

平成25年度ワークショップ

# やってみようよ！画像、統計、ゲノム・トランスクリプトーム解析ハンズオン実習

平成25年6月6日（木）奈良先端科学  
よ！画像、統計、ゲノム・トランスクリ  
ローバルトップ教育推進プログラム）  
だき、画像解析、生物統計、ゲノム  
皆様奮ってご参加下さい。

※ワークショップ・懇親会参加希望の  
※本実習においては、参加前に持込  
お願いします。対応可能なOS、及び

## <ImageJについて>

ImageJとそのプラグイン、解析用テ  
※コンピューターのスペックは特に特

## <Rについて>

Windows (64 bit, 4GB以上) 、R 2  
Mac OS X (10.6以降、4GB以上) 、  
[インストール参照HP \(1\)](#)  
[インストール参照HP \(2\)](#) (このサイ

## Rのインストールと起動

基本的には[こちら](#)または[こちら](#)をご覧ください。  
よく分からぬ方へはWindowsユーザーの方へ：

1. [R-3.0.1-win.exe](#) (32/64 bitともこれ) をクリックして「実行する」を選択
2. 聞かれるがままに「次へ」などを押してとにかくインストールを完了させる
3. **Windows Vista**の人は(パッケージのインストール中に書き込み権限に関するエラーが出るのを避けるために)「コントロールパネル」-「ユーザーアカウント」-「ユーザーアカウント制御の有効化または無効化」で、「ユーザーアカウント制御(UAC)を使ってコンピュータの保護に役立たせる」の**チェックをあらかじめ外しておくことを強くお勧めします**。
4. インストールが無事完了したら、デスクトップに出現する「R3.0.1 (32 bitの場合)」または「R x64 3.0.1 (64 bitの場合)」アイコンをダブルクリックして起動
5. 以下を、「R コンソール画面上」でコピー&ペーストする。10GB程度のディスク容量を要しますが一番お手軽です。(どこからダウンロードするか?と聞かれるので、その場合は自分のいる場所から近いサイトを指定)

```
----- ここから -----  
install.packages(available.packages()[,1])
```

```
source("http://www.bioconductor.org/biocLite.R")  
biocLite(all_group())
```

```
----- ここまで -----
```

#CRAN中にある全てのパッケージをインストール

#おまじない  
#Bioconductor中にある全てのパッケージをインストール

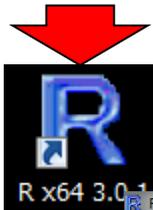
6. 「コントロールパネル」-「フォルダオプション」-「表示(タブ)」-「詳細設定」のところで、「登録されている拡張子は表示しない」のチェックを外してください。

cran.md.tsukuba.ac.jp から R-3.0.1-win.exe (51.5 MB) を実行しますか?

実行(R)

保存(S)

キャンセル(C)



```
RGui (64-bit)
ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ
[Icons]

R Console
R version 3.0.1 (2013-05-16) -- "Good Sport"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R は、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してください。

R は多くの貢献者による共同プロジェクトです。
詳しくは 'contributors()' と入力してください。
また、R や R のパッケージを出版物で引用する際の形式については
'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。

> |
```

```
RGui (64-bit)
ファイル 編集 閲覧 その他 パッケージ ウィンドウ ヘルプ
[Icons]

R Console
R version 3.0.1 (2013-05-16) -- "Good Sport"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R は、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してください。

R は多くの貢献者による共同プロジェクトです。
詳しくは 'contributors()' と入力してください。
また、R や R のパッケージを出版物で引用する際の形式については
'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。

> |
```

### Rのインストールと起動

基本的には[こちら](#)または[こちら](#)をご覧ください。  
よく分からない人でWindowsユーザーの方へ:

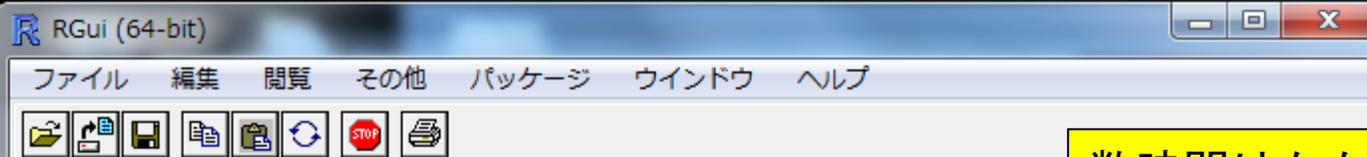
1. [R-3.0.1-win.exe](#) (32/64 bitともにこれ)をクリックして「実行する」を選択
2. 聞かれるがままに「次へ」などを押してとにかくインストールを完了させる
3. **Windows Vistaの人**は(パッケージのインストール中に書き込み権限に関するエラーが出るのを避けるために)「コントロールパネル」-「ユーザーアカウント」-「ユーザーアカウント制御の有効化または無効化」で、「ユーザーアカウント制御(UAC)を使ってコンピュータの保護に役立たせる」の**チェックをあらかじめ外しておくことを強くお勧め**します。
4. インストールが無事完了したら、デスクトップに出現する「R3.0.1 (32 bitの場合)」または「R x64 3.0.1 (64 bitの場合)」アイコンをダブルクリックして起動
5. 以下を、「R コンソール画面上」でコピー&ペーストする。10GB程度のディスク容量を要しますが一番お手軽です。(どこからダウンロードするか?と聞かれるので、その場合は自分のいる場所から近いサイトを指定)



