GANreproduce Quick Start Guide

Oct. 12, 2018

Program MATLAB (2016a for demonstration)

Type *Package (GANreproduce.m)*

Title *A package for reproducing the results in the manuscript entitled:*

A trio of genes in germinal center B cells mediates the pathogenesis of blood cancers

Version *1.4*

Date *2018/10/11*

Authors *Wei-Quan Fang and Ming-Jing Hwang*

Maintainer *Wei-Quan Fang < deleapoli@gmail.com >*

Description *In this work, we analyzed lymphoma-related gene expression and clinical data and identified a regulatory motif of germinal center B cell genes. We conduct multiple blood cancer prognoses and uncover tumor suppressor functions of BACH2 which are much more important than previous thought.*

Repository *GITHUB*

Publication *Submitted*

Quick Start Guide

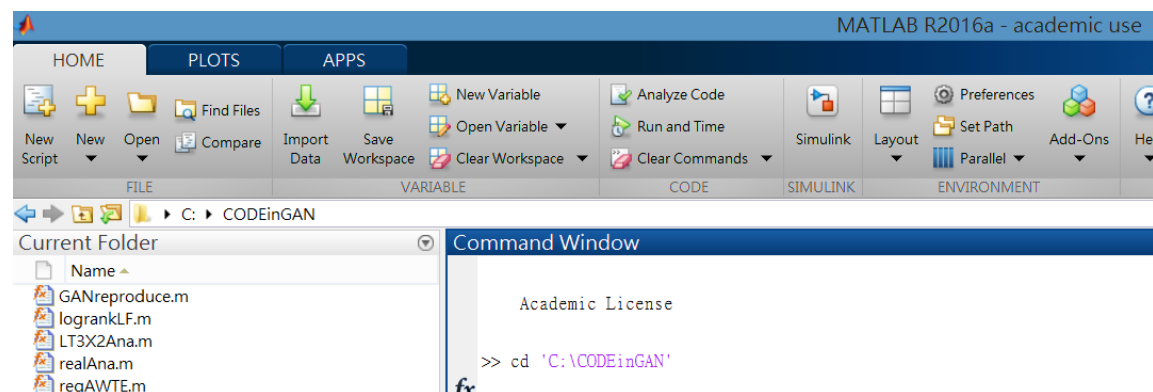Three steps are necessary before the launch of the *GANreproduce* Package.

Step1: Download the two files, **DATAinGAN** and **CODEinGAN**, from the web-site https://github.com/IBMSn121/GAN-paper1.

🔒 GitHub, Inc. [US] | https://github.com/IBMSn121/GAN-paper1		

IBMSn121 Update README.md		Latest commit 6f143b5 5 hours ago
📁 CODEinGAN	Create filenameCode.txt	6 hours ago
📁 DATAinGAN	Create filenameData.txt	6 hours ago
📄 LICENSE	Initial commit	3 months ago
📄 README.md	Update README.md	5 hours ago

and put them in **C:**

📁 CODEinGAN	2018/10/11
📁 DATAinGAN	2018/10/9
📁 Program Files	2017/5/18
📁 Program Files (x86)	2018/5/17

Step2: Open MATLAB program window and type **cd 'C:\CODEinGAN'** in it, then press **ENTER**.

MATLAB R2016a - academic use

HOME   PLOTS   APPS

New Script | New | Open | Find Files | Compare | Import Data | Save Workspace | New Variable | Open Variable ▼ | Clear Workspace ▼ | Analyze Code | Run and Time | Clear Commands ▼ | Simulink | Layout ▼ | Preferences | Set Path | Parallel ▼ | Add-Ons ▼ | Help

FILE   VARIABLE   CODE   SIMULINK   ENVIRONMENT

◀ ▶ 🔳 📁 📁 ▶ C: ▶ CODEinGAN

Current Folder

Name ▲
- GANreproduce.m
- logrankLF.m
- LT3X2Ana.m
- realAna.m
- regAWTE.m

Command Window

Academic License

>> cd 'C:\CODEinGAN'

fx

(***Please check*** that GANreproduce Package **GANreproduce.m** can be

seen in the MATLAB panel of **Current Folder** after this step)

