Genomic Data Explorer User Guide (07-10-20)

Contact:
General: Mingchu Xu, MXu3@mdanderson.org
DNA module: Xiaogang (Sean) Wu, XWu10@mdanderson.org
RNA module: Mingchu Xu, MXu3@mdanderson.org

URL: http://doprlrshinyweb.mdanderson.edu:8080/
Login username: ICON
Login password: shinyicon

If the login process is too slow, someone else are probably running the apps. Please be patient and try at a different time. The R Shiny apps are running on a test virtual machine (VM) with limited capability. Infrastructure upgrades are planned.

After logging in, click one of the apps to start. It usually takes 10-20s to start the app.

DNA Explorer

1. Choose your data source:
   Click one of the three tabs: "Example", "Upload", and "Select". (Currently, "Upload" is not supported)
   There are three types of files – Mutation Annotation Format (MAF), Copy Number Variation (CNV), and Meta Data in "Select":
   MAF Data: ICON_108samples_mutect_vaf02_pindel_vaf05.maf
   ICON_108samples_mutect_vaf02_pindel_vaf05_dropRepeats.maf
   CNV Data: ICON_108samples_exomecn.txt
   Meta Data: metadata.csv
   In "Example", the following files are used:
   MAF Data: ICON_108samples_mutect_vaf02_pindel_vaf05.maf
   CNV Data: ICON_108samples_exomecn.txt
   Meta Data: metadata.csv

2. Click "Launch" button. Adjust the parameters, and then click "Plot" at the left bottom. The app will start plotting figures for you.

3. Parameters:
   a. Oncoplot Options:
      Choose to display clinical features, TiTv, CNV, sample names, or whether sort by the annotation.
   b. Select Clinical Features:
      Choose which annotations to display on the bottom of oncoplot/heatmap.
      - *For example, if you choose Clinical Features and select Sex, choose Sort by Annotation, and click "Plot", you can view the mutation landscape (Oncoplot) by comparing male and female groups.*

c. Gene List:
Three ways to define genes: 1. Use Top Genes: use the most mutated genes (ranked by mutation frequency by maftools). Use the slide bar to adjust top n mutated genes, and click "Plot"; 2. Manual Input: type your favorite genes, split by semicolon or comma, and click "Plot"; 3. Upload File: upload your gene list, the file should be plain text file with one gene per row.

d. Samples List:
Three ways to define samples: 1. Use Top Samples: use the most mutated samples (ranked by mutation frequency by maftools). Use the slide bar to adjust top n mutated samples, and click "Plot"; 2. Manual Input: type your interested sample names, split by semicolon or comma, and click "Plot"; 3. Upload File: upload your sample list, the file should be plain text file with one sample per row.
   - *For example, if you upload a list of female patients, you can view the mutation landscape (Oncoplot) only for the female group.*

e. Plot Height, Plot Width:
Adjust the plot size in pixels (px).

f. Pathway Names:
Applies to Oncogenic Signaling Pathways. Choose specific pathway name to display pathway-gene association

g. VAF Column Name:
Applies to VAF Plot. Specify which column name in MAF file should be used for tumor (i.e., somatic mutation) variant allele frequency (VAF).

RNA Explorer

1. Choose your data source:
Click one of the three tabs: "Example", "Upload", and "Select". (Currently, "Upload" is not supported)
There are two datasets in "Select":
ICON dataset: "icon_rnaseq.rds"; Example dataset: "rnaseq.rds".

2. Click "Launch" button. The app will start plotting figures for you. (The first click will take the app 10-15s to load the libraries)

3. Parameters:
   a. Variable:
   1. Adjust which variable to label the rows and columns of Heatmap; 2. Adjust which variable to color in PCA plot; 3. Adjust which variable to compare in Gene Boxplot.

   b. Categorical Palette, Continuous Palette:
   Adjust the palette used for categorical or continuous variables. Can be used in Heatmap, PCA, Gene Boxplot, and Gene Clustering. Please see the last page for palette names and colors.

c. Reverse Scale Color Direction:
   Only applies to continuous palettes.