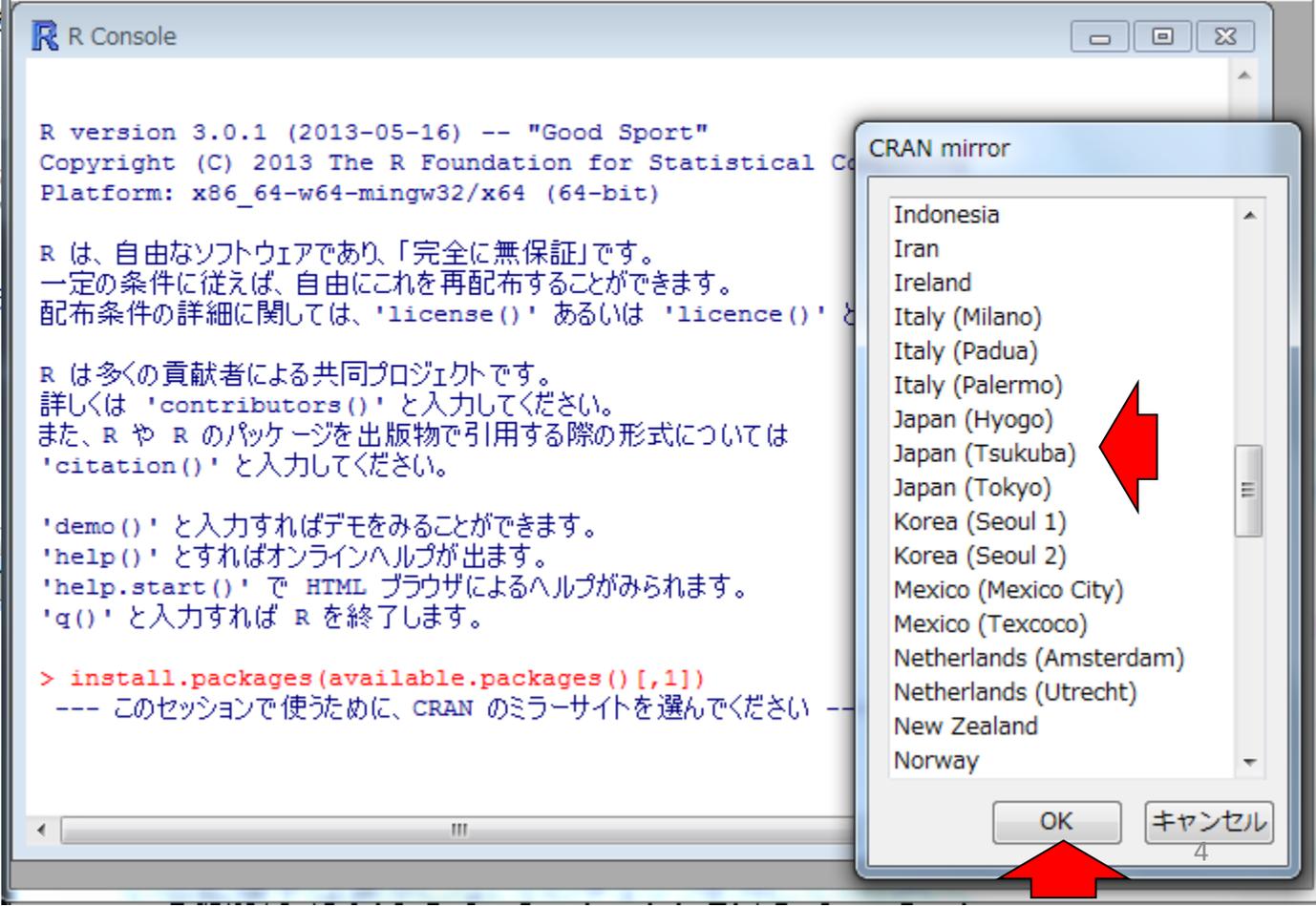
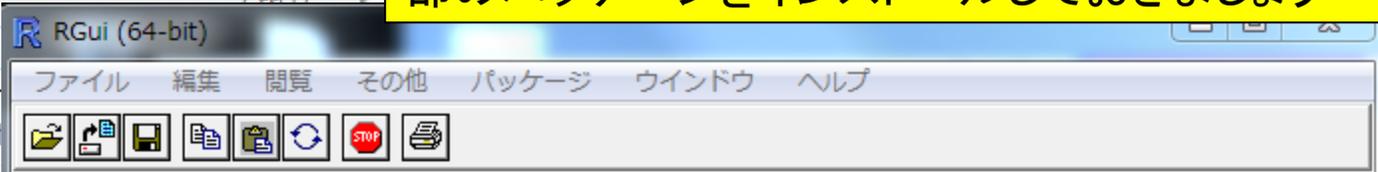
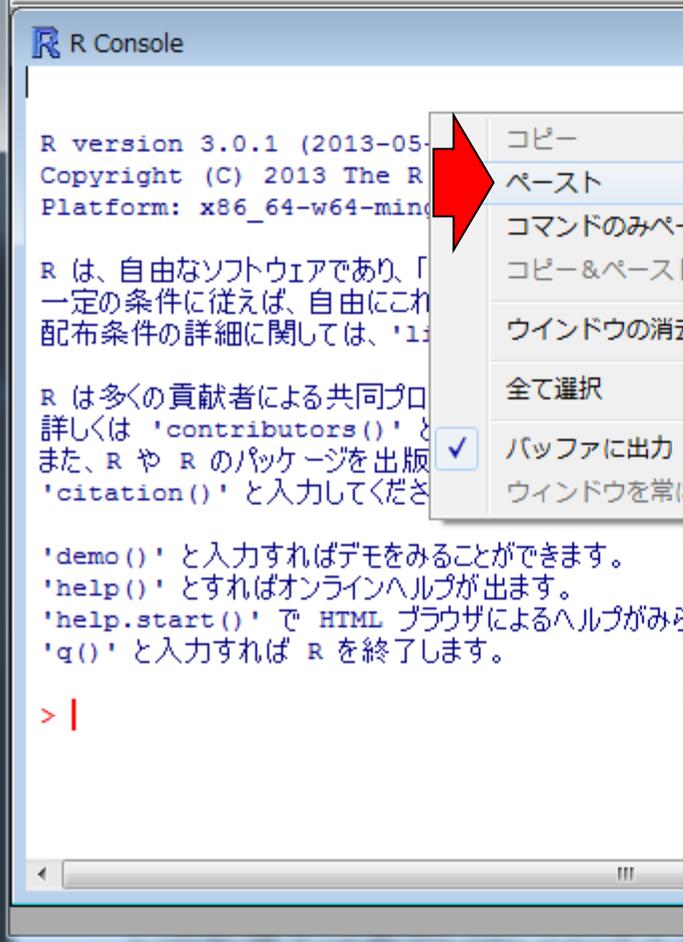
```
----- ここから -----
install.packages(available.packages()[,1])
source("http://www.bioconductor.org/biocLite.R")
biocLite(all_group())
----- ここまで -----
```

切り取り(T)  
**コピー(C)**  
 貼り付け  
 すべて選択(A)  
 印刷(I)...  
 印刷プレビュー(N)...  
 Bing でマップ  
 Bing で翻訳  
 Google で検索  
 電子メール (Windows Live Hotmail)

6. 「コントロールパネル」-「フォルダオプション」-「表示の拡張子は表示しない」のチェックを外してくだ



数時間はかかります。途中で何か問い合わせてきたりもしますが、基本的に言われるがままに全部のパッケージをインストールしておきましょう



- CRAN mirror
- Indonesia
  - Iran
  - Ireland
  - Italy (Milano)
  - Italy (Padua)
  - Italy (Palermo)
  - Japan (Hyogo)
  - Japan (Tsukuba)
  - Japan (Tokyo)
  - Korea (Seoul 1)
  - Korea (Seoul 2)
  - Mexico (Mexico City)
  - Mexico (Texcoco)
  - Netherlands (Amsterdam)
  - Netherlands (Utrecht)
  - New Zealand
  - Norway
- OK キャンセル

## Rのインストールと起動

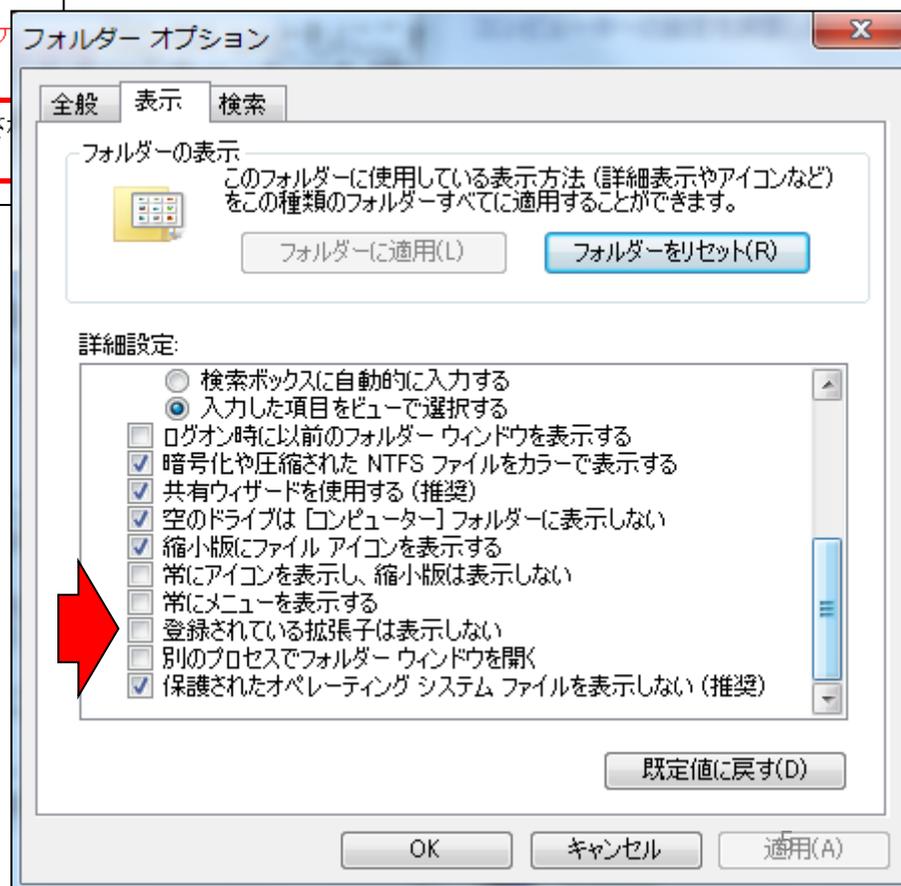
基本的には[こちら](#)または[こちら](#)をご覧ください。  
よく分からない人でWindowsユーザーの方へ:

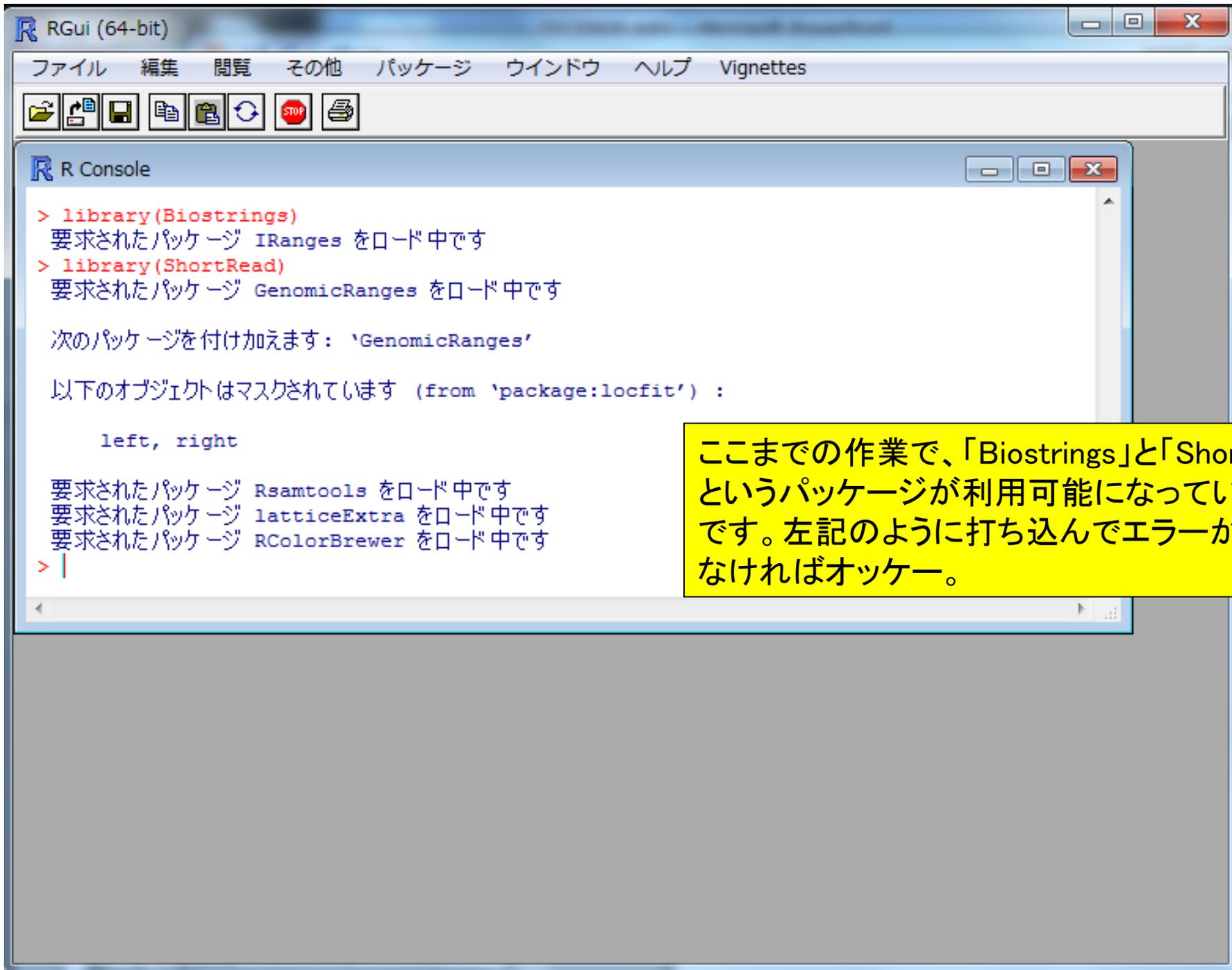
1. `R-3.0.1-win.exe` (32/64 bitとも(これ)をクリックして「実行する」を選択
2. 聞かれるがままに「次へ」などを押してとにかくインストールを完了させる
3. **Windows Vistaの人は**(パッケージのインストール中に書き込み権限に関するエラーが出るのを避けるために)「コントロールパネル」-「ユーザーアカウント」-「ユーザーアカウント制御の有効化または無効化」で、「ユーザーアカウント制御(UAC)を使ってコンピュータの保護に役立たせる」の**チェックを外す**ことを強くお勧めします。
4. インストールが無事完了したら、デスクトップに出現する「R3.0.1 (32 bitの場合)」または「R x64 3.0.1 (64 bitの場合)」アイコンをダブルクリックして起動
5. 以下を、「R コンソール画面上」でコピー&ペーストする。10GB程度のディスク容量を要しますが一番お手軽です。(どこからダウンロードするか?と聞かれるので、その場合は自分のいる場所から近いサイトを指定)

```
----- ここから -----  
install.packages(available.packages()[,1]) #CRAN中にある全てのパッケージをインストール  
  
source("http://www.bioconductor.org/biocLite.R") #おまじない  
biocLite(all_group()) #Bioconductor中にある全てのパッケージをインストール  
  
----- ここまで -----
```

6. 「コントロールパネル」-「フォルダオプション」-「表示(タブ)」-「詳細設定」のところで、「登録されている拡張子は表示しない」のチェックを外してください。

チェックが外れていればおっけーです。





ここまでの作業で、「Biostrings」と「ShortRead」というパッケージが利用可能になっているはずです。左記のように打ち込んでエラーが出ていなければオッケー。

http://www.iu.a.u-tokyo.ac.jp/~kadota/r\_seq.html#install Plant Global Educ... (Rで)塩基配列... x

## (Rで)塩基配列解析(主に次世代シーケンサーのデータ) (last modified 2013/05/23, since 2010)

What's new?

RパッケージTCCの最新版は1.1.99です。6月6日に開催されるNAIST植物グローバル教育プロジェクト・平成25年度ワークショップでは、R(ver. 3.0.1)とTCC(ver. 1.1.99)をベースに話を予定しています。Rのインストールと起動を実行したあとにTCC(ver. 1.1.99)のインストールも忘れずに行っておいてください。(2013/05/23) **NEW**

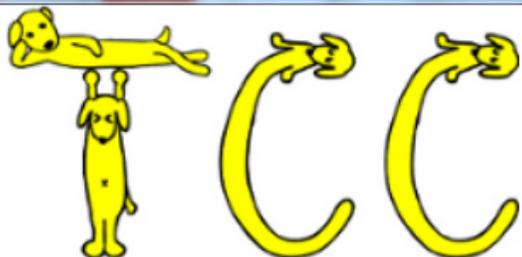
↑

- 平成25年6月27日、7月3, 4日にこのウェブページ関連の実習を含む講義(農学生命情報科学特論)を行います。東大生以外の外部の方も受講可能です。詳しくは事務局までお問い合わせください。(2013/05/23) **NEW**
- R3.0.1がリリースされていたのでこれに変更しました。(2013/05/17) **NEW**

