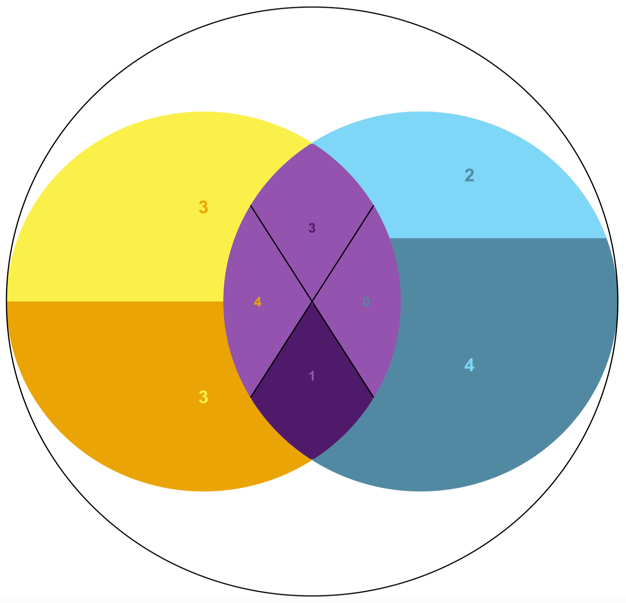
**How to use Kaleidoscope** **R script**

*Background:*

The Kaleidoscope diagram is a modified Venn diagram designed to compare the directionality of shared and unique sets of dysregulated genes from two different RNAseq analyses related to the same species. While a traditional Venn diagram can depict the number of genes that are shared or unique between two RNAseq analyses, the Kaleidoscope diagram takes things one step further by dividing these groups into up/downregulated categories. In the outer circles representing the dysregulated genes unique to each group, the number of upregulated genes is labeled in the top lighter regions while the number of downregulated genes is labeled in the bottom darker regions. The relative ratio of each is also represented by a horizontal midline that scales vertically along the height of the circle’s diameter. The overlapping region representing the dysregulated genes shared between the two groups is divided into four quadrants; the bottom darker quadrant representing the genes that are downregulated in both groups, the top lighter quadrant representing the genes that are upregulated in both groups, the left light quadrant with text sharing the same color as the left group representing genes that are upregulated in the left group but downregulated in the right group, and the right light quadrant with text sharing the same color as the right group representing genes that are upregulated in the right group but downregulated in the left group.

**Formatting data**

* Unzip the compressed Kaleidoscope\_v0.1.3 file
* You will need to compile your data from two separate RNAseq analyses pertaining to the same organism into two separate tables.
* These tables should each have 3 columns, “Treatment”, “Gene”, and “Change”.
  + Note: If you change the names of these headers, the Rscript will have to be adjusted accordingly.
* “Treatment” corresponds to the name for you study and should be the same for every entry in your file.
* “Gene” corresponds to the Gene ID with significant dysregulation.
* “Change” should either be “Up” or “Down” indicating the direction of dysregulation.
  + Note: this can be done quickly in excel by ordering the data by log2fold change and assigning the names Up/Down to all positive/negative values from the prospective of your treatment group.

*E.g.,*  **Study1 Study2**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Treatment** | **Gene** | **Change** |  | **Treatment** | **Gene** | **Change** |
| Study1 | Gene\_00001 | Up |  | Study2 | Gene\_00001 | Up |
| Study1 | Gene\_00002 | Down |  | Study2 | Gene\_00002 | Up |
| Study1 | Gene\_00003 | Down |  | Study2 | Gene\_00003 | Down |
| Study1 | Gene\_00004 | Up |  | Study2 | Gene\_00004 | Down |

* Save these documents as separate .csv files in the unzipped Kaleidoscope folder.

**Running the script**

Before generating the Kaleidoscope diagram, you must first input your file information into Kaleidoscope\_script.R.

* Open the Kaleidoscope\_script.R file using a text editing program, e.g., Notepad (windows) or TextEdit (mac)
* Enter the names of your .csv files here:

left <- read.csv("Study1.csv")

right <- read.csv("Study2.csv")

* Save the changes and open the edited Kaleidoscope\_script.R file with RStudio
* In the top left quadrant of RStudio, click the button.
* If the data files are set up properly, the Rscript should run automatically and produce a Kaleidoscope diagram comparing the directionality of expression between your two studies.
  + Note: The script will not run properly if there is not at least 1 unique gene ID for each group and at least 1 shared gene ID.

