Cogent™ NGS Discovery Software v2.0.14 (beta) Quick Start Guide

Cogent NGS Discovery Software (Cogent DS) enables analysis of both single-cell and bulk RNA-seq sequencing data prepared with Cogent NGS Analysis Pipeline Software. This guide is meant to be a high-level overview of how to install, run, and use the software; for more detailed instructions, please contact technical_support@takarabio.com.

Before you begin

A. Supported operating systems
- Windows 11
- macOS Mojave (10.14) or higher
- Linux CentOS 6.9 / Red Hat 7 or higher

B. Hardware requirements
- Standard laptop, desktop, or server
- Memory: 8 GB RAM or higher
- Free disk space: 2 GB or higher

C. Additional dependencies
- Internet connectivity on the computer
- R version 4.4.0 or higher, and:
  - For Windows: RTools 4.4 or higher, based on the R version
  - For MacOS: Homebrew; after R installation, please install xcode tools and run brew install libomp and brew install gcc
  - For Linux (CentOS): use yum install to install libomp-devel and gcc
- RStudio (IDE for R)

Required input files

CogentDS can take either of the following options for input (choose one):

1. `CogentDS.analysis.rds`, an rds file output from Cogent NGS Analysis Pipeline (CogentAP v3.0.19 (beta))—recommended input for full analysis capabilities
2. `.rda` file, including backward compatibility with .rda files from CogentAP v2.0.1.

   NOTE: Certain functionalities may be limited when using an RDA file from a previous CogentAP version.
3. Raw gene-count matrix and stats/metadata files (.csv/.csv.gz)

Installation

1. Sign up to download the installation package from our website.
2. Run RStudio.
3. (Uninstall previous software versions): If you have an older version of CogentDS or the Takara Bio hanta software installed, type the following command(s) at the RStudio prompt to uninstall them:

   CogentDS: `remove.packages("CogentDS")`

   hanta: `remove.packages("hanta")`
4. Type the following commands at the RStudio prompt to install CogentDS:

```
setwd('<PATH>')
source ('setup_CogentDS.R')
```

   where **<PATH>** is replaced by the full path of the folder created by unzipping the software zip file.
5. Select a CogentDS installation option from the selection prompt options:
   - Selection 1 is the default and recommended for most users
   - Selection 2 is an alternative and can be used if default gives an error during installation

Running Cogent NGS Discovery Software

1. Run RStudio.
2. At the RStudio prompt, run the command:

```
setwd('<PATH>')
source ('launch_CogentDS.R')
```

   where **<PATH>** is replaced by the full path of the folder created by unzipping the software zip file.

   The CogentDS user interface (GUI) will display in the default browser for the workstation or server.

Single-cell RNA-seq analysis

1. Click [Launch scRNA app] in the GUI. A new window will appear. Click [Analysis Mode].

   NOTE: Clicking [Discovery Mode] will allow for the viewing and manipulation of already-processed data
2. *Select Input Data*: Upload the required input files by selecting one of the options from the radio buttons**.**

   NOTE: To explore CogentDS functionalities, example data can be used by choosing the "Example data" radio button.

Click [Prepare Data for downstream analysis (Required)]. After data preparation is complete, the "Data preparation is complete" window will appear. Click [Ok] and then [Next: Perform Ambient RNA Filtering] on the bottom right of the GUI.

3. *Ambient RNA Filtering:* The results of ambient RNA decontamination can be viewed as a dot plot of the top 10 highly variable genes across estimated clusters.

   Click [Perform ambient RNA decontamination] to perform ambient RNA decontamination. Click [Skip ambient RNA decontamination] to skip this step.

   When finished, click [Next: Perform QC] on the bottom right of the GUI.

4. *Pre-QC and QC Analysis*: Perform cell-level filtering at this step and visualize pre- and post-QC violin and scatter plots.

   Set QC parameters as desired. Click [Apply QC filters] to perform QC analysis and view post-QC plots. Click [Skip QC] to skip this step.

   NOTE: First-time users should start by using the default settings for the QC options.

   The "QC is complete" window will appear when QC is complete. Click [Ok] and then [Next: Perform normalization, feature selection & scaling] on the bottom right of the GUI.