次はTCCというRパッケージのインストールです。

- はじめに (last modified 2012/03/29)
- Rのインストールと起動 (last modified 2013/05/17) **NEW**
- サンプルデータ (last modified 2013/01/23)
- イントロダクション | NGS | 各種覚書 (last modified 2010/12/10)
- イントロダクション | NGS | 様々なプラットフォーム (last modified 2011/07/15)
- イントロダクション | NGS | リファレンス配列取得(マップされる側) (last modified 2011/02/03)
- イントロダクション | NGS | リファレンス配列取得後の各種情報抽出(特にRefSeq) (last modified 2011/03/20)
- イントロダクション | NGS | リファレンス配列取得後の各種情報抽出2(readFASTA関数の利用) (last modified 2011/04/07)
- イントロダクション | NGS | アノテーション情報取得(refFlatファイル) (last modified 2010/12/07)

ファイル(F) 検索(N) 無効化(S) 表示(V) イメージ(I) キャッシュ(C) ツール(T) 検証(A)



TCC ver. 1.1.99というパッケージの圧縮ファイル(\*.tar.gz形式のファイル)をデスクトップにダウンロードします。

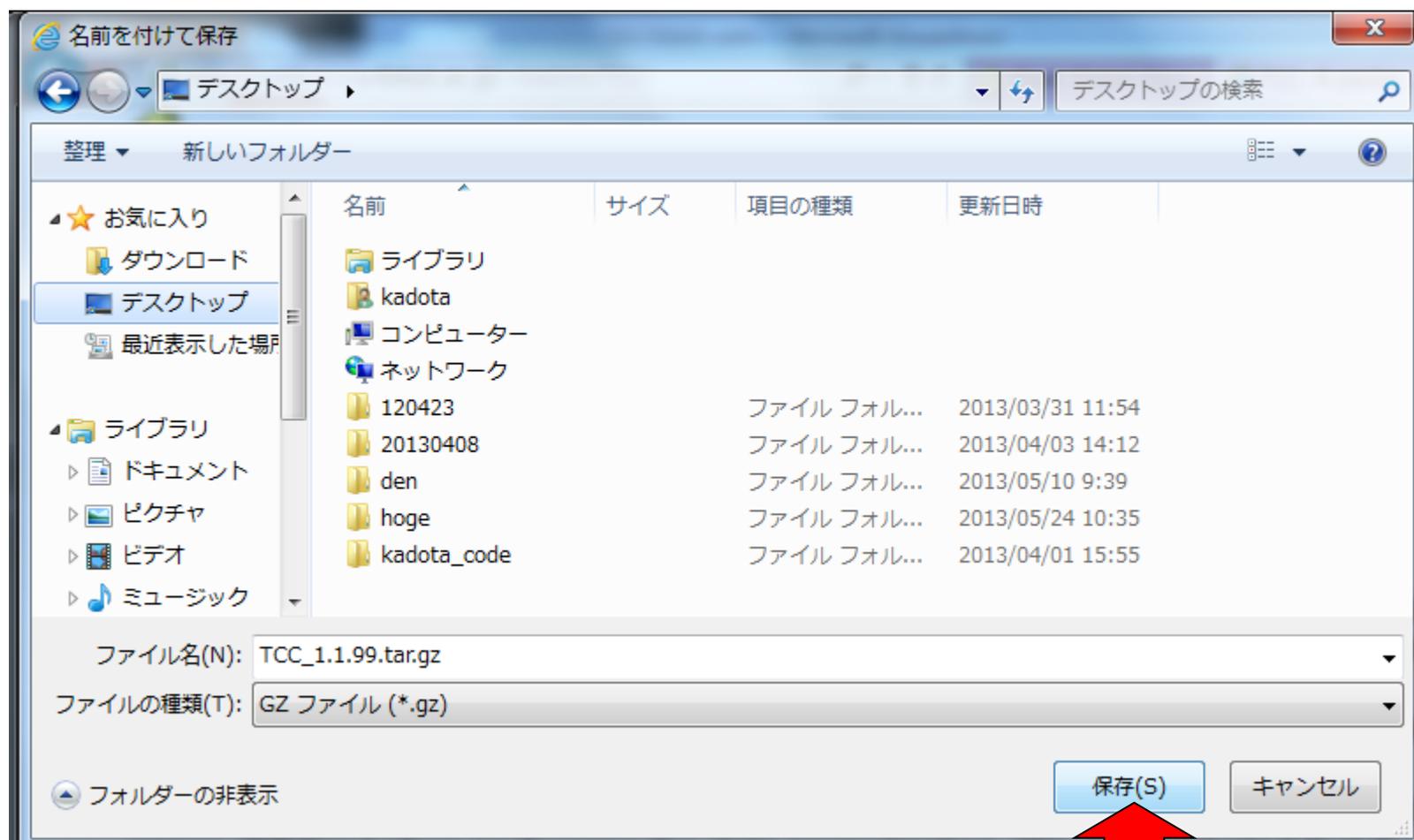
## TCC: an R package for comparing tag count data with robust normalization strategies

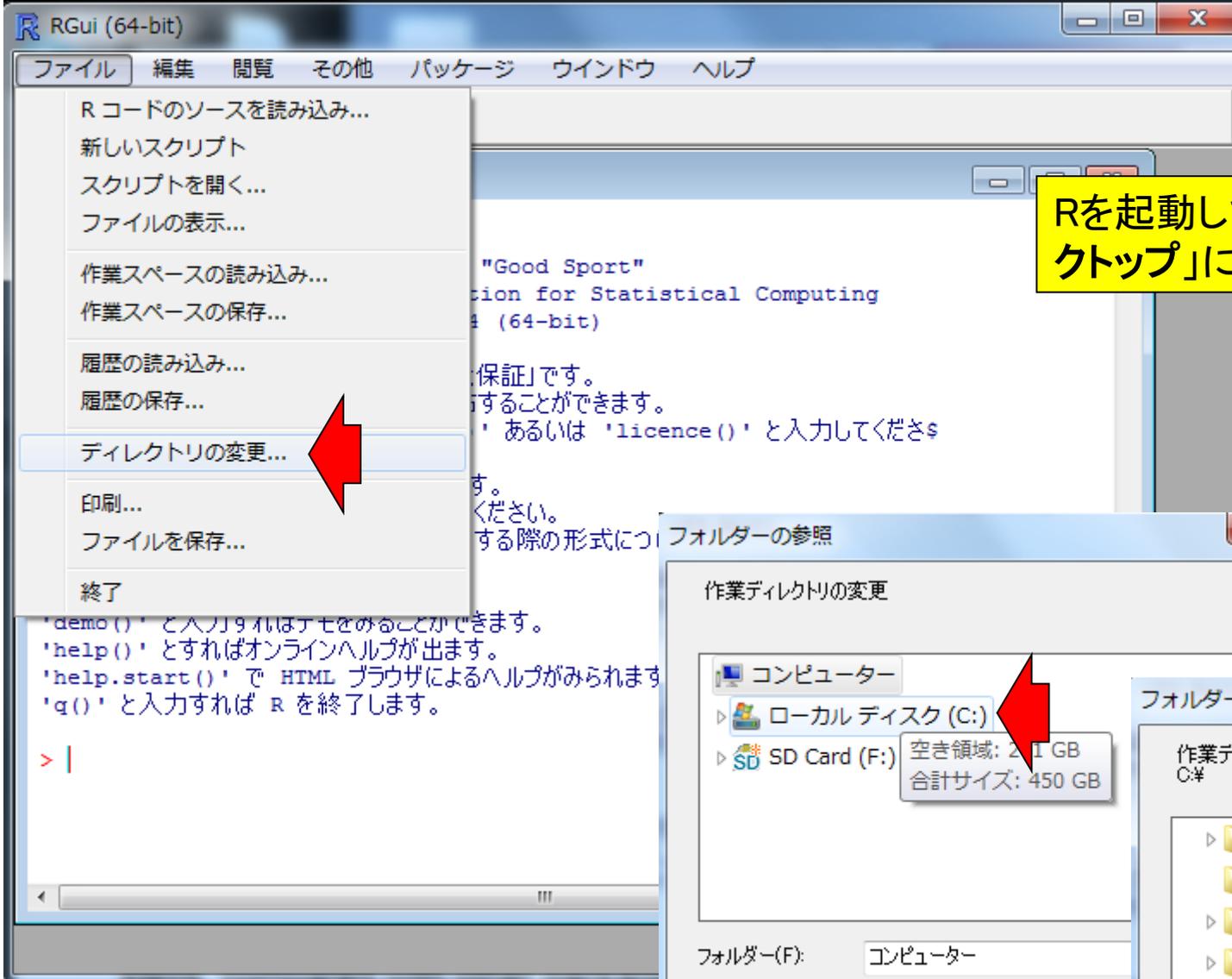
The R package, **TCC** provides users with a robust and accurate framework to perform differential expression analysis of tag count data. Differential expression analysis of tag count data (such as RNA-seq) from high-throughput sequencing technologies is a fundamental means of studying gene expression. We developed a normalization method ([TbT](#); Kadota et al., 2012) for replicates. The strategy is to remove data that are poor genes (DEGs) before performing the data normalization. We developed an elimination strategy (called DEGES) for data normalization. We provide a well-ranked gene list in which true DEGs are top-ranked. **TCC** provides integrated analysis pipelines in several steps, compared with other packages such as [edgeR](#) and [DESeq2](#) by appropriately combining their functionalities.