Step3: Type **GANreproduce** and press **ENTER** to launch the GANreproduce

Package and start ***dialog boxes*** for reproducing the results.
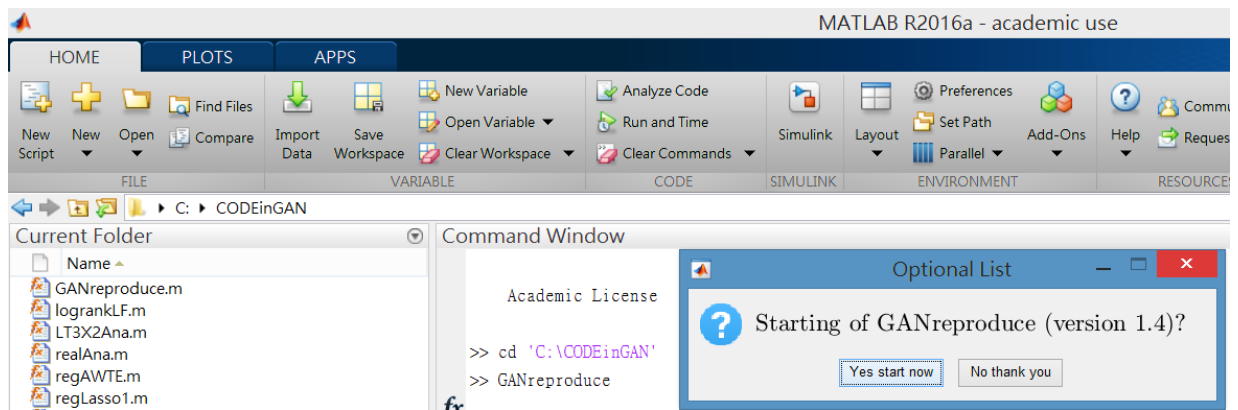
Examples

Three examples, concerning **Real Data Analysis**, **Linear Trend Analysis** and **Clinical Controversy Analysis**, were provided for reproducing the results via GANreproduce Package with selections of *user friendly dialog boxes*.

Example 1 (Real Data Analysis)

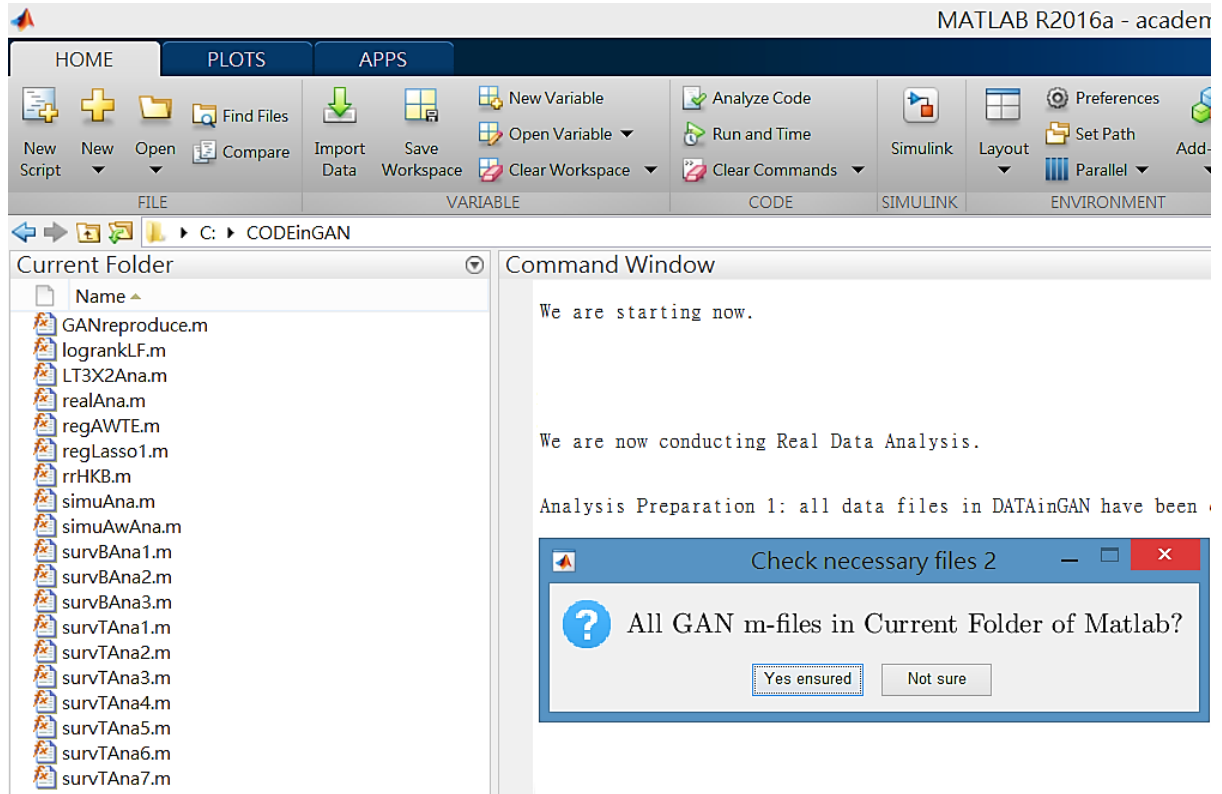Click **Yes start now** after the launch of the Package, then click **Real Data**.