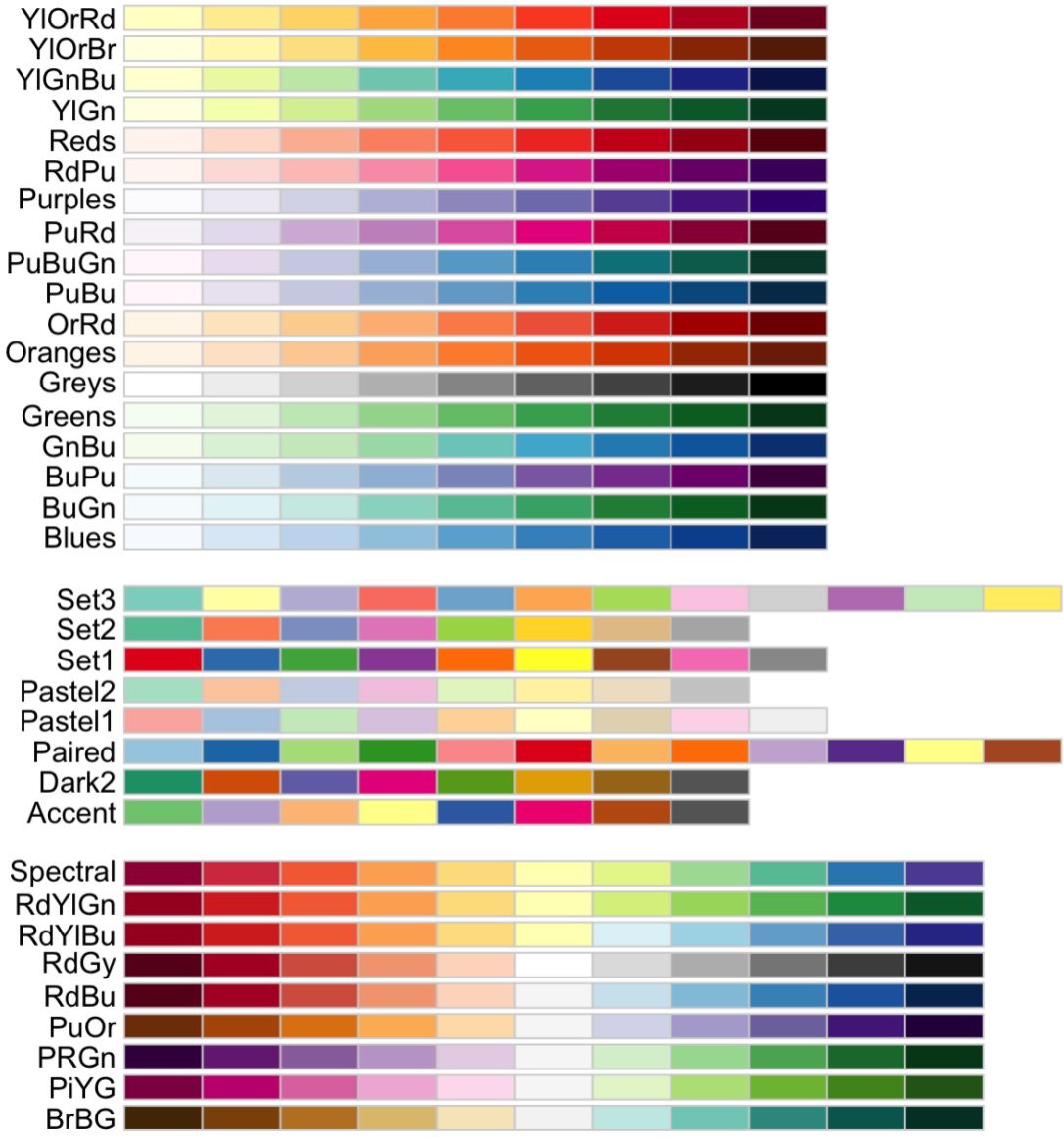
d. Plot Height, Plot Width:
   Adjust the plot size in pixels (px). Maximum 4000px.

e. Adjusted P-value Cutoff, Log2 Fold Change cutoff:
   Applies to MA Plot, Volcano Plot, and Table.

f. Adjusted P-value Squash, Log2 Fold Change Squash:
   Applies to MA Plot, Volcano Plot. Used to squash outlier points to acceptable ranges.

g. Gene List:
   Three ways to define genes: 1. Use Top Genes: use the most significantly differentially expressed genes (ranked by p-values). Use the slide bar to adjust top n genes, and click "Plot"; 2. Manual Input: type your favorite genes, split by semicolon or comma, and click "Plot"; 3. Upload File: upload your gene list, the file should be plain text file with one gene per row.

h. Pathway List:
   Two ways to define pathways: 1. Use Hallmarks: use 50 hallmark gene sets defined in this link: https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp
   2. Upload File: upload your pathway list, the file should be two-column plain text file with the first column as pathway names and the second row as gene names.


YlOrRd
YlOrBr
YlGnBu
YlGn
Reds
RdPu
Purples
PuRd
PuBuGn
PuBu
OrRd
Oranges
Greys
Greens
GnBu
BuPu
BuGn
Blues
Set3
Set2
Set1
Pastel2
Pastel1
Paired
Dark2
Accent
Spectral
RdYlGn
RdYlBu
RdGy
RdBu
PuOr
PRGn
PiYG
BrBG