the right shows "Environment is empty".

The Files pane at the bottom right shows the directory structure: C:\Users\kadota\Desktop\hoge, containing files JSLAB18.xlsx (186 KB, May 2, 2022, 9:50 PM) and JSLAB18.R (5.2 KB, May 3, 2022, 7:45 PM).

The Console pane at the bottom shows the R startup output:

```
R version 4.2.0 (2022-04-22 ucrt) -- "Vigorous Calisthenics"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
```

W05: パッケージのイ...2

Rエディタ内の2-9行目を反転させて、① Run。実行直後の状態。

The screenshot shows the RStudio interface with the following content:

```
1 ##### w04:パッケージのインストール #####
2 if (!requireNamespace("BiocManager", quietly=T))
3   install.packages("BiocManager")
4 BiocManager::install("TCC", update=F)
5 BiocManager::install("MBCluster.Seq", update=F)
6 BiocManager::install("openxlsx", update=F)
7 BiocManager::install("ggplot2", update=F)
8 BiocManager::install("ggsci", update=F)
9 BiocManager::install("reshape2", update=F)
10
11
```

The console output shows the following text:

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
citation() on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> setwd("C:/Users/kadota/Desktop/hoge")
> if (!requireNamespace("BiocManager", quietly=T))
+   install.packages("BiocManager")
WARNING: Rtools is required to build R packages but is not current
ly installed. Please download and install the appropriate version
of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
```

The Environment pane shows "Environment is empty". The Files pane shows the directory structure:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	5.2 KB	May 3, 2022, 7:45 PM

W05: パッケージのイ...3

Rエディタ内の2-9行目を反転させて、① Run。実行直後の状態。実行完了後の状態。約5分。②「reshape2は無事に展開され…」というフレーズからも、③最後のコマンドが無事実行されていることが想像できる。

The screenshot shows the RStudio interface with a script editor, a console, and a file explorer. The script editor contains R code for installing packages. The console shows the execution progress, including the download of 'reshape2_1.4.4.zip' and the successful installation of 'plyr' and 'reshape2'. A file explorer window shows the local file system with files 'JSLAB18.xlsx' and 'JSLAB18.R'.

```
1 ##### w04:パッケージのインストール #####
2 if (!requireNamespace("BiocManager", quietly=T))
3   install.packages("BiocManager")
4 BiocManager::install("TCC", update=F)
5 BiocManager::install("MBCluster.Seq", update=F)
6 BiocManager::install("openxlsx", update=F)
7 BiocManager::install("ggplot2", update=F)
8 BiocManager::install("ggsci", update=F)
9 BiocManager::install("reshape2", update=F)
10
11
```

Console output:

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
Content type application/zip length 1151186 bytes (1.1 MB)
downloaded 1.1 MB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/reshape2_1.4.4.zip'
Content type 'application/zip' length 446586 bytes (436 KB)
downloaded 436 KB

パッケージ 'plyr' は無事に展開され、MD5 サムもチェックされました
パッケージ 'reshape2' は無事に展開され、MD5 サムもチェックされました

ダウンロードされたパッケージは、以下にあります
C:\Users\kadota\AppData\Local\Temp\RtmpQJ8LS6\downloaded_packages
>
```

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	5.2 KB	May 3, 2022, 7:45 PM

Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W06: Tips (ウィンドウ) 1

- ①11行目が先頭行になるようにした状態。
- ②赤線あたりにマウスポインタをおいて上下に動かすことで、エディタ(or Console)画面サイズを調整できます。

The screenshot shows the RStudio interface with the following components:

- Code Editor:** Contains R code for loading packages. Line 11 is highlighted with a red arrow labeled '1'. A red horizontal line is drawn across the editor, with a red arrow labeled '2' pointing to it.
- Environment Panel:** Shows 'Global Environment' and 'Environment is empty'.
- Files Panel:** Shows a file explorer view of 'C:\Users\kadota\Desktop\hoge' with files 'JSLAB18.xlsx' and 'JSLAB18.R'.
- Console:** Shows the output of the R script, including package download progress and messages for 'plyr' and 'reshape2'.

```
11 ##### w07:パッケージのロード(1回目) #####
12 library(TCC)
13 library(MBCluster.Seq)
14 library(openxlsx)
15 library(ggplot2)
16 library(ggsci)
17 library(reshape2)
18
19 ##### w07:パッケージのロード(2回目) #####
20 library(TCC)
21
```

Console Output:

```
R 4.2.0 · C:\Users\kadota\Desktop\hoge/
Content type application/zip length 1151186 bytes (1.1 MB)
downloaded 1.1 MB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/reshape2_1.4.4.zip'
Content type 'application/zip' length 446586 bytes (436 KB)
downloaded 436 KB

パッケージ 'plyr' は無事に展開され、MD5 サムもチェックされました
パッケージ 'reshape2' は無事に展開され、MD5 サムもチェックされました

ダウンロードされたパッケージは、以下にあります
C:\Users\kadota\AppData\Local\Temp\RtmpQJ8LS6\downloaded_packages
>
```

W06: Tips (ウィンドウ) 2

- ① 11行目が先頭行になるようにした状態。
- ② 赤線あたりにマウスポインタをおいて上下に動かすことで、エディタ(or Console)画面サイズを調整できます。
- ③ の部分を左右に調整することもできます。

The screenshot shows the RStudio interface with the following components:

- Code Editor:** Contains R code for loading packages. Line 11 is highlighted as the first line of the first code block. A red vertical line is positioned between the code editor and the environment pane.
- Environment Pane:** Shows "Environment is empty".
- Files Pane:** Shows a file explorer view of the directory "C:\Users\kadota\Desktop\hoge" with files "JSLAB18.xlsx" and "JSLAB18.R".
- Console:** Shows the output of the R script, including package download progress and messages for 'plyr' and 'reshape2'.

A red arrow points to a vertical separator between the code editor and the environment pane, labeled with a circled '3'.

W06: Tips (ウィンドウ) 3

- ①11行目が先頭行になるようにした状態。
- ②赤線あたりにマウスポインタをおいて上下に動かすことで、エディタ(or Console)画面サイズを調整できます。
- ③の部分左右に調整することもできます。
- 一般的なソフトウェアのウィンドウサイズ調整同様に、④を押せばREディタ画面の最小化…

The screenshot shows the RStudio interface. The script editor displays R code for loading packages. The console shows the execution output, including package installation progress and messages for 'plyr' and 'reshape2'. A red arrow labeled '4' points to the window control buttons (minimize, maximize, close) in the top right corner of the script editor window.

```
11 ##### w07:パッケージのロード(1回目) #####
12 library(TCC)
13 library(MBCluster.Seq)
14 library(openxlsx)
15 library(ggplot2)
16 library(ggsci)
17 library(reshape2)
18
19 ##### w07:パッケージのロード(2回目) #####
20 library(TCC)
21 library(MBCluster.Seq)
22 library(openxlsx)
23 library(ggplot2)
24 library(ggsci)
25 library(reshape2)
26
```

Console output:

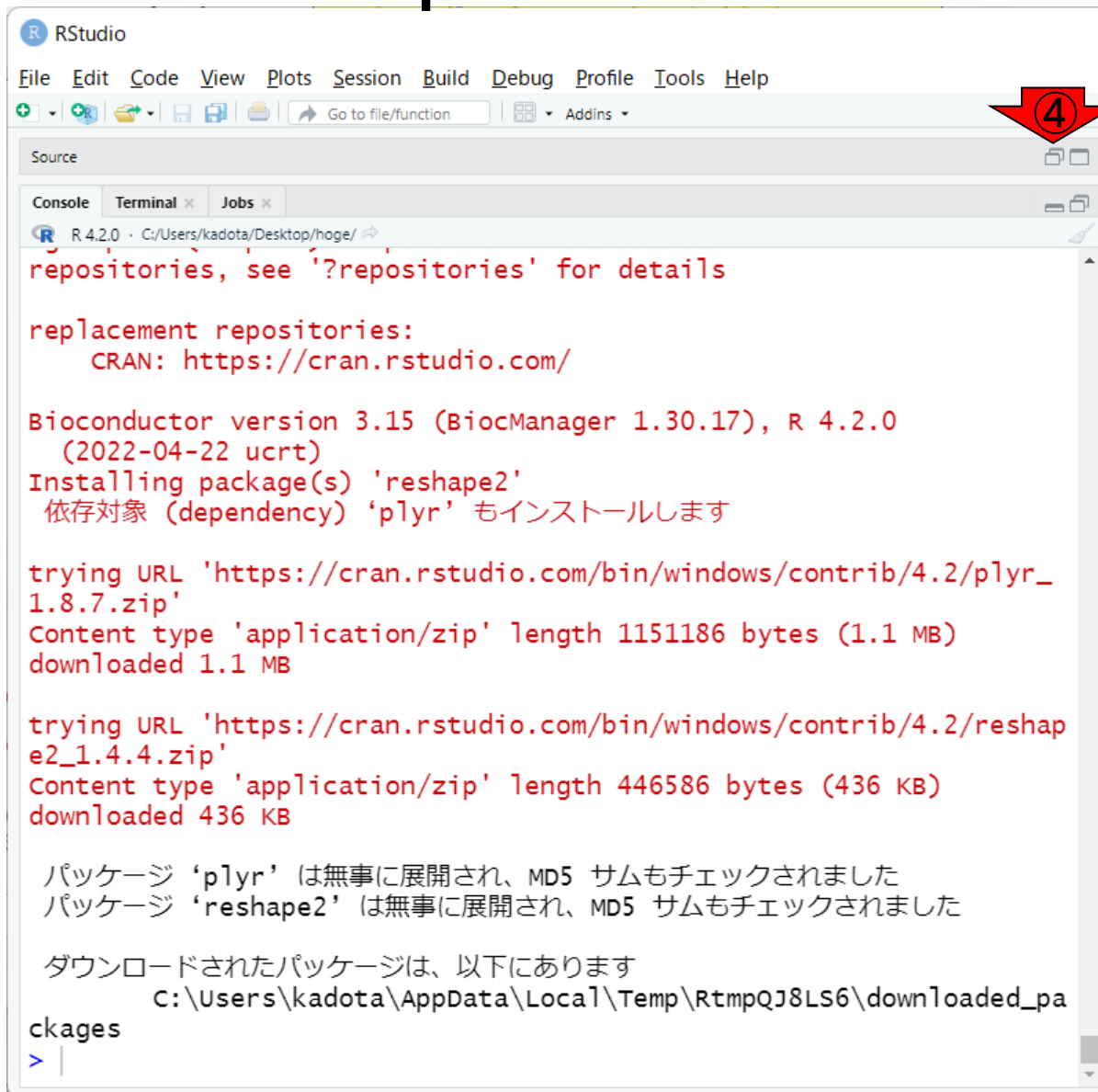
```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
downloaded 1.1 MB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/reshape2_1.4.4.zip'
Content type 'application/zip' length 446586 bytes (436 KB)
downloaded 436 KB

パッケージ 'plyr' は無事に展開され、MD5 サムもチェックされました
パッケージ 'reshape2' は無事に展開され、MD5 サムもチェックされました
```