**User inputs**

The Kaleidoscope\_script.R file was designed to allow for customization of text, colors, and overlapping regions of the Venn diagram:

*Color*

* The default colors settings are based on a color-blind-friendly pallet for people with protanopia, deuteranopia, and tritanopia, but can be modified by changing the color hex codes corresponding to each region of the diagram:

##Right (r), Left (l), and Middle (m) circle colors

r\_col\_upreg = "#7FD7F7" *Top “upregulated” region of the right circle*

r\_col\_downreg = "#528AA1" *Bottom “downregulated” region of the right circle*

l\_col\_upreg = "#FBEF49" *Top “upregulated” region of the left circle*

l\_col\_downreg = "#EAA407" *Bottom “downregulated” region of the left circle*

m\_col\_upreg = "#9353AF" *Top “upregulated” regions of the center circle*

m\_col\_downreg = "#501b6b" *Bottom “downregulated” region of the center circle*

* + Note: you can design your own color pallets and test their visibility for the main three types of colorblindness using this online tool: <https://davidmathlogic.com/colorblind/#%23000000-%23EAA407-%23528AA1-%237FD7F7-%23FBEF49-%23501B6B-%23EC44BF-%23000000>

A picture containing circle, colorfulness, graphics, diagram

Description automatically generated

*Overlap*

* The degree of overlap between the two circles can be modified by adjusting the “nudge” value:

##Circle arrangement variables. default "nudge" is -0.1

nudge = -.1

A picture containing circle, colorfulness, diagram, graphics

Description automatically generatedA picture containing circle, colorfulness, graphics, diagram

Description automatically generated

nudge = 0.4 nudge = -0.4

*Formatting text*

* A picture containing circle, colorfulness, graphics, diagram

  Description automatically generatedAs can be seen in the previous example for nudge = -0.4, the text may not always appear where you would like or may appear too small depending on your datasets. The position of the text can be adjusted by manipulating the x and y coordinate values under ##text padding:

##text padding default is 0

r\_x\_padding = -.5

r\_y\_padding = 0

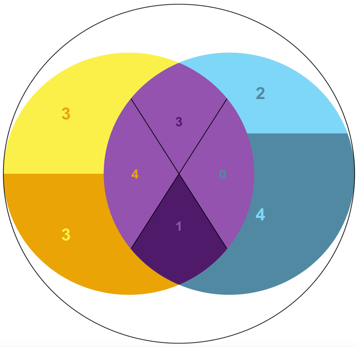
l\_x\_padding = 0

l\_y\_padding = 0

m\_h\_padding = .1

m\_v\_padding = -.05

* The size of the text can also be changed under ##scaling



##scaling text default is 11. units are "pts" like in microsoft word.

m\_b\_text = 12

m\_r\_text = 12

m\_l\_text = 12

m\_t\_text = 12

l\_t\_text = 16

l\_b\_text = 16

r\_t\_text = 16

r\_b\_text = 16

* If you would like to remove the text, simply change “TRUE” to “FALSE” under ##draw text. This will remove the text from the diagram and display the numbers for each region in the RStudio console instead.

A picture containing circle, colorfulness, graphics, diagram

Description automatically generated

##draw text. Either TRUE or FALSE

draw\_text <- TRUE

Output

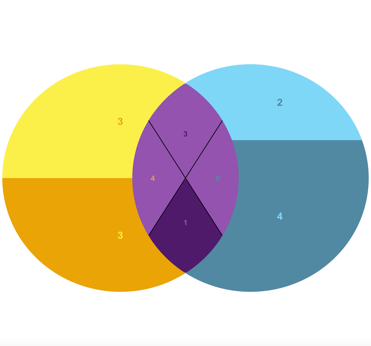
[1] "NUMBERS FOR LABELS"

[1] "Left Upregulated: 3"

[1] "Left Downregulated: 3"

[1] "Right Upregulated: 2"

[1] "Right Downregulated: 4"

[1] "Shared, Left quadrant Upregulated: 4"

[1] "Shared, Right quadrant Upregulated: 0"

[1] "Shared, Top quadrant Upregulated: 3"

[1] "Shared, Bottom quadrant Downregulated: 1"

*Other options*

* If you would like to remove the outer circle, simply change “TRUE” to “FALSE” under ##draw text.

A picture containing circle, colorfulness, graphics, diagram

Description automatically generated##draw encompassing circle. Either TRUE or FALSE

draw\_e\_circle <- TRUE

* You can change the angle of the cross section in center overlapping region with regards to the top quadrant by adjusting the ##angle.

##angle of lines for circle center; default is 45

line\_angle = 45