5. *Normalization*: Select how to normalize and scale the data. Scaling can be done using variable features or all features.

   Set normalization, features, and scaling parameters as desired. Click [Perform normalization, feature selection & scaling]. The "Normalization, feature selection, and scaling completed" window will appear when normalization is complete. Click [Ok].

   A minus vs. average (MA) plot showing variable features, with standard variance on the y-axis and average expression on x-axis, will appear. When finished with the plot, click [Next: Perform PCA Analysis] on the bottom right of the GUI.

6. *PCA Analysis*: Select the numbers of principal components and visualize results in the form of elbow plots, Viz Dim Loadings plots, and heatmaps.

   Choose the desired number of principal components. Click [Run PCA]. The "PCA analysis is complete" window will appear when the analysis is complete. Click [Ok].

   The different plot types generated after the analysis can be viewed by clicking on the tabs at the top of the window. When finished with the plots, click [Next: Perform Clustering] on the bottom right of the GUI.

7. *Clustering:* Adjust the number of dimensions and resolution for clustering and visualize cell clusters using UMAP or tSNE plots. Click [Cluster cells & Perform non-linear reduction]. The "Clustering & Non-linear reduction completed" window will appear. Click [Ok].

   After analysis, you can choose to visualize clustering with a UMAP plot or tSNE plot.

   - To choose a UMAP plot, select the "UMAP" radio button, set the desired parameters, and click [Apply UMAP]. The "UMAP completed" window will appear. Click [Ok] to view the plot.
   - To choose a tSNE plot, select the "tSNE" radio button, set the desired parameters, and click [Apply tSNE]. The "tSNE completed" window will appear. Click [Ok] to view the plot.

   When finished viewing the plots, click [Next: Perform Cell Type Annotation] on the bottom right of the GUI.

8. *Annotate cell types*: Upload your own reference dataset for annotation or use an existing reference to annotate cell types and visualize annotations using UMAP or tSNE plots.

   Select the "Upload reference" radio button to upload a reference data set in .rds format. To use a pre-existing reference from the celldex R package, select "Existing reference" and choose a pre-existing dataset from the dropdown menu.

   Click [Perform cell-type annotation]. The "Cell-type annotation complete" window will appear. Click [Ok].

   A cell annotation UMAP or tSNE plot will appear with the option to label clusters based on the reference dataset. When finished viewing the plot, click [Next: Perform Custom Lasso Selection] on the bottom right of the GUI.

9. *Custom Cell Selection*: Choose or define custom cell clusters using the UMAP or tSNE plot created in the previous step.

Select the portion of the cells on the UMAP or tSNE plot you wish to define as a custom selection using the lasso tool. When chosen, a "Save Lasso Selection" window will appear. Name and save the custom selection.

When finished, click [Next: Perform Differential Gene Expression Analysis].

10. *Differential Expression Analysis:* Choose how to perform differential expression analysis by identifying clusters, cell types, or custom clusters/cell types, and visualize differential expression using a volcano plot.

    Select the "Fine markers" option to compare specific clusters or the "Find all markers" option to identify markers across all clusters. After selecting the remaining desired parameters, click [Perform DE analysis]. The "DE analysis complete" window will appear. Click [Ok].

    To visualize the results using a volcano plot, select the desired volcano plot parameters, and click [Generate plots]. A volcano plot will appear on the screen. When finished, click [Next: perform pathway analysis] on the bottom right of the GUI.

11. *Pathway Analysis*: Choose how to perform pathway enrichment analysis using differential expression analysis results, and tabulate and visualize results.

    Select the enrichment method from the dropdown menu. Options include the GSEA approach, which includes pre-packages GMT files, or exploring functional insights through GO enrichment, WIkiPathways, and Reactome.

    After setting the desired parameters for gene selection, click [Perform Enrichment Analysis]. The "Enrichment Analysis complete" window will appear. Click [Ok].

    A plot and table of the results will appear. When finished, click [Generate additional plots] on the bottom right of the GUI.

12. *Additional Visualization*: Visualize specific features using feature plots, dot plots, violin plots, and ridge plots.