↓
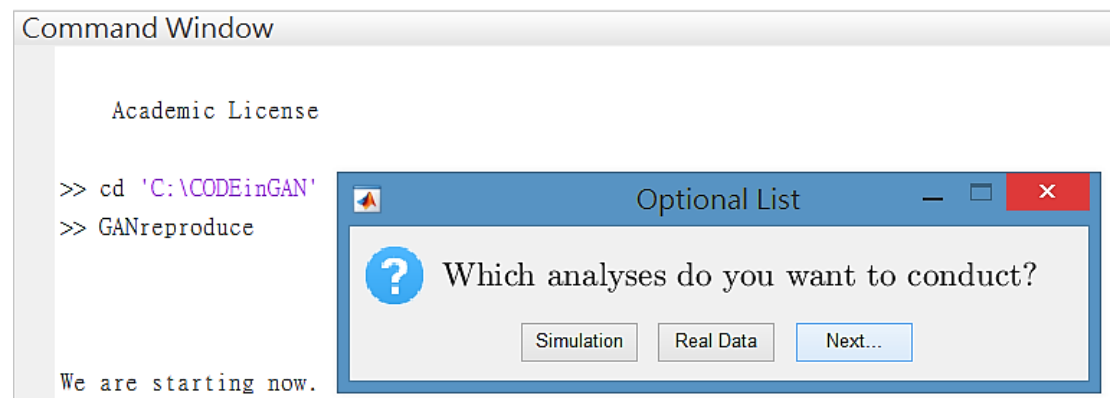
Ensure all the data in DATAinGAN were set up right and click **Yes ensured**.
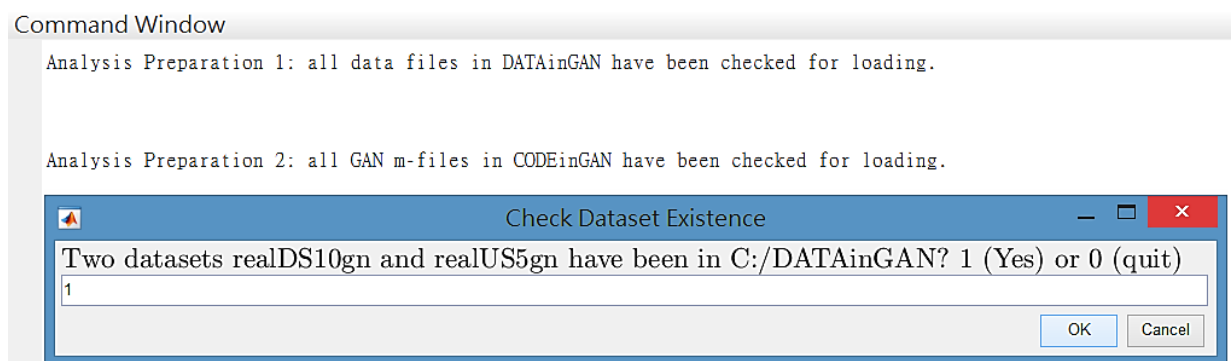