### Important note! (last modified: May 22, 2013)

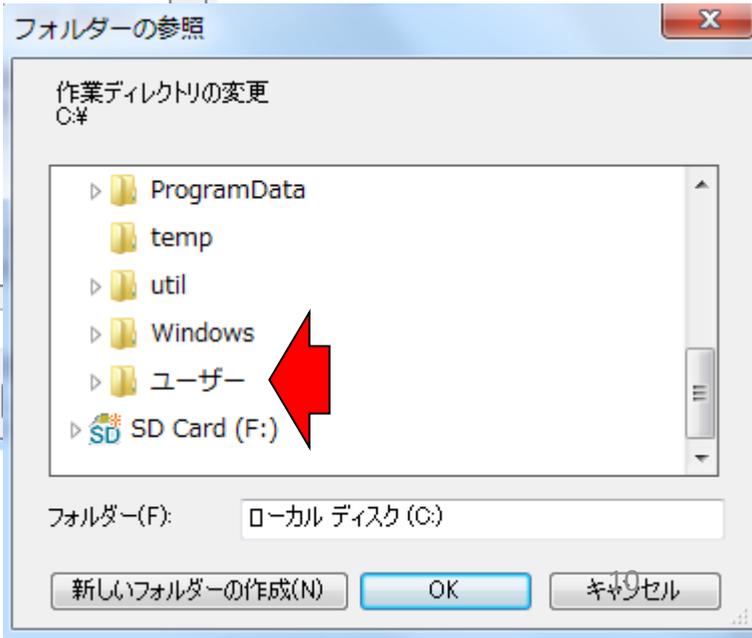
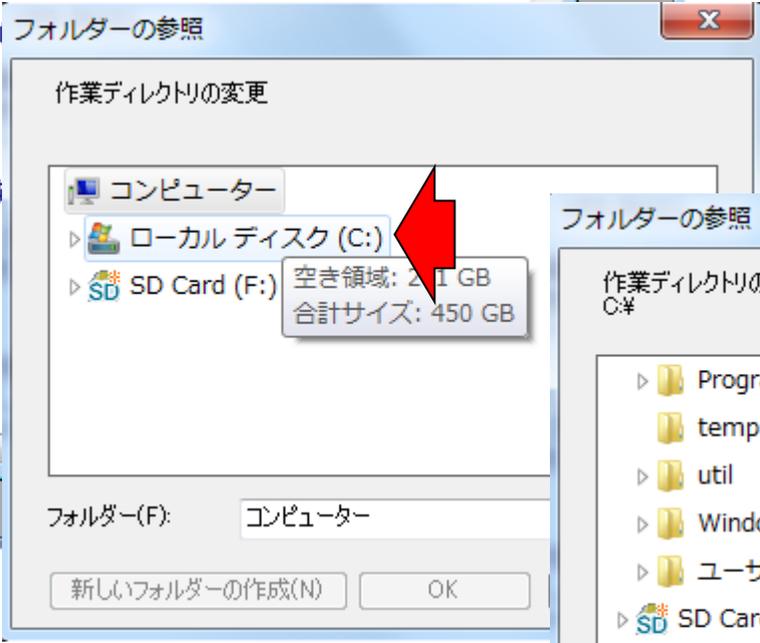
While the older version ([ver. 1.1.3](#)) of this package is available in the CRAN repository, we are now moving it from [CRAN](#) to [Bioconductor](#) until the next release (perhaps, ver. 1.2.0) of TCC is available. The latest version available on this webpage is [ver. 1.1.99](#).