W06: Tips (ウィンドウ) 4

- ①11行目が先頭行になるようにした状態。
- ②赤線あたりにマウスポインタをおいて上下に動かすことで、エディタ(or Console)画面サイズを調整できます。
- ③の部分左右に調整することもできます。
- ④一般的なソフトウェアのウィンドウサイズ調整同様に、④を押せばRエディタ画面の最小化、もう1回押せば元に戻ります。



```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source
Console Terminal Jobs
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
repositories, see '?repositories' for details

replacement repositories:
  CRAN: https://cran.rstudio.com/

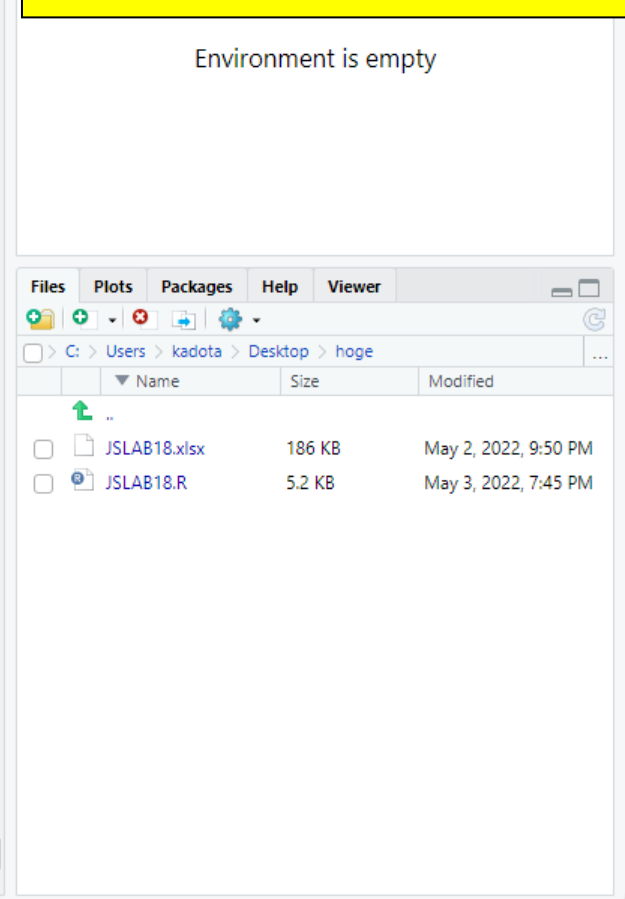
Bioconductor version 3.15 (BiocManager 1.30.17), R 4.2.0
(2022-04-22 ucrt)
Installing package(s) 'reshape2'
依存対象 (dependency) 'plyr' もインストールします

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/plyr_
1.8.7.zip'
Content type 'application/zip' length 1151186 bytes (1.1 MB)
downloaded 1.1 MB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/reshap
e2_1.4.4.zip'
Content type 'application/zip' length 446586 bytes (436 KB)
downloaded 436 KB

パッケージ 'plyr' は無事に展開され、MD5 サムもチェックされました
パッケージ 'reshape2' は無事に展開され、MD5 サムもチェックされました

ダウンロードされたパッケージは、以下にあります
C:\Users\kadota\AppData\Local\Temp\RtmpQJ8LS6\downloaded_pa
ckages
> |
```



Environment is empty

Files	Plots	Packages	Help	Viewer
C: > Users > kadota > Desktop > hoge				
Name	Size	Modified		
..				
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM		
JSLAB18.R	5.2 KB	May 3, 2022, 7:45 PM		

W06: Tips (ウィンドウ) 5

- ① 11行目が先頭行になるようにした状態。
- ② 赤線あたりにマウスポインタをおいて上下に動かすことで、エディタ(or Console)画面サイズを調整できます。
- ③ の部分を左右に調整することもできます。一般的なソフトウェアのウィンドウサイズ調整同様に、④を押せばREディタ画面の最小化、もう1回押せば元に戻ります。
- ⑤の部分についても同様です。以後は特に説明しないので、適宜自分で調整してください。

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

```
11 ##### w07:パッケージのロード(1回目) #####
12 library(TCC)
13 library(MBCluster.Seq)
14 library(openxlsx)
15 library(ggplot2)
16 library(ggsci)
17 library(reshape2)
18
19 ##### w07:パッケージのロード(2回目) #####
20 library(TCC)
21 library(MBCluster.Seq)
22 library(openxlsx)
23 library(ggplot2)
24 library(ggsci)
25 library(reshape2)
26
```

Console Terminal Jobs

R 4.2.0 · C:/Users/kadota/Desktop/hoge/

downloaded 1.1 MB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/reshape2_1.4.4.zip'

Content type 'application/zip' length 446586 bytes (436 KB)

downloaded 436 KB

パッケージ 'plyr' は無事に展開され、MD5 サムもチェックされました

パッケージ 'reshape2' は無事に展開され、MD5 サムもチェックされました

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	5.2 KB	May 3, 2022, 7:45 PM

Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W07: パッケージのロ... 1

インストールしたパッケージをlibrary関数を使って読み込んで、使える状態にする作業を行います。12-17行目を反転させて、①Run。

The screenshot shows the RStudio interface with the following content:

Source Editor:

```
11 ##### w07:パッケージのロード(1回目) #####
12 library(TCC)
13 library(MBCluster.Seq)
14 library(openxlsx)
15 library(ggplot2)
16 library(ggsci)
17 library(reshape2)
18
19 ##### w07:パッケージのロード(2回目) #####
20 library(TCC)
21 library(MBCluster.Seq)
22 library(openxlsx)
23 library(ggplot2)
24 library(ggsci)
25 library(reshape2)
26
```

Environment Panel: Environment is empty

Files Panel:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	5.2 KB	May 3, 2022, 7:45 PM

Console:

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
downloaded 1.1 MB

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/reshap
e2_1.4.4.zip'
Content type 'application/zip' length 446586 bytes (436 KB)
downloaded 436 KB

パッケージ 'plyr' は無事に展開され、MD5 サムもチェックされました
パッケージ 'reshape2' は無事に展開され、MD5 サムもチェックされました
```

W07: パッケージのロ...2

インストールしたパッケージをlibrary関数を使って読み込んで、使える状態にする作業を行います。12-17行目を反転させて、①Run。②Console画面はこんな感じになります。画面がざーっと流れていくさまを眺めながら、エラーメッセージっぽいものが出ていないかを確認するのが一般的です。慣れてくると③のあたりを眺めることで、「無事ロードに成功してるな」と判断できます。

The screenshot shows the RStudio interface with a script editor and a console window. The script editor contains two blocks of code for loading packages. The first block (lines 11-17) is highlighted in blue. The second block (lines 19-25) is also highlighted. A red arrow with the number '2' points to the first block. The console window shows the output of the first block, with a red arrow and the number '3' pointing to the output of the library(MBCluster.Seq) command.