↓
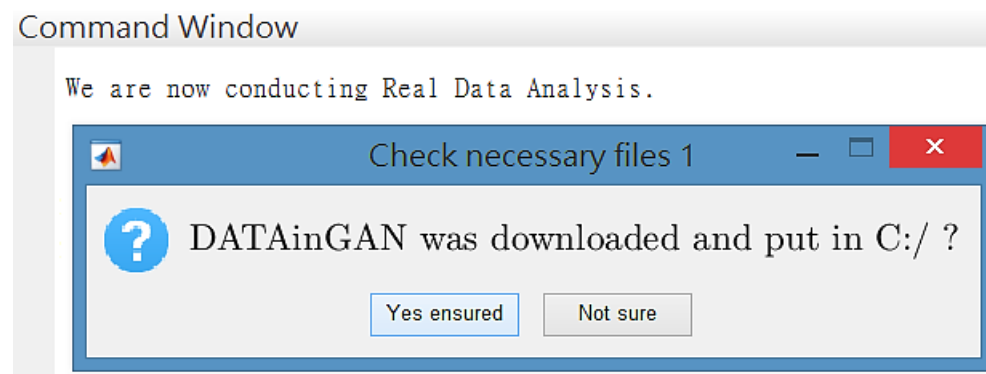
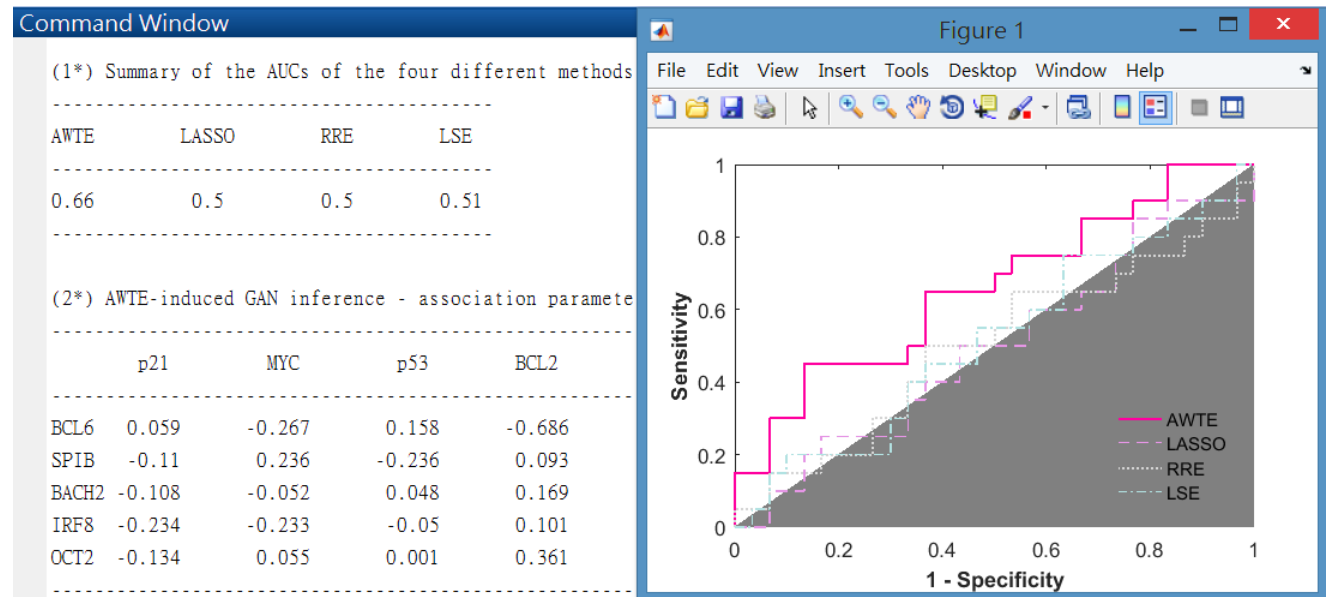Ensure all the codes in CODEinGAN were set up right and click **Yes ensured**.