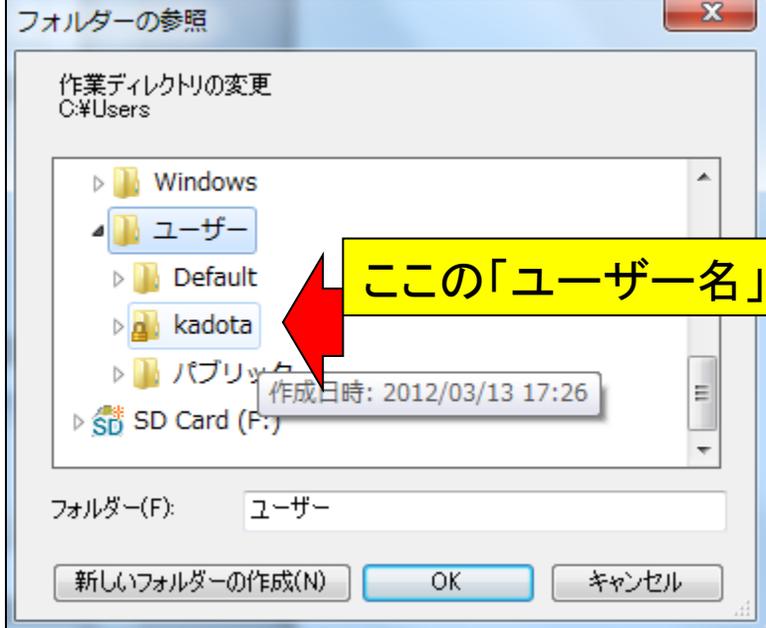




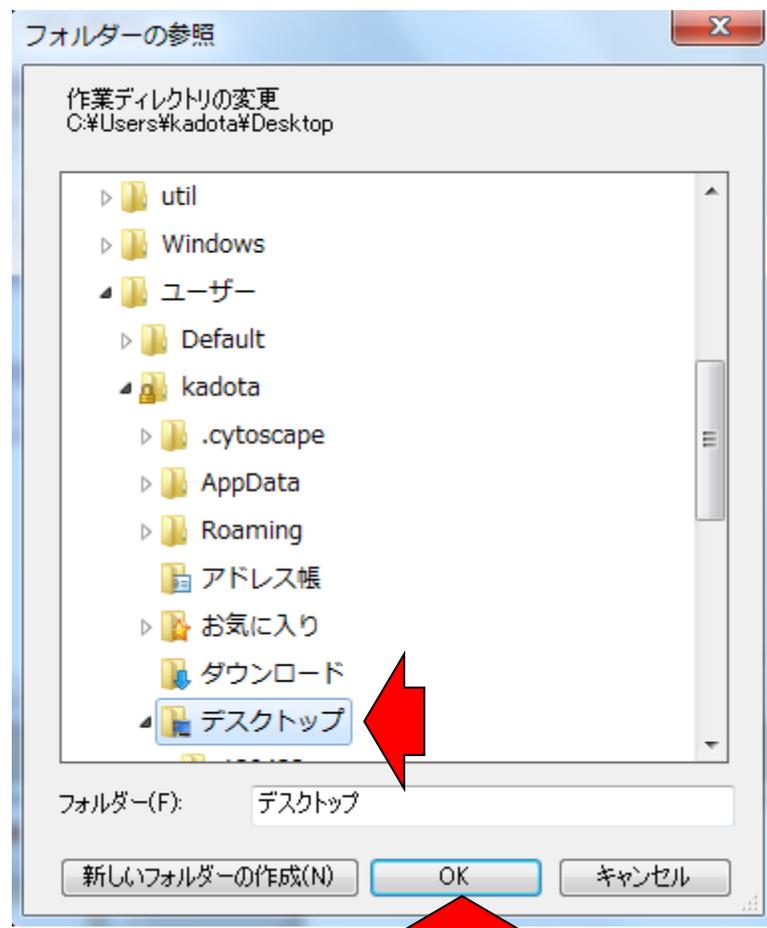


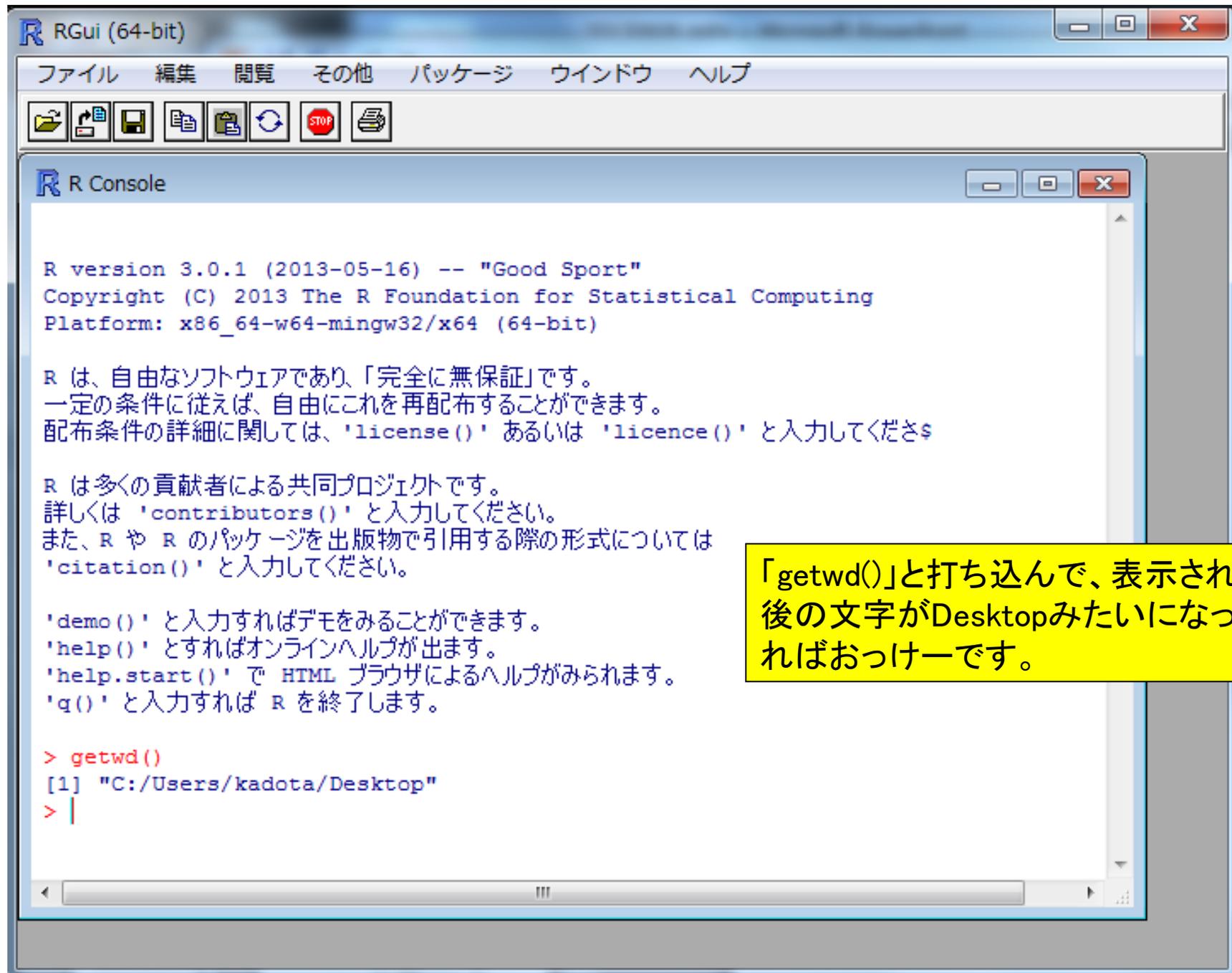
Rを起動して、作業ディレクトリを「デスクトップ」に変更します。





この「ユーザー名」はひとそれぞれ





## TCC: an R package for comparing tag count data with robust normalization strategies

The R package, **TCC** provides users with a robust and accurate framework to perform differential expression analysis of tag count data. Differential expression analysis of tag count data (such as RNA-seq) from high-throughput sequencing technologies is a fundamental means of studying gene expression. We recently developed a multi-step normalization method ([TbT](#); Kadota et al., 2012) for two-group RNA-seq data with replicates. The strategy is to remove data that are potential differentially expressed genes (DEGs) before performing the data normalization. We demonstrated that the DEG elimination strategy (called DEGES) for data normalization is essential for obtaining a well-ranked gene list in which true DEGs are top-ranked and non-DEGs are bottom ranked. **TCC** provides integrated analysis pipelines with improved data normalization steps, compared with other packages such as [edgeR](#), [DESeq](#), and [baySeq](#), appropriately combining their functionalities.

TCCパッケージのウェブサイトに行って、左下部分のコマンドをコピー

### Important note! (last modified: May 22, 2013)

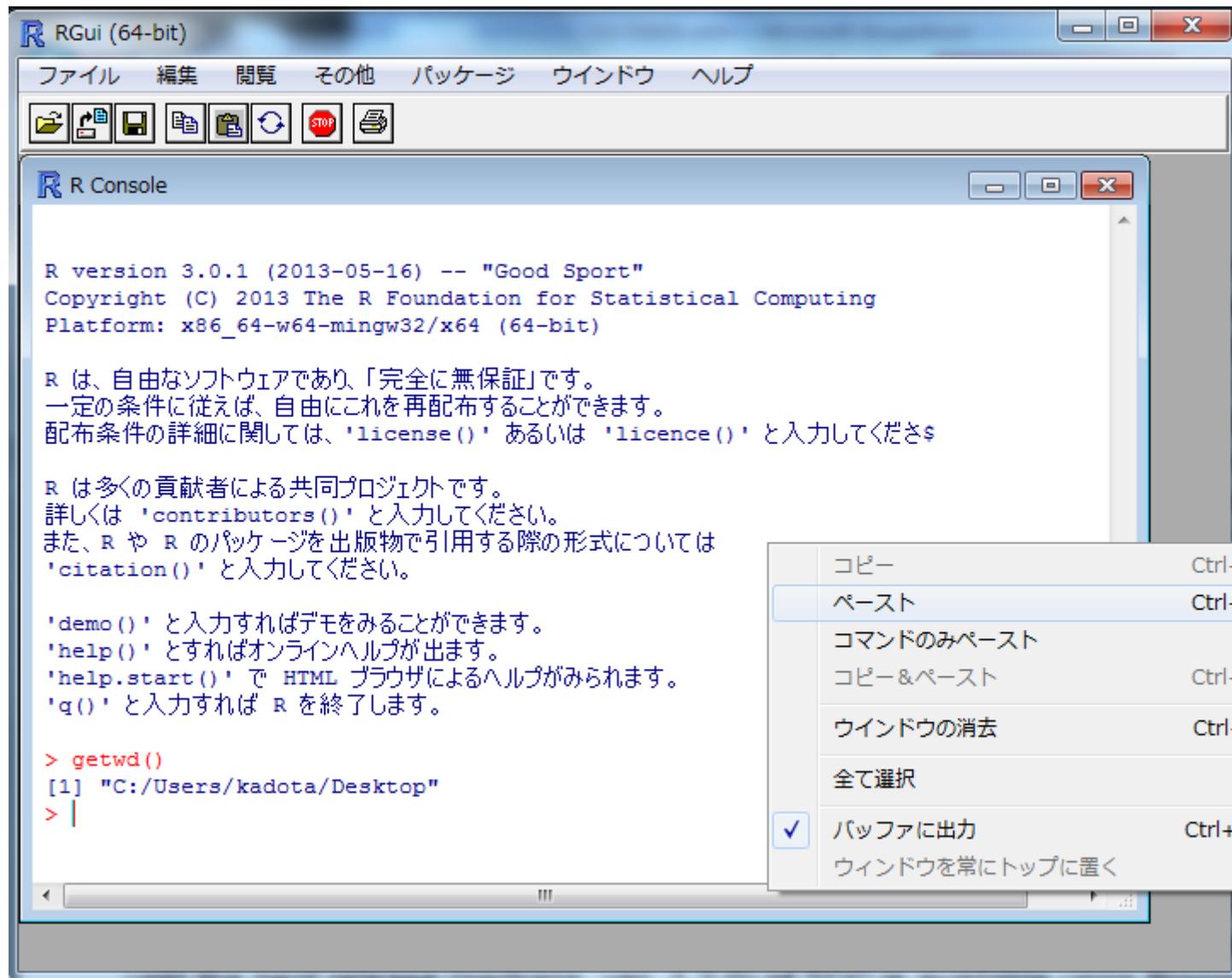
While the older version ([ver. 1.1.3](#)) of this package is currently available at repository, we are now moving it from [CRAN](#) to [Bioconductor](#). This webpage will remain available until the next release (perhaps, ver. 1.2.0) of TCC is available upon Bioconductor. The latest version available on this webpage is [ver. 1.1.99](#).