```
11 ##### W07:パッケージのロード(1回目) #####
12 library(TCC)
13 library(MBCluster.Seq)
14 library(openxlsx)
15 library(ggplot2)
16 library(ggsci)
17 library(reshape2)
18
19 ##### W07:パッケージのロード(2回目) #####
20 library(TCC)
21 library(MBCluster.Seq)
22 library(openxlsx)
23 library(ggplot2)
24 library(ggsci)
25 library(reshape2)
```

Console Terminal Jobs

R 4.2.0 · C:/Users/kadota/Desktop/hoge/

以下のオブジェクトは "package:edgeR" からマスクされています:

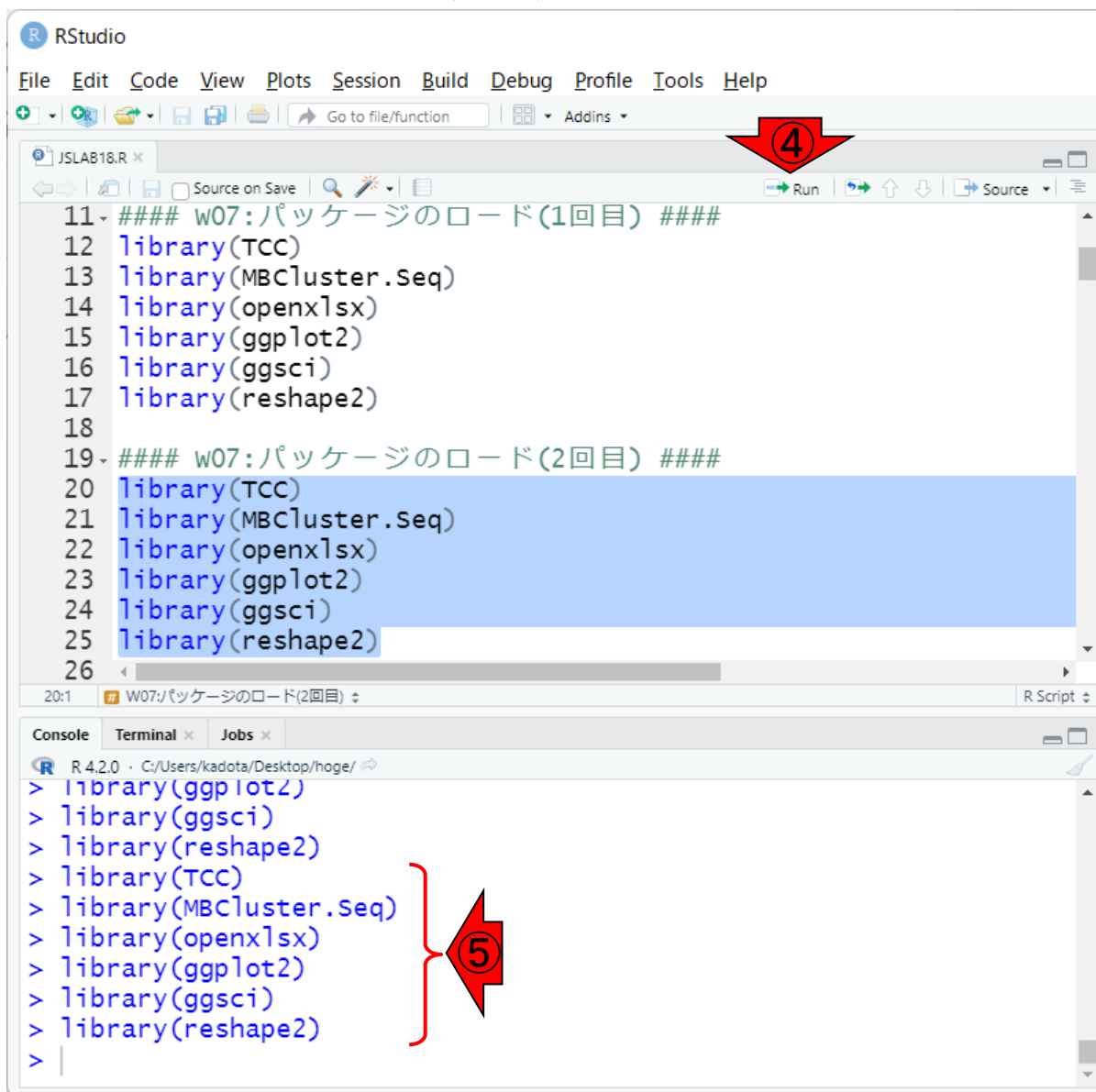
```
calcNormFactors
> library(MBCluster.Seq)
> library(openxlsx)
> library(ggplot2)
> library(ggsci)
> library(reshape2)
> |
```

The screenshot shows a file explorer window displaying the contents of a folder named 'hoge'. The files listed are JSLAB18.xlsx (186 KB, May 2, 20), JSLAB18.R (5.2 KB, May 3, 20), and .Rhistory (604 B, May 4, 20).

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 20
JSLAB18.R	5.2 KB	May 3, 20
.Rhistory	604 B	May 4, 20

W07: パッケージのロ...3

インストールしたパッケージをlibrary関数を使って読み込んで、使える状態にする作業を行います。12-17行目を反転させて、①Run。②Console画面はこんな感じになります。画面がざーっと流れていくさまを眺めながら、エラーメッセージっぽいものが出ていないかを確認するのが一般的です。慣れてくると③のあたりを眺めることで、「無事ロードに成功してるな」と判断できます。しかし最初のうちやロードするパッケージ数が多い場合は、なかなか判断が付きづらいので、同じ作業(つまり20-25行目)をもう一度実行することをお勧めします。④2回目の実行結果は⑤に相当しますが、1回目と2回目の区切り部分の特定は重要ではなく、実用上は**何のエラーメッセージも出ていないことの確認のみが重要です**。



The screenshot shows the RStudio interface. The script editor contains R code for loading packages. A red arrow labeled '4' points to the 'Run' button. The console shows the output of the code, with a red arrow labeled '5' pointing to the output of the second block of code.

```
#### w07:パッケージのロード(1回目) ####
library(TCC)
library(MBCluster.Seq)
library(openxlsx)
library(ggplot2)
library(ggsci)
library(reshape2)
#### w07:パッケージのロード(2回目) ####
library(TCC)
library(MBCluster.Seq)
library(openxlsx)
library(ggplot2)
library(ggsci)
library(reshape2)
```

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
> library(ggplot2)
> library(ggsci)
> library(reshape2)
> library(TCC)
> library(MBCluster.Seq)
> library(openxlsx)
> library(ggplot2)
> library(ggsci)
> library(reshape2)
```

Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- **W08: サンプルのクラスタリング**
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W08: サンプルのクラ...1

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for hierarchical clustering. Line 27 is highlighted with a red arrow and a circled '1'. The code includes comments in Japanese and uses functions like `read.xlsx`, `clustersample`, and `png`.
- Environment:** Shows 'Global Environment' and 'Environment is empty'.
- Files:** A file explorer showing the directory `C:\Users\kadota\Desktop\hoge` with files `JSLAB18.xlsx`, `JSLAB18.R`, and `.Rhistory`.
- Console:** Shows the execution of `library` functions for `TCC`, `MBCcluster.Seq`, `openxlsx`, `ggplot2`, `ggsci`, and `reshape2`.

W08: サンプルのクラ...2

①27行目が先頭行になるようにした状態。
28-36行目を反転させて、②Runした結果です。

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for reading data, performing hierarchical clustering, and plotting. A red arrow labeled '2' points to the Run button. The code is:

```
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clustersample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
```
- Console:** Shows the execution output of the code:

```
> library(ggplot2)
> library(ggsci)
> library(reshape2)
> ## データの読み込み ##
> data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
> ## サンプル間クラスタリング ##
+ out <- clustersample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
      1
>
```
- Environment:** Shows the current environment with 'data' (2949 obs. of 9 variables) and 'out' (List of 7).
- Files:** Shows the file explorer with a list of files:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 20
JSLAB18.R	5.2 KB	May 4, 20
.Rhistory	604 B	May 4, 20
fig7.png	4.5 KB	May 4, 20

W08: サンプルのクラ...3

①27行目が先頭行になるようにした状態。
28-36行目を反転させて、②Runした結果です。
③Console画面でエラーメッセージは出ておらず、④作業ディレクトリ上に、⑤で指定した出力ファイルが生成されていることがわかります。

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for clustering and plotting. Lines 28-36 are highlighted in blue, and a red arrow labeled ⑤ points to the `png` function call.
- Environment/Plots:** Shows the current data environment with 'data' (2949 obs. of 9 variables) and 'out' (List of 7).
- Files:** A file explorer window shows the current directory containing 'JSLAB18.xlsx', 'JSLAB18.R', '.Rhistory', and 'fig7.png'. A red arrow labeled ④ points to 'fig7.png'.
- Console:** Shows the execution of the R code. A red arrow labeled ③ points to the 'null device' output.

```
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clustersample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
```

```
> library(ggplot2)
> library(ggsci)
> library(reshape2)
> ## データの読み込み ##
> data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
> ## サンプル間クラスタリング ##
> out <- clustersample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
  1
```

W08: サンプルのクラ...4

⑥は、⑦read.xlsx関数で読み込んだ入力ファイルの中身を、⑧dataというオブジェクトに格納せよという指令です。⑨ Environmentタブに⑩同じのがみえます。

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for reading data, clustering, and plotting. Red arrows point to specific lines: ⑧ points to the `data` variable assignment, ⑦ points to the `read.xlsx` function, ⑥ points to the `data` object, and ⑩ points to the `out` object.
- Environment Panel:** Shows the `data` object with 2949 observations and 9 variables, and the `out` object as a list of 7 items.
- Console:** Shows the execution of the code, resulting in a `null device` and the number `1`.
- Files Panel:** Shows the local file system with files `JSLAB18.xlsx`, `JSLAB18.R`, `.Rhistory`, and `fig7.png`.

```
27- W08: サンプルのクラスタリング ####
28- データの読み込み ##
29- data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
30- ## サンプルのクラスタリング ##
31- out <- clustersample(data, dist.method="spearman",
32-                       hclust.method="average", unique.pattern=TRUE)
33- ## ファイルに保存(図7の作成) ##
34- png("fig7.png", pointsize=13, width=500, height=400)
35- plot(out)
36- dev.off()
37-
```

```
> library(ggplot2)
> library(ggsci)
> library(reshape2)
> ## データの読み込み ##
> data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
> ## サンプル間クラスタリング ##
+   out <- clustersample(data, dist.method="spearman",
+                         hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
      1
>
```


Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W09: Tips (オブジェ... 1

①の部分をクリックすると、入力ファイル中のヘッダ行情報や、②最初の数行分の情報が見られます。①をもう一度押すと元に戻ります。

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for reading an Excel file, performing hierarchical clustering, and saving a plot. A red arrow labeled ① points to the `read.xlsx` function call.
- Console:** Shows the execution of the code, including library loading and the output of `read.xlsx` and `clusterSample`.
- Environment:** Shows the `data` object as a data frame with 2949 observations and 9 variables. A red arrow labeled ② points to the first few rows of data.
- Files Panel:** Shows the file explorer with `JSLAB18.xlsx`, `JSLAB18.R`, `.Rhistory`, and `fig7.png`.

```
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
```

```
> library(ggplot2)
> library(ggsci)
> library(reshape2)
> ## データの読み込み ##
> data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
> ## サンプル間クラスタリング ##
+ out <- clusterSample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
      1
>
```

Environment	History	Connections	Tutorial
R	Global Environment	312 MIB	

Data	
data	2949 obs. of 9 varia...
\$ pH4.5_1h_1	num 21 8 657...
\$ pH4.5_1h_2	num 262 96 2...
\$ pH4.5_1h_3	num 275 124 ...
\$ pH4.5_24h_1	num 143 96 2...
\$ pH4.5_24h_2	num 323 318 ...
\$ pH4.5_24h_3	num 171 151 ...
\$ pH7_CCG_1	num 216 98 6...
\$ pH7_CCG_2	num 319 85 6...
\$ pH7_CCG_3	num 124 27 2...

Files	Plots	Packages	Help	Viewer
C:\Users\kadota\Desktop\hoge				
Name	Size	Modified		
..				
JSLAB18.xlsx	186 KB	May 2, 20		
JSLAB18.R	5.2 KB	May 4, 20		
.Rhistory	604 B	May 4, 20		
fig7.png	4.5 KB	May 5, 20		

W09: Tips (オブジェ...2)

①の部分をクリックすると、入力ファイル中のヘッダ行情報や、②最初の数行分の情報が見られます。①をもう一度押すと元に戻ります。③のあたりをクリックすると、エディタ画面に④dataタブが追加されて表形式で中身が見られるので便利です。⑤×を押すと元に戻ります。

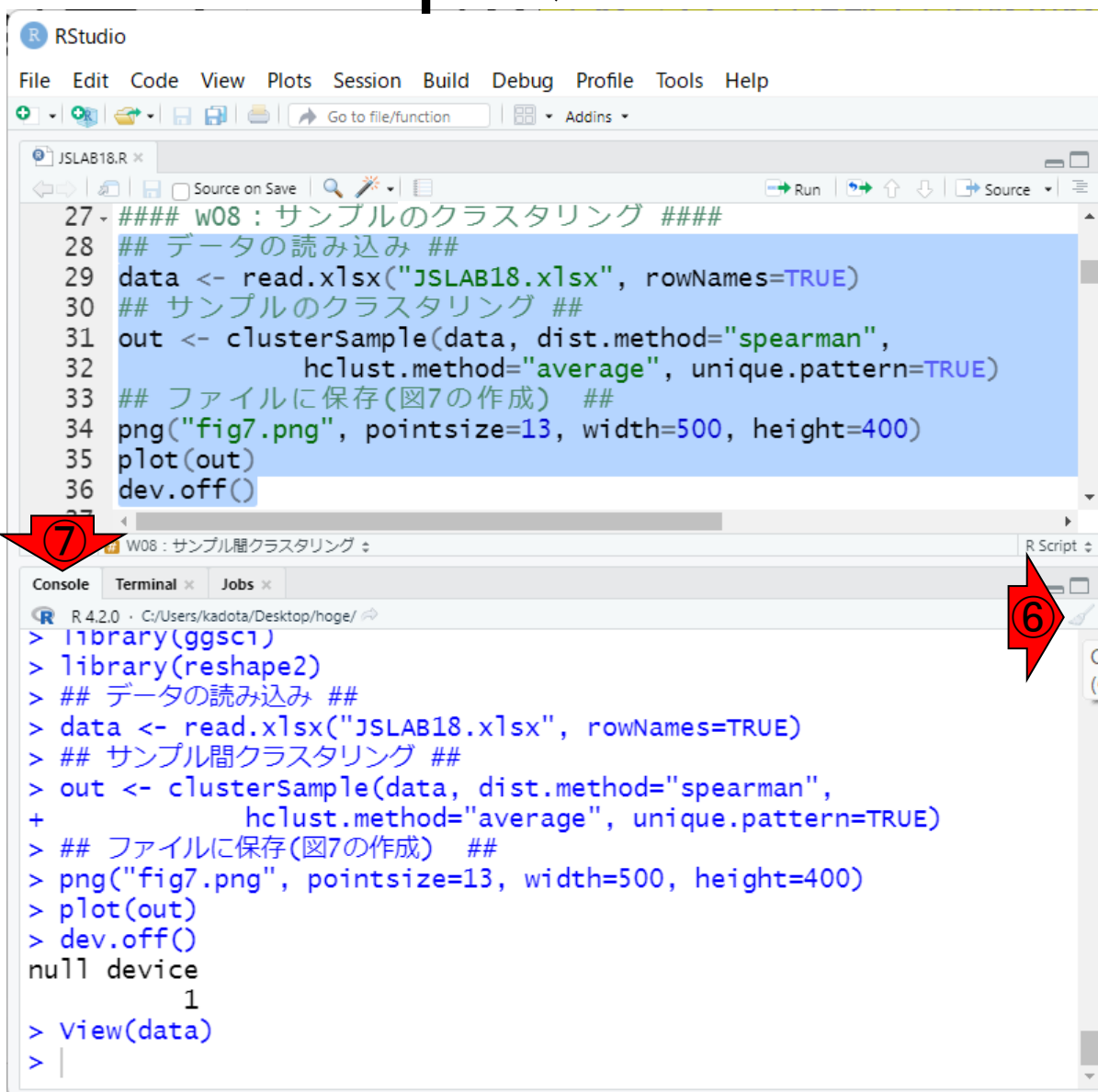
The screenshot shows the RStudio interface. At the top, the menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with various icons. A red arrow labeled '4' points to the 'View' menu, and another red arrow labeled '5' points to the 'data' tab in the environment pane. The environment pane displays a data table with columns: pH4.5_1h_1, pH4.5_1h_2, pH4.5_1h_3, pH4.5_24h_1, pH4.5_24h_2, pH4.5_24h_3, pH7_CCG_1, and pH7. The table contains 8 rows of data. Below the table, the console shows the following R code:

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
> library(ggsci)
> library(reshape2)
> ## データの読み込み ##
> data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
> ## サンプル間クラスタリング ##
> out <- clustersample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
  1
> view(data)
>
```

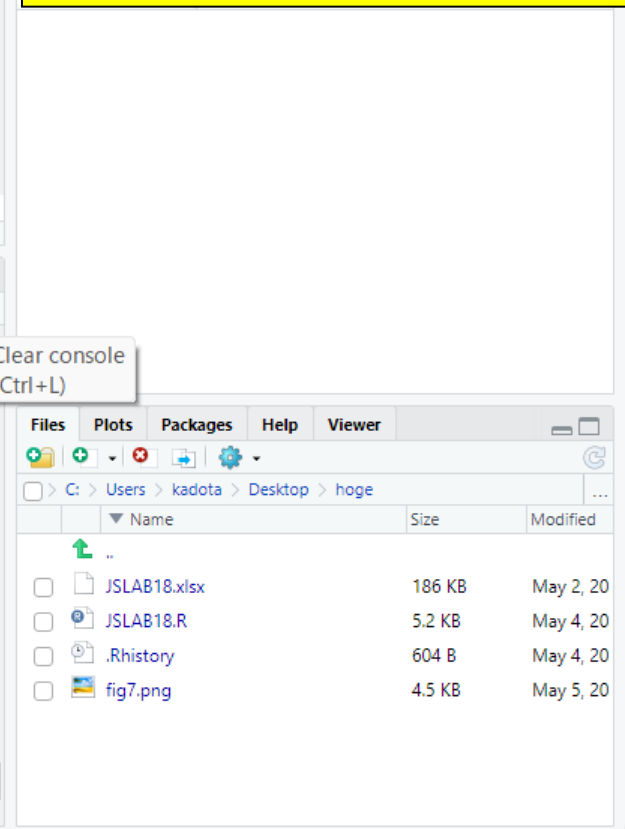
The screenshot shows the RStudio interface. The Data pane is visible, showing a table with two rows: 'data' (2949 obs. of 9 varia...) and 'out' (List of 7). A red arrow labeled '3' points to the 'data' row. Below the Data pane is the Files pane, showing a directory listing for 'C:\Users\kadota\Desktop\hoge'. The files listed are: JSLAB18.xlsx (186 KB, May 2, 20), JSLAB18.R (5.2 KB, May 4, 20), .Rhistory (604 B, May 4, 20), and fig7.png (4.5 KB, May 5, 20).

W09: Tips (オブジェ...3

①の部分をクリックすると、入力ファイル中のヘッダ行情報や、②最初の数行分の情報が見られます。①をもう一度押すと元に戻ります。③のあたりをクリックすると、エディタ画面に④dataタブが追加されて表形式で中身が見られるので便利です。⑤×を押すと元に戻ります。⑥を押すと、⑦Console画面をクリアできます。