↓

Check the two datasets realDS10gn and realUS5gn used in real data analysis were set up right, then select **1 (Yes)** and click **Ok**.

↓

Results for Real Data Analysis were reproduced on the **Figure** and **MATLAB Command Window**.

Command Window

```
(1*) Summary of the AUCs of the four different methods

-------------------------------------------
AWTE        LASSO        RRE        LSE
-------------------------------------------
0.66         0.5         0.5        0.51
-------------------------------------------


(2*) AWTE-induced GAN inference - association paramete

-------------------------------------------
         p21         MYC         p53        BCL2
-------------------------------------------
BCL6    0.059     -0.267       0.158      -0.686
SPIB   -0.11       0.236      -0.236       0.093
BACH2  -0.108     -0.052       0.048       0.169
IRF8   -0.234     -0.233      -0.05        0.101
OCT2   -0.134      0.055       0.001       0.361
-------------------------------------------
```


Figure 1
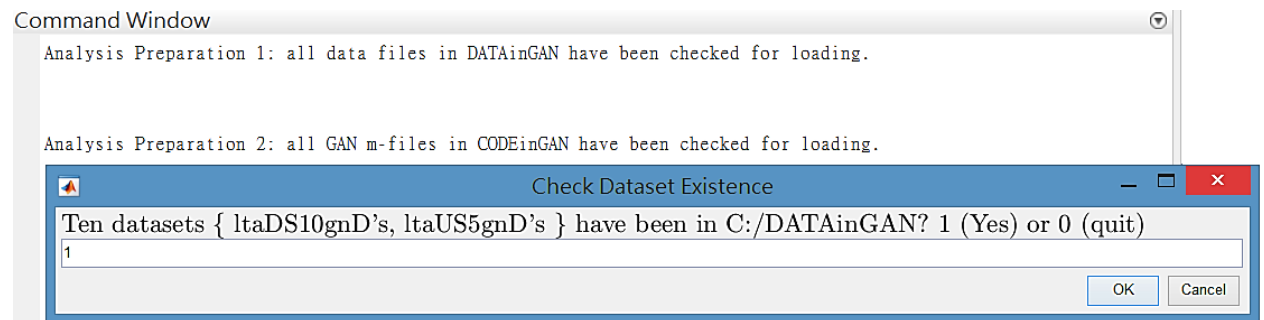
Example 2 (Linear Trend Analysis)

Click **Yes start now** after the launch of the Package, then click **Next…** and **Linear Trend**.
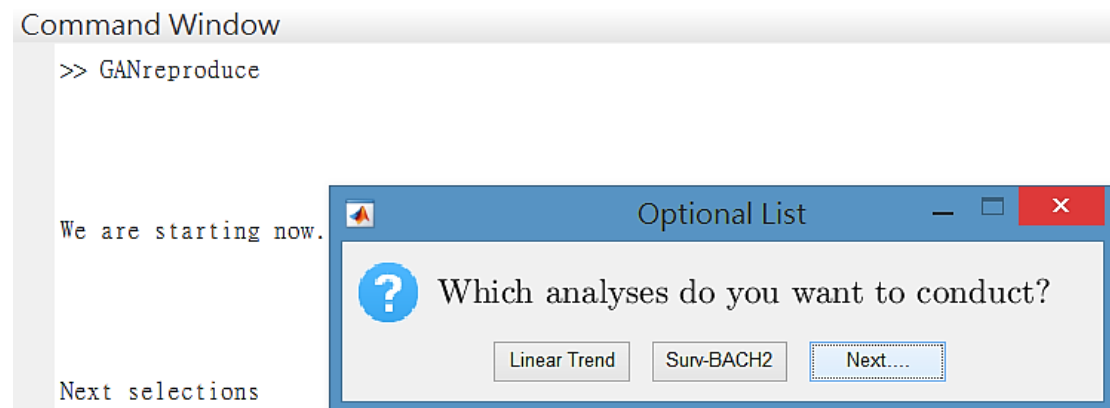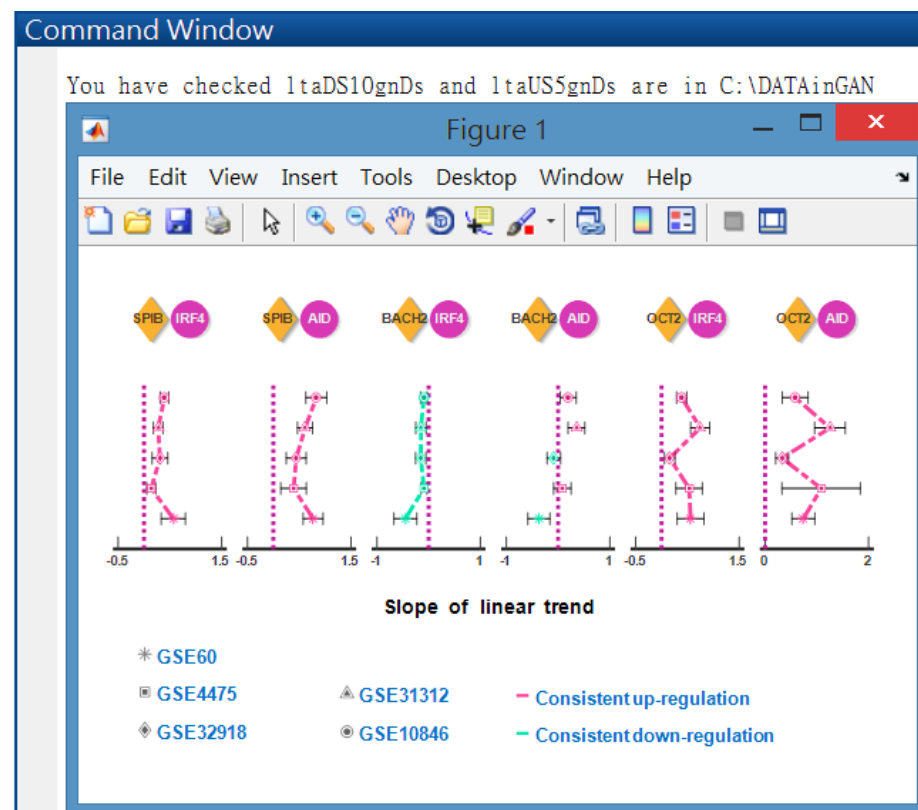
↓