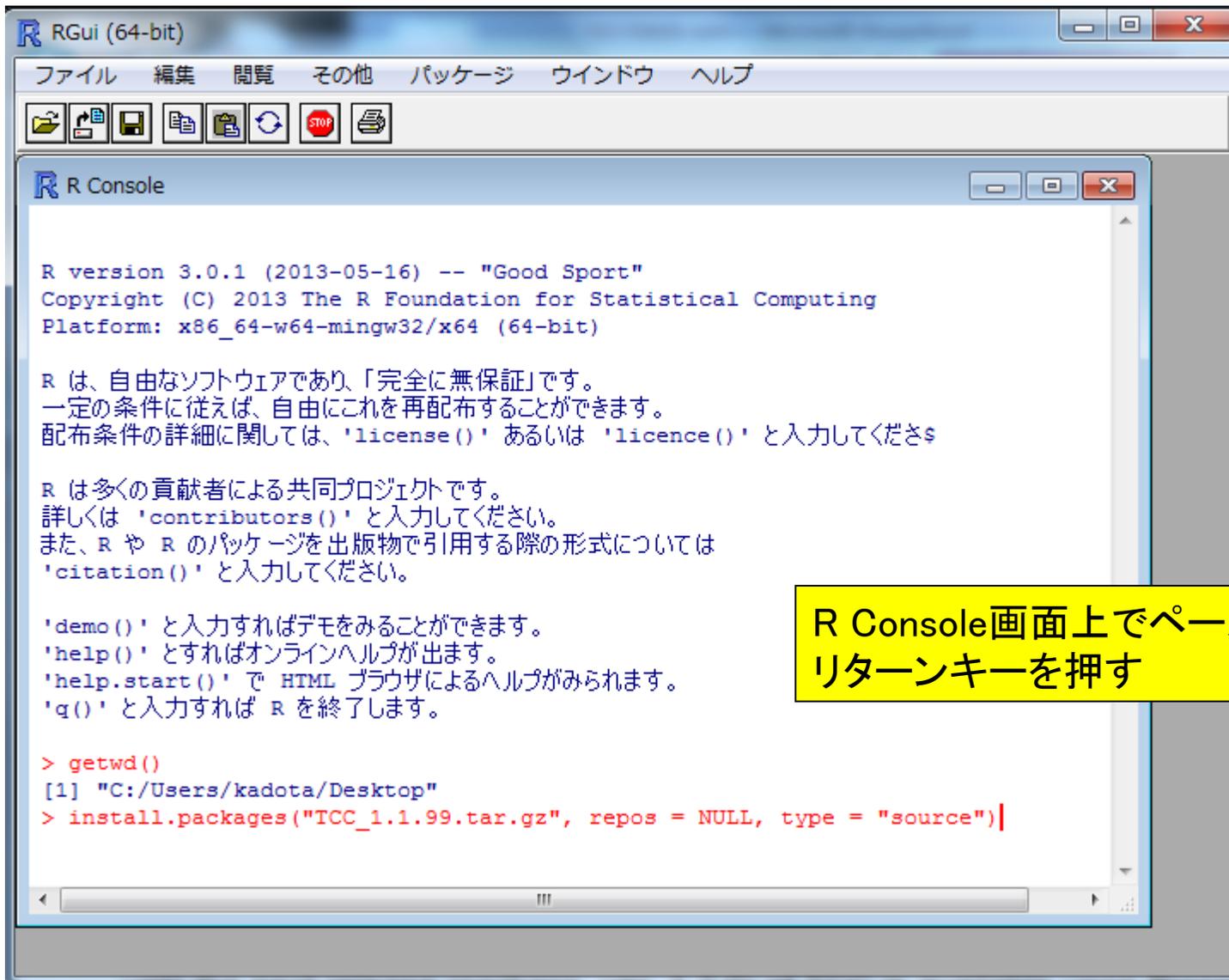
### Installation

To install the latest version ([ver. 1.1.99](#)) of this package, download the [source code](#) and enter the following command after starting [R](#):