```
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
```



```
> library(ggsci)
> library(reshape2)
> ## データの読み込み ##
> data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
> ## サンプル間クラスタリング ##
+ out <- clusterSample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
      1
> view(data)
>
```

W09: Tips (オブジェ...4

①の部分をクリックすると、入力ファイル中のヘッダ行情報や、②最初の数行分の情報が見られます。①をもう一度押すと元に戻ります。③のあたりをクリックすると、エディタ画面に④dataタブが追加されて表形式で中身が見られるので便利です。⑤×を押すと元に戻ります。⑥を押すと、⑦Console画面をクリアできます。実行結果。

The screenshot shows the RStudio interface with the following R code in the editor:

```
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
```

The console shows the command prompt: `R 4.2.0 · C:/Users/kadota/Desktop/hoge/ >`

The screenshot shows a file explorer window with the following contents:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 20
JSLAB18.R	5.2 KB	May 4, 20
.Rhistory	604 B	May 4, 20
fig7.png	4.5 KB	May 5, 20

W09: Tips (オブジェ...5)

次はオブジェクトの消去。慣れないうちは、何らかのエラーによって入力ファイルの読み込みがうまく行かなかった場合でも、以前作成した同名オブジェクト(例: data)が残っちゃっていて、解析自体は進められてしまうことがあります。ここはかなり重要な部分です。

```
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
```

Data	
data	2949 obs. of 9 varia...
out	List of 7



Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 20
JSLAB18.R	5.2 KB	May 4, 20
.Rhistory	604 B	May 4, 20
fig7.png	4.5 KB	May 5, 20

W09: Tips (オブジェ...6

例えば作業ディレクトリ上に存在しない
①hoge.xlsxを入力ファイルとして与えて
実行すると、②の部分で「ファイルが存
在しない」というエラーが正しく出ます。

The screenshot shows the RStudio interface. The script editor contains the following code:

```
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("hoge.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
```

A red arrow labeled ① points to the file path "hoge.xlsx" in line 29. The console shows the execution output:

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
> ## データの読み込み ##
> data <- read.xlsx("hoge.xlsx", rowNames=TRUE)
Error in read.xlsx.default("hoge.xlsx", rowNames = TRUE) :
  File does not exist.
> ## サンプル間クラスタリング ##
> out <- clusterSample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
      1
>
```

A red arrow labeled ② points to the error message "File does not exist." in the console. The Environment pane shows the data objects: data (2949 obs. of 9 variables) and out (List of 7). The Files pane shows the current directory (C:/Users/kadota/Desktop/hoge) containing files: JSLAB18.xlsx (186 KB), JSLAB18.R (5.2 KB), .Rhistory (604 B), and fig7.png (4.5 KB).

W09: Tips (オブジェ...7

例えば作業ディレクトリ上に存在しない
①hoge.xlsxを入力ファイルとして与えて
実行すると、②の部分で「ファイルが存
在しない」というエラーが正しく出ます。し
かし、以前作成した③dataオブジェクトが
存在するので、④以降はエラーなく実行
できてしまう。

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
JSLAB18.R* x
Source on Save
Run
Source
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("hoge.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
28:1 W08 : サンプル間クラスタリング R Script
Console Terminal x Jobs x
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
> ## データの読み込み ##
> data <- read.xlsx("hoge.xlsx", rowNames=TRUE)
Error in read.xlsx.default("hoge.xlsx", rowNames = TRUE) :
  File does not exist.
> ## サンプル間クラスタリング ##
> out <- clusterSample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
      1
>
```



Data

data	2949 obs. of 9 varia...
out	List of 7

Files Plots Packages Help Viewer

C:\Users\kadota\Desktop\hoge

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 20
JSLAB18.R	5.2 KB	May 4, 20
.Rhistory	604 B	May 4, 20
fig7.png	4.5 KB	May 5, 20

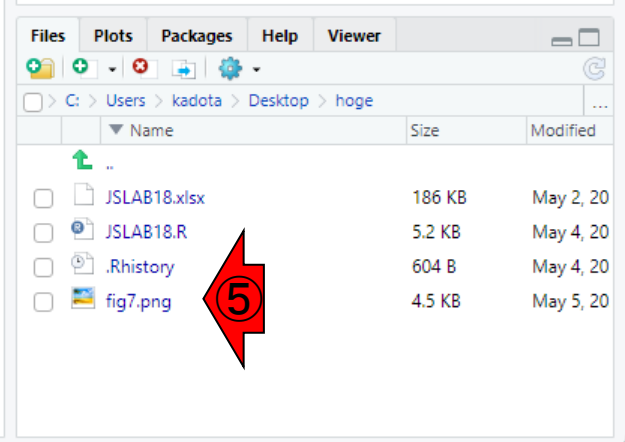
W09: Tips (オブジェ...8

例えば作業ディレクトリ上に存在しない
①hoge.xlsxを入力ファイルとして与えて
実行すると、②の部分で「ファイルが存
在しない」というエラーが正しく出ます。し
かし、以前作成した③dataオブジェクトが
存在するので、④以降はエラーなく実行
できてしまう。結果として、⑤作成された
クラスタリング結果は、本来やりたかった
①のデータではないものになってしまう
のです！アグリバイオの講義レポート課
題のありがちなミスです。

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
JSLAB18.R* x
Source on Save Run Source
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("hoge.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(① data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37

28:1 W08 : サンプル間クラスタリング R Script

Console Terminal Jobs
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
> ## データの読み込み ##
> data <- read.xlsx("hoge.xlsx", rowNames=TRUE)
Error in read.xlsx.default("hoge.xlsx", rowNames = TRUE) :
  File does not exist.
> ## サンプル間クラスタリング ##
> out <- clusterSample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
      1
>
```



W09: Tips (オブジェ...9

このような残念なミスを防げるのが、「大事な解析実行前の、①オブジェクトの全消去」。②Yes。

The screenshot shows the RStudio interface with a script editor, console, and environment pane. A red arrow labeled '1' points to the 'Global Environment' dropdown in the Environment pane. A red arrow labeled '2' points to the 'Yes' button in the 'Confirm Remove Objects' dialog box.

```
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("hoge.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
```

Console output:

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
> ## データの読み込み ##
> data <- read.xlsx("hoge.xlsx", row
Error in read.xlsx.default("hoge.xlsx",
  File does not exist.
> ## サンプル間クラスタリング ##
> out <- clusterSample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
> dev.off()
null device
      1
>
```

Environment pane:

Environment	History	Connections	Tutorial
R	Global Environment	693 MIB	
Data			
data	2949 obs. of 9 varia...		
out	List of 7		

Files pane:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 20
JSLAB18.R	5.2 KB	May 4, 20
.Rhistory	604 B	May 4, 20
fig7.png	4.5 KB	May 5, 20

W09: Tips (オブジェ... 10)

このような残念なミスを防げるのが、「大事な解析実行前の、①オブジェクトの全消去」。②Yes。③コンソール画面をクリアして、④作業ディレクトリ上に存在しないファイル名のままで、⑤Run。

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for hierarchical clustering. A red arrow labeled '4' points to the file name 'hoge.xlsx' in the `read.xlsx` function. Another red arrow labeled '5' points to the `Run` button.
- Environment Panel:** Shows 'Global Environment' and 'Environment is empty'.
- Console:** Shows the R prompt `>` and the current directory `C:/Users/kadota/Desktop/hoge/`. A red arrow labeled '3' points to the console area.
- Files Panel:** Shows a file explorer view of the `hoge` directory with the following table:

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 20
JSLAB18.R	5.2 KB	May 4, 20
.Rhistory	604 B	May 4, 20
fig7.png	4.5 KB	May 5, 20

W09: Tips (オブジェ... 11)

このような残念なミスを防げるのが、「大事な解析実行前の、①オブジェクトの全消去」。②Yes。③コンソール画面をクリアして、④作業ディレクトリ上に存在しないファイル名のままで、⑤Run。実行結果。沢山のエラーメッセージが出ていることがわかります。妥当ですね。

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
JSLAB18.R* x
Source on Save Run Source
27 ##### w08 : サンプルのクラスタリング #####
28 ## データの読み込み ##
29 data <- read.xlsx("hoge.xlsx", rowNames=TRUE)
30 ## サンプルのクラスタリング ##
31 out <- clusterSample(data, dist.method="spearman",
32                       hclust.method="average", unique.pattern=TRUE)
33 ## ファイルに保存(図7の作成) ##
34 png("fig7.png", pointsize=13, width=500, height=400)
35 plot(out)
36 dev.off()
37
28:1 W08 : サンプル階層クラスタリング R Script
Console Terminal x Jobs x
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
> ## サンプル階層クラスタリング ##
> out <- clusterSample(data, dist.method="spearman",
+                       hclust.method="average", unique.pattern=TRUE)
Error in rowSums(data) :
  'x' は少なくとも二次元の配列でなければなりません
> ## ファイルに保存(図7の作成) ##
> png("fig7.png", pointsize=13, width=500, height=400)
> plot(out)
Error in h(simpleError(msg, call)) :
  引数 'x' の評価中にエラーが起きました (関数 'plot' に対するメソッドの選択
時): object 'out' not found
> dev.off()
null device
      1
> |
```

Environment is empty

Files Plots Packages Help Viewer

C:\Users\kadota\Desktop\hoge

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 20
JSLAB18.R	5.2 KB	May 4, 20
.Rhistory	604 B	May 4, 20
fig7.png	289 B	May 5, 20



Contents

- W01: RStudio起動
- W02: ファイルの保存
- W03: 作業ディレクトリの変更
- W04: エディタの起動
- W05: パッケージのインストール
- W06: Tips (ウィンドウサイズの調整)
- W07: パッケージのロード
- W08: サンプルのクラスタリング
- W09: Tips (オブジェクトの確認や消去、コンソール画面のクリア)
- W10: MBCdegの実行

W10: MBCdegの実行1

W10の反転部分実行結果。①で入力ファイルを再読み込みし、②で4-9列目のデータを抽出している。

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for reading data and processing it. Red arrows labeled ① and ② point to the `read.xlsx` and `data` assignment lines, respectively.
- Environment Pane:** Shows the objects created in the Global Environment:

Object	Description
data	2949 obs. of 6 variables
data_all	2949 obs. of 9 variables
- Console:** Shows the execution output of the code, including the file path and the objects created.
- Files Pane:** Shows the project files: `JSLAB18.xlsx` (186 KB) and `JSLAB18.R` (4.2 KB).