Click **Yes ensured** twice and check the necessary files, datasets and codes, as in Example 1, then select **1 (Yes)** and click **Ok**.
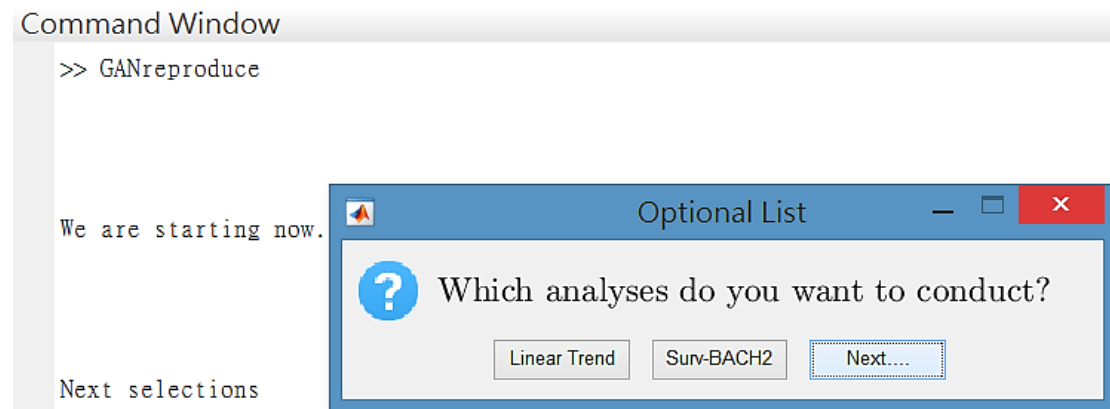

↓

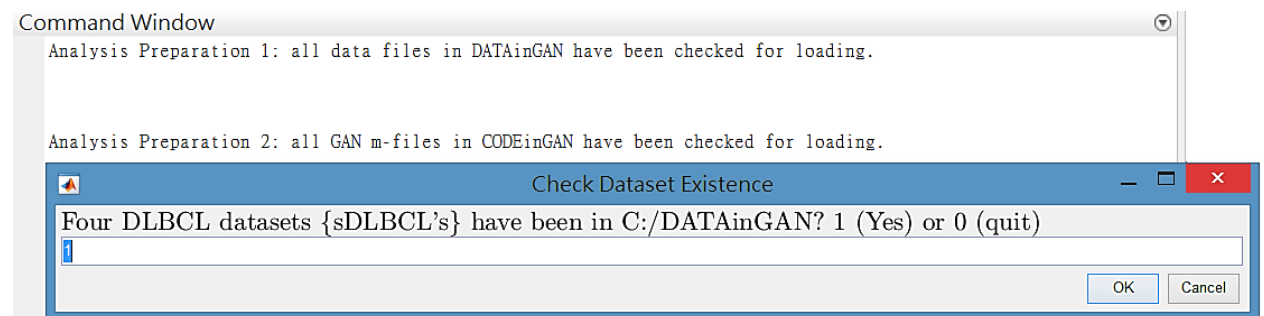Results for Linear Trend Analysis were reproduced on the **Figure**.