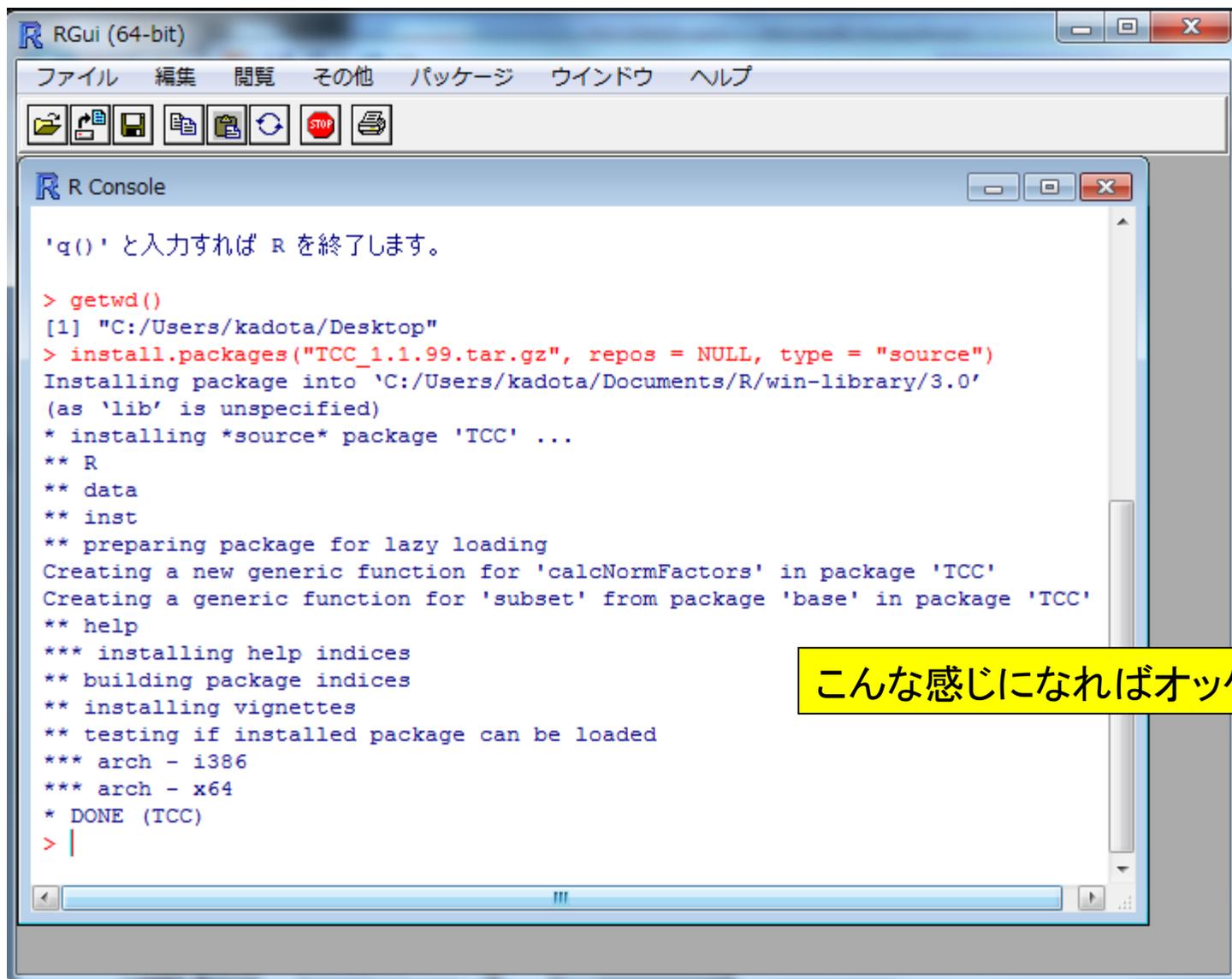
```
install.packages("TCC 1.1.99.tar.gz", repos = NULL, type = "source")
```

- 切り取り(T)
- コピー(C)
- 貼り付け
- すべて選択(A)
- 印刷(I)...
- 印刷プレビュー(N)...
- Bing でマップ
- Bing で翻訳
- Google で検索
- 電子メール (Windows Live Hotm...)
- すべてのアクセラレータ
- OneNote に送る(N)

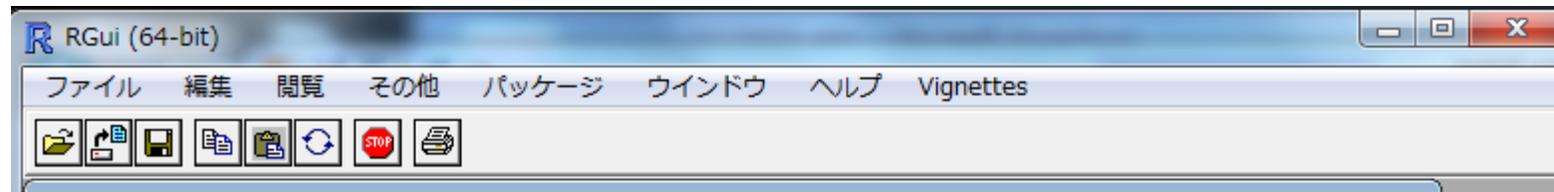




R Console画面上でペーストしたのちに、リターンキーを押す







```
R Console
> sessionInfo()
R version 3.0.1 (2013-05-16)
Platform: x86_64-w64-mingw32/x64 (64-bit)

locale:
 [1] LC_COLLATE=Japanese_Japan.932 LC_CTYPE=Japanese_Japan.932
 [3] LC_MONETARY=Japanese_Japan.932 LC_NUMERIC=Japanese_Japan.932
 [5] LC_TIME=Japanese_Japan.932

attached base packages:
 [1] parallel stats graphics grDevices

other attached packages:
 [1] TCC_1.1.99 ROC_1.36.0
 [5] limma_3.16.4 DESeq_1.12.0
 [9] Biobase_2.20.0 BiocGenerics_0.6.0

loaded via a namespace (and not attached):
 [1] annotate_1.38.0 AnnotationDbi_1.36.0
 [4] genefilter_1.42.0 geneplotter_1.38.0 grid_3.0.1
 [7] IRanges_1.18.1 RColorBrewer_1.0-5 RSQlite_0.11.3
 [10] splines_3.0.1 stats4_3.0.1 survival_2.37-4
 [13] XML_3.96-1.1 xtable_1.7-1
```

```
R Console
> sessionInfo()
R version 3.0.1 (2013-05-16)
Platform: x86_64-w64-mingw32/x64 (64-bit)

locale:
 [1] LC_COLLATE=Japanese_Japan.932 LC_CTYPE=Japanese_Japan.932
 [3] LC_MONETARY=Japanese_Japan.932 LC_NUMERIC=Japanese_Japan.932
 [5] LC_TIME=Japanese_Japan.932

attached base packages:
 [1] parallel stats graphics grDevices

other attached packages:
 [1] TCC_1.1.99 ROC_1.36.0
 [5] limma_3.16.4 DESeq_1.12.0
 [9] Biobase_2.20.0 BiocGenerics_0.6.0

loaded via a namespace (and not attached):
 [1] annotate_1.38.0 AnnotationDbi_1.36.0
 [4] genefilter_1.42.0 geneplotter_1.38.0 grid_3.0.1
 [7] IRanges_1.18.1 RColorBrewer_1.0-5 RSQlite_0.11.3
 [10] splines_3.0.1 stats4_3.0.1 survival_2.37-4
 [13] XML_3.96-1.1 xtable_1.7-1
```

コピー	Ctrl+C
ペースト	Ctrl+V
コマンドのみペースト	
コピー&ペースト	Ctrl+X
ウインドウの消去	Ctrl+L
全て選択	
<input checked="" type="checkbox"/> バッファに出力	Ctrl+W

もしエラーが出る場合は「sessionInfo()」と打ち込んで表示される情報をコピーして...

インストールできません - メッセージ (テキスト形式)

ファイル    メッセージ    挿入    オプション    書式設定    校閲    活用しよう! アウトLOOK

送信(S)    差出人(M)    kadota@iu.a.u-tokyo.ac.jp

宛先...    CC(C)...    BCC(B)...

件名(U):    インストールできません

```
たすけてください。。。↵
↵
> sessionInfo()↵
R version 3.0.1 (2013-05-16)↵
Platform: x86_64-w64-mingw32/x64 (64-bit)↵
↵
locale:↵
[1] LC_COLLATE=Japanese_Japan.932 LC_CTYPE=Japanese_Japan.932 ↵
[3] LC_MONETARY=Japanese_Japan.932 LC_NUMERIC=C ↵
[5] LC_TIME=Japanese_Japan.932 ↵
↵
attached base packages:↵
[1] parallel stats graphics grDevices utils datasets methods base ↵
↵
other attached packages:↵
[1] TCC_1.1.99 ROC_1.36.0 baySeq_1.14.0 edgeR_3.2.3 ↵
[5] limma_3.16.4 DESeq_1.12.0 lattice_0.20-15 locfit_1.5-9.1 ↵
[9] Biobase_2.20.0 BiocGenerics_0.6.0↵
↵
loaded via a namespace (and not attached):↵
[1] annotate_1.38.0 AnnotationDbi_1.22.5 DBI_0.2-7 ↵
[4] genefilter_1.42.0 geneplotter_1.38.0 grid_3.0.1 ↵
[7] IRanges_1.18.1 RColorBrewer_1.0-5 RSQLite_0.11.3 ↵
[10] splines_3.0.1 stats4_3.0.1 survival_2.37-4 ↵
[13] XML_3.96-1.1 xtable_1.7-1 ↵
>↵
↵
```

メールを出す