W10: MBCdegの実行2

W10の反転部分実行結果。①で入力ファイルを再読み込みし、②で4-9列目のデータを抽出している。③data_allオブジェクトの概要を見てもわかるように、④の列が4-9列目に相当。

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for loading and processing data. Red arrows point to specific lines: ② points to the data extraction line, and ④ points to the normalization line.
- Environment:** Shows the 'Data' environment with a table of variables and their dimensions.
- Console:** Shows the output of the executed code.
- Files Panel:** Shows the file structure in the current directory.

```
38 - ##### w10 : MBCdeg2の実行 #####
39 ## データの読み込み ##
40 data_all <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
41
42 ## pH4.5_24h群とpH7_CCG群の抽出 ##
43 data <- data_all[, 4:9]
44
45 ## 群ラベル情報の作成とTCC正規化 ##
46 data.c1 <- c(1,1,1,2,2,2)
47 tcc <- new('TCC',data,data.c1)
48 tcc <- calcNormFactors(tcc,norm.method="tmm",test.method="")
49
```

Variable	Dimensions
data	2949 obs. of 6 variables
data_all	2949 obs. of 9 variables
\$ pH4.5_1h_1	num 21 8 6579 16 0 ...
\$ pH4.5_1h_2	num 262 96 20948 33...
\$ pH4.5_1h_3	num 275 124 28807 3...
\$ pH4.5_24h_1	num 143 96 22790 16...
\$ pH4.5_24h_2	num 323 318 57552 4...
\$ pH4.5_24h_3	num 171 151 28927 2...
\$ pH7_CCG_1	num 216 98 65311 32...
\$ pH7_CCG_2	num 319 85 67686 40...
\$ pH7_CCG_3	num 124 27 23262 11...

```
> ## データの読み込み ##
> data_all <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
>
> ## pH4.5_24h群とpH7_CCG群の抽出 ##
> data <- data_all[, 4:9]
>
```

Name	Size	Modified
..		
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.2 KB	May 8, 2022, 1:05 PM

W10: MBCdegの実行3

W10の反転部分実行結果。①で入力ファイルを再読み込みし、②で4-9列目のデータを抽出している。③data_allオブジェクトの概要を見てもわかるように、④の列が4-9列目に相当。⑤dataオブジェクトの概要でもわかるように、解析したい列を抽出できていることがわかります。

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for loading and processing data. Lines 38-49 are visible. A blue highlight covers lines 39-41, and another blue highlight covers lines 42-43. Red arrows labeled ⑤ point to these highlighted sections.
- Console:** Shows the execution output of the code, with corresponding blue highlights and a red arrow labeled ④ pointing to the output of the data extraction step.
- Environment:** Shows two data objects: 'data' (2949 obs. of 6 variables) and 'data_all' (2949 obs. of 9 variables). A red bracket labeled ④ groups the variables in 'data_all' that correspond to the columns 4-9 of the original data.
- Files Panel:** Shows the current working directory as 'C:\Users\kadota\Desktop\hoge'.

```
38 - ##### w10 : MBCdeg2の実行 #####
39 ## データの読み込み ##
40 data_all <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
41
42 ## pH4.5_24h群とpH7_CCG群の抽出 ##
43 data <- data_all[, 4:9]
44
45 ## 群ラベル情報の作成とTCC正規化 ##
46 data.cl <- c(1,1,1,2,2,2)
47 tcc <- new('TCC',data,data.cl)
48 tcc <- calcNormFactors(tcc,norm.method="tmm",test.met
49
```

```
> ## データの読み込み ##
> data_all <- read.xlsx("JSLAB18.xlsx", rowNames=TRUE)
>
> ## pH4.5_24h群とpH7_CCG群の抽出 ##
> data <- data_all[, 4:9]
>
```

Object	Obs	Vars
data	2949	6
data_all	2949	9

Variable	Unit	Value 1	Value 2	Value 3	Value 4
\$ pH4.5_24h_1	num	143	96	22790	16...
\$ pH4.5_24h_2	num	323	318	57552	4...
\$ pH4.5_24h_3	num	171	151	28927	2...
\$ pH7_CCG_1	num	216	98	65311	32...
\$ pH7_CCG_2	num	319	85	67686	40...
\$ pH7_CCG_3	num	124	27	23262	11...

W10: MBCdegの実行4

①TCCパッケージのURIである頑健なDEGES正規化法の実行部分。ややこしいので詳細は割愛しますが、②で指定しているのは、③解析対象データのどの列がどの群に属するかを指定する群ラベル情報に相当するものです。④でもわかりますが、最初の3列がG1群、残りの3列分がG2群という意味です。

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for TCC normalization and cluster analysis. Red arrows point to specific lines: ① points to the comment, ② to the data vector, ③ to the `calcNormFactors` function call, and ④ to the `size.factors` calculation.
- Console:** Shows the execution output, including the TCC normalization process and the resulting object structure.
- Environment:** Lists the objects created, such as `data`, `data_all`, `data.c1`, `ef.libsize...`, `norm.fact...`, `size.fact...`, and `tcc`.

```
45 ## 群ラベル情報の作成とTCC正規化 ##
46 data.c1 <- c(1,1,1,2,2,2)
47 tcc <- new('TCC',data,data.c1)
48 tcc <- calcNormFactors(tcc,norm.method="tmm",test.met
49 iteration=3, FDR=0.1,floorPDEG=0.05)
50 norm.factors <- tcc$norm.factors
51 ef.libsizes <- colSums(data)*norm.factors
52 size.factors <- ef.libsizes/mean(ef.libsizes)
53
54 ## MBCcluster.seq実行の共通部分 ##
55 set.seed(2022)
56
```

```
> data.c1 <- c(1,1,1,2,2,2)
> tcc <- new('TCC',data,data.c1)
> tcc <- calcNormFactors(tcc,norm.method="tmm",test.method
+ iteration=3, FDR=0.1,floorPDEG=0.05)
TCC::INFO: Calculating normalization factors using DEGES
TCC::INFO: (iDEGES pipeline : tmm - [ edger - tmm ] x 3 )
TCC::INFO: Done.
> norm.factors <- tcc$norm.factors
> ef.libsizes <- colSums(data)*norm.factors
> size.factors <- ef.libsizes/mean(ef.libsizes)
>
```

Object	Variables
data	2949 obs. of 6 variables
data_all	2949 obs. of 9 variables
data.c1	num [1:6] 1 1 1 2 2 2
ef.libsize...	Named num [1:6] 997916 291...
norm.fact...	Named num [1:6] 1.063 1.10...
size.fact...	Named num [1:6] 0.637 1.86...
tcc	<Object containing active ...

W10: MBCdegの実行5

反転部分実行結果。MBCcluster.Seqは内部的に乱数を発生させるK-meansを使用するが、試行ごとに結果が異なるとややこしいので、①で乱数を発生させるタネ番号を固定させている。ここでは2022としているが、別に1000でも73とかでもなんでもよい。

```
45 ## 群ラベル情報の作成とTCC正規化 ##
46 data.c1 <- c(1,1,1,2,2,2)
47 tcc <- new('TCC',data,data.c1)
48 tcc <- calcNormFactors(tcc,norm.method="tmm",test.method="edgeR",
49                       iteration=3, FDR=0.1,floorPDEG=0.001)
50 norm.factors <- tcc$norm.factors
51 ef.libsizes <- colSums(data)*norm.factors
52 size.factors <- ef.libsizes/mean(ef.libsizes)
53
54 ## MBCcluster.Seq実行の共通部分 ##
55 set.seed(2022)
56 hoge <- RNASeq.Data(Count = data, Treatment = data.c1,
57                    Normalizer= log(size.factors))
58
```

Console output:

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
TCC::INFO: calculating normalization factors using DEGES
TCC::INFO: (iDEGES pipeline : tmm - [ edgeR - tmm ] X 3 )
TCC::INFO: Done.
> norm.factors <- tcc$norm.factors
> ef.libsizes <- colSums(data)*norm.factors
> size.factors <- ef.libsizes/mean(ef.libsizes)
> set.seed(2022)
> hoge <- RNASeq.Data(Count = data, Treatment = data.c1,
+                    Normalizer= log(size.factors))
>
```

Environment pane:

Data	
data	2949 obs. of 6 variables
\$ pH4.5_24h_1:	num 143 96 22790 16...
\$ pH4.5_24h_2:	num 323 318 57552 4...
\$ pH4.5_24h_3:	num 171 151 28927 2...
\$ pH7_CCG_1 :	num 216 98 65311 32...
\$ pH7_CCG_2 :	num 319 85 67686 40...
\$ pH7_CCG_3 :	num 124 27 23262 11...
data_all	2949 obs. of 9 variables
hoge	List of 7
Values	
data.c1	num [1:6] 1 1 1 2 2 2
ef.libsiz...	Named num [1:6] 997916 291...
norm.fact...	Named num [1:6] 1.063 1.10...
size.fact...	Named num [1:6] 0.637 1.86...
tcc	<Object containing active ...

W10: MBCdegの実行6

56-57行目では、①dataオブジェクト、②群ラベル情報、そして③TCC正規化係数をちょっと変換したもの(サイズファクター)の、④logをとったものを入力として与えてRNASeq.Data関数を実行し、その結果をhogeオブジェクトに格納している。

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for creating data, calculating TCC normalization factors, and running MBCcluster.Seq. Red arrows point to specific parts of the code: ① points to the `data` object, ② points to the `data.c1` vector, ③ points to the `size.factors` calculation, and ④ points to the `log(size.factors)` normalization step in the `RNASeq.Data` function call.
- Environment Pane:** Shows the objects created in the global environment. The `data` object is a matrix with 2949 observations and 6 variables. The `hoge` object is a list of 7 elements. The `Values` section shows the structure of the objects: `data.c1` is a numeric vector [1, 1, 1, 2, 2, 2], `ef.libsize...` is a named numeric vector, `norm.fact...` is a named numeric vector, `size.fact...` is a named numeric vector, and `tcc` is an object containing active data.
- Console:** Shows the execution output, including the TCC pipeline information and the successful execution of the normalization steps and the `RNASeq.Data` function call.

W10: MBCdegの実行7

クラスター数Kを3として実行する部分。
①60行目でKの値、61行目で実行結果ファイル名をMBCdeg_K3.xlsxとするように指定している。赤枠部分を異なるKの値のときでも使いまわせるようにして、最終的にK = 3の結果情報を②72行目でcls_k3という名前で格納して(上書きされないように)退避させているのがポイント。

```
59 ## MBCcluster.Seq(K=3)の実行部分 ##
60 K <- 3 #クラスター数を指定
61 out_f <- "MBCdeg_K3.xlsx" #出力ファイル名を指定
62 c0 <- KmeansPlus.RNASeq(data=hoge, nK=K,
63 model="nbinom", print.steps=F)
64 cls <- Cluster.RNASeq(data=hoge, model="nbinom",
65 centers=c0$centers, method="EM")
66 PP <- cls$probability
67 colnames(PP) <- paste("c", 1:K, sep="")
68 tmp <- cbind(rownames(data), data, PP, cls$cluster)
69 write.xlsx(x=tmp, file=out_f)
70 cls$centers
71 table(cls$cluster)
72 cls_k3 <- cls
73
```

Console output:

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
[5,] 0.1601454 -0.1601454
attr(,"class")
[1] "logFC"
> table(cls$cluster)

 1   2   3
1303 222 1424
> cls_k3 <- cls
>
```

Name	Size	Modified
..		
MBCdeg_K3.xlsx	231.3 KB	May 14, 2022, 8:44 PM
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.2 KB	May 8, 2022, 1:05 PM

W10: MBCdegの実行8

クラスター数Kを3として実行する部分。
①60行目でKの値、61行目で実行結果ファイル名をMBCdeg_K3.xlsxとするように指定している。赤枠部分を異なるKの値のときでも使いまわせるようにして、最終的にK = 3の結果情報を②72行目でcls_k3という名前で格納して(上書きされないように)退避させているのがポイント。
③70行目のcls\$centersは、クラスター中心のlog比の情報になります。④がどのクラスター番号のものが何個あったかという情報になります。これらの表示されている情報を見て、原稿中の図8に書き込んでいます。

```
## MBCcluster.Seq(K=3)の実行部分 ##
K <- 3 #クラスター数を指定
out_f <- "MBCdeg_K3.xlsx" #出力ファイル名を指定
c0 <- kmeansPlus.RNASeq(data=hoge, nK=K,
  model="nbinom", print.steps=F)
cls <- Cluster.RNASeq(data=hoge, model="nbinom",
  centers=c0$centers, method="EM")
PP <- cls$probability
colnames(PP) <- paste("c", 1:K, sep="")
tmp <- cbind(rownames(data), data, PP, cls$cluster)
cls_k3 <- cls
```

Console Terminal

```
> cls$centers
      1      2
[1,] -0.2996681  0.2996681
[2,]  1.7270041 -1.7270041
[3,]  0.1801454 -0.1801454
attr(,"class")
[1] "logFC"
> table(cls$cluster)
 1  2  3
1303 222 1424
> cls_k3 <- cls
```

Files Explorer

Name	Size	Modified
..		
MBCdeg_K3.xlsx	231.3 KB	May 14, 2022, 8:44 PM
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.2 KB	May 8, 2022, 1:05 PM

W10: MBCdegの実行9

クラスター数Kを3として実行する部分。
 ①60行目でKの値、61行目で実行結果ファイル名をMBCdeg_K3.xlsxとるように指定している。赤枠部分を異なるKの値のときでも使いまわせるようにして、最終的にK = 3の結果情報を②72行目でcls_k3という名前で格納して(上書きされないように)退避させているのがポイント。
 ③70行目のcls\$centersは、クラスター中心のlog比の情報になります。④がどのクラスター番号のものが何個あったかという情報になります。これらの表示されている情報を見て、原稿中の図8に書き込んでいます。今はK = 3の結果ですので、赤字で書いてあるものに相当します。

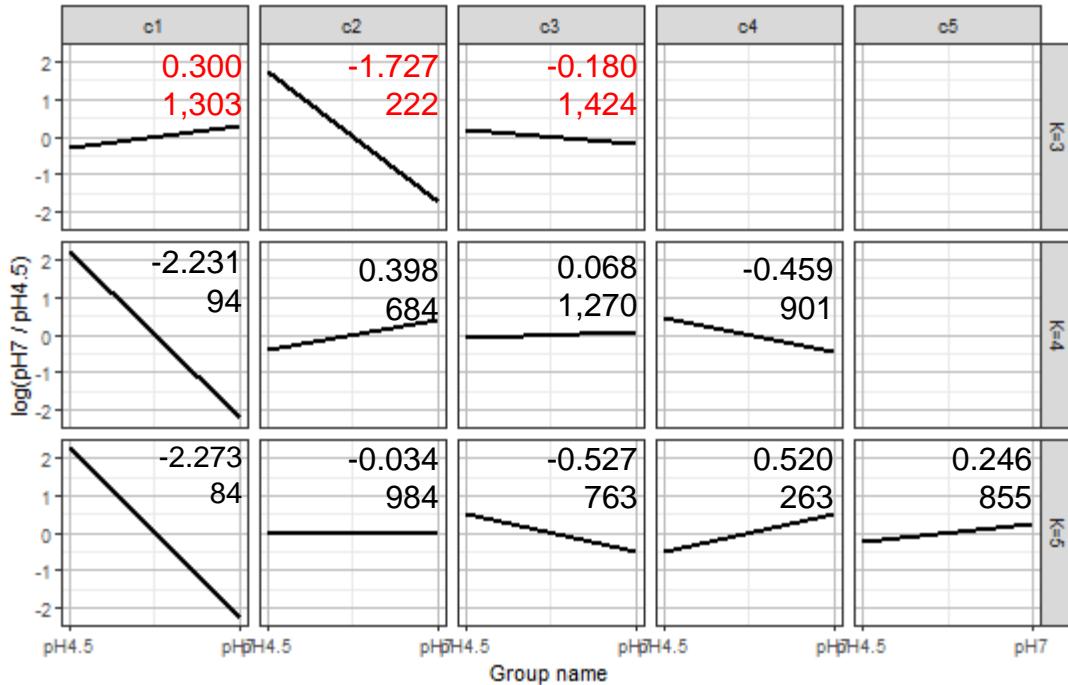


図8. MBCdeg実行結果。スクリプト実行結果として自動生成されるfig8.pngをベースとして、右上部分の数値情報を後から追加している。

W10: MBCdegの実行10

K = 4実行部分。K = 3のときと同様、①と②以外の、赤枠部分は同じものを使いまわしています。

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for `MBCcluster.Seq(K=4)`. A red box highlights lines 77-87, with a red arrow labeled '1' pointing to the top and another labeled '2' pointing to the bottom.
- Environment:** Lists objects in the Global Environment, including `c0`, `cls`, `cls_k3`, `cls_k4`, `data`, `data_all`, `hoge`, `PP`, and `tmp`.
- Files:** Shows a file explorer view of the `hoge` directory, listing `MBCdeg_K4.xlsx`, `MBCdeg_K3.xlsx`, `JSLAB18.xlsx`, and `JSLAB18.R`.
- Console:** Shows the output of `table(cls$cluster)`, displaying a 4x4 matrix of counts.

```
74 ## MBCcluster.Seq(K=4)の実行部分 ##
75 K <- 4 #クラスター数を指定
76 out_f <- "MBCdeg_k4.xlsx" #出力ファイル名を指定
77 c0 <- kmeansPlus.RNASeq(data=hoge, nK=K,
78 model="nbinom", print.steps=F)
79 cls <- Cluster.RNASeq(data=hoge, model="nbinom",
80 centers=c0$centers, method="EM")
81 PP <- cls$probability
82 colnames(PP) <- paste("c", 1:K, sep="")
83 tmp <- cbind(rownames(data), data, PP, cls$cluster)
84 write.xlsx(x=tmp, file=out_f)
85 cls$centers
86 table(cls$cluster)
87 cls_k4 <- cls
88
```