Example 3 (Clinical Controversy Analysis)

Click **Yes start now** after the launch of the Package, then click **Next...** and **Surv-BACH2**.
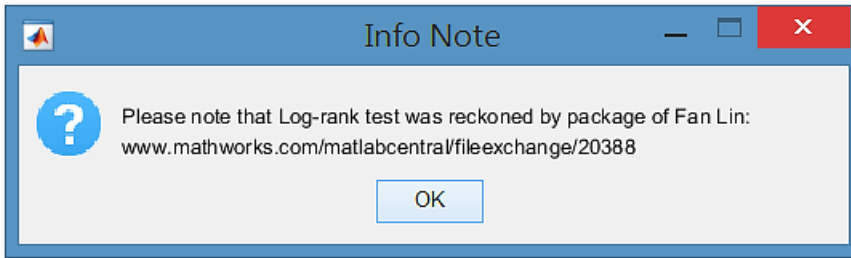
Command Window
```
>> GANreproduce



We are starting now.



Next selections
```

Optional List — □ ×

? Which analyses do you want to conduct?

[Linear Trend] [Surv-BACH2] [Next....]

↓

Click **Yes ensured** twice and check the necessary files, datasets and codes, as in Example 1, then select **1 (Yes)** and click **Ok**.

Command Window
```
Analysis Preparation 1: all data files in DATAinGAN have been checked for loading.


Analysis Preparation 2: all GAN m-files in CODEinGAN have been checked for loading.
```

Check Dataset Existence — □ ×

Four DLBCL datasets {sDLBCL's} have been in C:/DATAinGAN? 1 (Yes) or 0 (quit)

1

[OK] [Cancel]

↓

Receive the information on the package used for log-rank statistical testing and click **Ok**.

Command Window

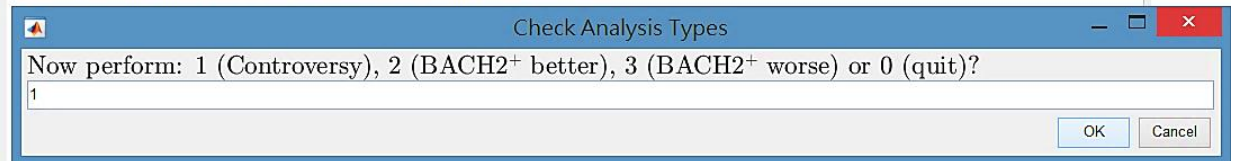You have checked 4 DLBCL datasets are in C:\DATAinGAN

Info Note — □ ✕

? Please note that Log-rank test was reckoned by package of Fan Lin:
www.mathworks.com/matlabcentral/fileexchange/20388

OK

↓

Select analytical type **1 (Controversy)** and click **Ok** .

Command Window

You have checked 4 DLBCL datasets are in C:\DATAinGAN
Note that Log-rank test was reckoned by package of Fan Lin

Check Analysis Types — □ ✕

Now perform: 1 (Controversy), 2 (BACH2$^+$ better), 3 (BACH2$^+$ worse) or 0 (quit)?

1

OK   Cancel

↓

Results for Clinical Controversy Analysis were reproduced.

Command Window

You have checked 4 DLBCL datasets are in C:\D.
Note that Log-rank test was reckoned by packa;
Now we perform Clinical Controversy for BACH2

(1.1*) Controversy of BACH2's role - GSE4475

Summary of the Number of Censored and Uncenso

GROUP	Total	Failed	Censored
X	124	59	65
Y	31	11	20
Total	155	70	85

Chi-square:1.0630   P:0.3025