    Select one or more features from the dropdown menu and click [Feature plot]. The "Feature plot generated" window will appear. Click [Ok].

    View the feature plot and additional plot types by clicking on the tabs in the top of the window. When finished, click [Next: isoform analysis] on the bottom right of the GUI.

13. *Isoform Analysis*: analyze and cross-visualize differential isoform expression with regards to differential gene expression on UMAP or tSNE plots.

    Select the desired parameters for isoform analysis and click [Analyze isoform data]. After the analysis is complete, the option for overlay type and gene selection will appear. Choose the desired overlay type and select one or more genes from the dropdown menu. Click [Generate plots].

    When finished, click [Next: gene fusion analysis] on the bottom right of the GUI.

14. *Gene Fusion Analysis*: Overlay fusion genes on UMAP or tSNE plots.

    Choose the desired overlay type and select one or more fusion genes from the dropdown menu. Click [Overlay plot]. The plot will appear in the window.

    When finished, click [Next: clonotype analysis] on the bottom right of the GUI.

15. *Clonotype Analysis*: Overlay clonotypes on UMAP or tSNE plots.

    Choose the desired overlay type and select one or more clonotypes from the dropdown menu. Click [Overlay plot]. The plot will appear in the window.

    When finished, click [Next: download Seurat object] on the bottom right of the GUI.

16. *Processed Seurat Object Download*: Download the complete analysis report and the processed Seurat object.

    Click [Download processed scRNA-seq Seurat v5 object]. A File Explorer window will pop up and allow you to save the file.

Bulk RNA-seq analysis

1. Click [Launch bulkRNA app] in the GUI, then click [Launch scRNA app] in the GUI. A new window will appear. Click [Analysis Mode].

   NOTE: there is no "discovery mode" option for bulk RNA-seq analysis.

2. *Select Input Data*: Upload the required input files by selecting one of the options from the radio buttons.

   NOTE: For the "Raw counts matrix" and "Processed data for CogentAP" upload options, a sample metadata file where conditions for each sample are defined, is required.

   Click [Upload Raw Data]. When complete, the "Data Upload Successful" window will appear. Click [Ok] and then [Next: Check Data Quality] in the bottom right of the GUI.

3. *QC Visualization*: Check data quality by visualizing data using a PCA plot, a sample distance plot, or an MA plot.

   Use the radio buttons to select the desired plot for visualization. Click [Generate Plot]. The plot you selected will appear.

   When finished, click [Next: Perform Differential Expression] in the bottom right of the GUI.

4. *DE Analysis*: Perform differential expression analysis based on information from uploaded metadata or manually-selected samples.

   Use the radio buttons to select how to analyze the data for differential expression and set the remaining parameters as desired. Click [Perform DE Analysis]. The "DE Analysis is complete" window will appear. Click [Ok] to reveal a table of differentially-expressed transcripts or genes.

   When finished, click [Next: Visualization].

5. *Differential Expression Visualization*: visualize the results of differential expression analysis using volcano plots, heatmaps, or clustered by biological pathway.

   Use the radio buttons to select the desired plot for visualization. Click [Generate Plot]. The plot you selected will appear.

   When finished, click [Go to Analysis Report] in the bottom right of the GUI.

6. *Analysis Report*: Download the bulk RNA-seq analysis report.

   Click [Download bulkRNA Analysis Report]. The "Analysis Report Downloaded" window will appear when the download is complete. Click [Ok].

Contact Us	
Customer Service/Ordering	**Technical Support**
tel: 800.662.2566 (toll-free)	tel: 800.662.2566 (toll-free)
fax: 800.424.1350 (toll-free)	fax: 800.424.1350 (toll-free)
web: takarabio.com/service	web: takarabio.com/support
e-mail: ordersUS@takarabio.com	e-mail: technical_support@takarabio.com

Takara Bio USA, Inc.

United States/Canada: +1.800.662.2566 • Asia Pacific: +1.650.919.7300 • Europe: +33.(0)1.3904.6880 • Japan: +81.(0)77.565.6999

This document has been reviewed and approved by the Quality Department.

10.24 US