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
[4,] 0.45920365 -0.45920365
attr(,"class")
[1] "logFC"
> table(cls$cluster)

 1  2  3  4
94 684 1270 901
> cls_k4 <- cls
>
```

Name	Size	Modified
..		
MBCdeg_K4.xlsx	272.8 KB	May 14, 2022, 9:08 PM
MBCdeg_K3.xlsx	231.3 KB	May 14, 2022, 8:44 PM
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.2 KB	May 8, 2022, 1:05 PM

W10: MBCdegの実行11

K = 5実行部分。K = 3のときと同様、①と②以外の、赤枠部分は同じものを使いまわしています。

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for `MBCcluster.Seq(K=5)`. Lines 90-92 are highlighted in blue and marked with a red arrow labeled ①. Lines 93-102 are enclosed in a red box and marked with a red arrow labeled ②.
- Environment:** Lists objects in the Global Environment, including `c0`, `cls`, `cls_k3`, `cls_k4`, `cls_k5`, `data`, `data_all`, `hoge`, `PP`, and `tmp`.
- Files:** Shows a file explorer view of the `hoge` directory, listing files like `MBCdeg_K5.xlsx`, `MBCdeg_K4.xlsx`, `MBCdeg_K3.xlsx`, `JSLAB18.xlsx`, and `JSLAB18.R`.
- Console:** Shows the output of the `table(cls$cluster)` command, displaying a matrix of counts for each cluster.

```
## MBCcluster.Seq(K=5)の実行部分 ##
K <- 5 #クラスター数を指定
out_f <- "MBCdeg_k5.xlsx" #出力ファイル名を指定
c0 <- kmeansPlus.RNASeq(data=hoge, nK=K,
  model="nbinom", print.steps=F)
cls <- Cluster.RNASeq(data=hoge, model="nbinom",
  centers=c0$centers, method="EM")
PP <- cls$probability
colnames(PP) <- paste("c", 1:K, sep="")
tmp <- cbind(rownames(data), data, PP, cls$cluster)
write.xlsx(x=tmp, file=out_f)
cls$centers
table(cls$cluster)
cls_k5 <- cls
```

```
R 4.2.0 · C:/Users/kadota/Desktop/hoge/
[5,] -0.24576894  0.24576894
attr(,"class")
[1] "logFC"
> table(cls$cluster)

 1  2  3  4  5
84 984 763 263 855
> cls_k5 <- cls
> |
```

Name	Size	Modified
..		
MBCdeg_K5.xlsx	315 KB	May 14, 2022, 9:11 PM
MBCdeg_K4.xlsx	272.8 KB	May 14, 2022, 9:08 PM
MBCdeg_K3.xlsx	231.3 KB	May 14, 2022, 8:44 PM
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.2 KB	May 8, 2022, 1:05 PM

W10: MBCdegの実行12

図8を作成するための基礎情報をまとめている部分。106-110行目の左辺で見えているc1, c2, ..., c5はクラスター番号情報。K = 3の場合はc1からc3まで、K = 4の場合はc1からc4まで、そしてK = 5の場合はc1からc5までである。したがって、K = 3の場合は、c4とc5に相当する部分はNot Availableを意味するNAという情報で埋めているようなイメージで捉えるとよいです。①は気にしなくてかまいません。

The screenshot shows the RStudio interface with the following content:

```
104 ## ファイルに保存(図8の作成) ##
105 dat <- data.frame(
106   c1 = c(c1s_k3$centers[1,], c1s_k4$centers[1,], c1s_k5$centers[1,]),
107   c2 = c(c1s_k3$centers[2,], c1s_k4$centers[2,], c1s_k5$centers[2,]),
108   c3 = c(c1s_k3$centers[3,], c1s_k4$centers[3,], c1s_k5$centers[3,]),
109   c4 = c(rep(NA, 2), c1s_k4$centers[4,], c1s_k5$centers[4,]),
110   c5 = c(rep(NA, 2), rep(NA, 2), c1s_k5$centers[5,])
111 )
112 melt(dat)
113 matome <- data.frame(melt(dat), k=rep(c(rep('K=3', 2), rep('K=4', 2), rep('K=5', 2)), 5), u=rep(1:2, 15))
114
115
```

Console output:

```
23   c4 -0.52035924
24   c4  0.52035924
25   c5          NA
26   c5          NA
27   c5          NA
28   c5          NA
29   c5 -0.24576894
30   c5  0.24576894
> matome <- data.frame(melt(dat), k=rep(c(rep('K=3', 2), rep('K=4', 2), rep('K=5', 2)), 5), u=rep(1:2, 15))
No id variables; using all as measure variables
>
```

Environment pane:

data_all	2949 obs. of 9 variables
hoge	List of 7
matome	30 obs. of 4 variables
PP	num [1:2949, 1:5] 3.73e-8...
tmp	2949 obs. of 13 variables

Files pane:

Name	Size	Modified
..		
MBCdeg_K5.xlsx	315 KB	May 14, 2022, 9:11 PM
MBCdeg_K4.xlsx	272.8 KB	May 14, 2022, 9:08 PM
MBCdeg_K3.xlsx	231.3 KB	May 14, 2022, 8:44 PM
JSLAB18.xlsx	186 KB	May 2, 2022, 9:50 PM
JSLAB18.R	4.2 KB	May 8, 2022, 1:05 PM

A red arrow with the number 1 points to the console output line "No id variables; using all as measure variables".

W10: MBCdegの実行13

最後の部分を実行。①ggplot関数を用いています。出力ファイルである②fig8.pngの中身は…

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains R code for creating a ggplot. A red arrow labeled '1' points to the `ggplot` function call on line 115.
- Environment:** Lists objects in the Global Environment, including `data` (2949 obs. of 6 variables) and `matome` (30 obs. of 4 variables).
- Files Panel:** Shows the file explorer for the directory `C:\Users\kadota\Desktop\hoge`. A red arrow labeled '2' points to the file `fig8.png`, which is 5.7 KB and was modified on May 14, 2022, at 9:26 PM.
- Console:** Shows the output of the executed code, including the command `png("fig8.png", pointsize=13, width=550, height=350)` and the resulting `null device`.

W10: MBCdegの実行14

最後の部分を実行。①ggplot関数を用いています。出力ファイルである②fig8.pngの中身は、こんな感じです。

The screenshot shows the RStudio interface with the following code in the editor:

```
114
115 g <- ggplot(data = matome)
116 g <- g + geom_line(aes(x = u, y = value), size
117 g <- g + theme_bw()
118 g <- g + theme(panel.grid.major = element_line(
119               aspect.ratio = 1)
120 g <- g + labs(x='Group name', y='log(pH7 / pH4.5)')
121 g <- g + scale_x_continuous(breaks=1:2, labels
122 g <- g + facet_grid(k ~ variable)
123
124 png("fig8.png", pointsize=13, width=550, height=350)
125 plot(g)
126 dev.off()
127
```

The console shows the execution of the final lines of code:

```
5', 'pH7'))
> g <- g + facet_grid(k ~ variable)
>
> png("fig8.png", pointsize=13, width=550, height=350)
> plot(g)
> dev.off()
null device
      1
> |
```

The plot, titled "fig8.png", is a 3x5 grid of line plots. The y-axis is labeled "log(pH7 / pH4.5)" and ranges from -2 to 2. The x-axis is labeled "Group name" and has two major ticks: "pH4.5" and "pH7". The grid is faceted by "variable" (columns: c1, c2, c3, c4, c5) and "k" (rows: K=3, K=4, K=5). The plots show the relationship between the log ratio and the group name for each combination of variable and k.

Variable	k	Group name	log(pH7 / pH4.5)
c1	K=3	pH4.5	-0.2
		pH7	0.2
	K=4	pH4.5	2.2
		pH7	-2.2
	K=5	pH4.5	2.2
		pH7	-2.2
c2	K=3	pH4.5	1.8
		pH7	-1.8
	K=4	pH4.5	-0.2
		pH7	0.2
	K=5	pH4.5	0.2
		pH7	-0.2
c3	K=3	pH4.5	0.2
		pH7	-0.2
	K=4	pH4.5	0.2
		pH7	0.2
	K=5	pH4.5	0.5
		pH7	-0.5
c4	K=3	pH4.5	0.2
		pH7	-0.2
	K=4	pH4.5	0.5
		pH7	-0.5
	K=5	pH4.5	-0.5
		pH7	0.5
c5	K=3	pH4.5	-0.2
		pH7	0.2
	K=4	pH4.5	-0.2
		pH7	0.2
	K=5	pH4.5	-0.2
		pH7	0.2

The file explorer at the bottom right shows the following files:

- MBCdeg_K4.xlsx (272.8 KB, May 14, 2022, 9:08 PM)
- MBCdeg_K3.xlsx (231.3 KB, May 14, 2022, 8:44 PM)
- JSLAB18.xlsx (186 KB, May 2, 2022, 9:50 PM)
- JSLAB18.R (4.2 KB, May 8, 2022, 1:05 PM)
- fig8.png (5.7 KB, May 14, 2022, 9:26 PM)