

Ball and stick graph for Clade models



Status

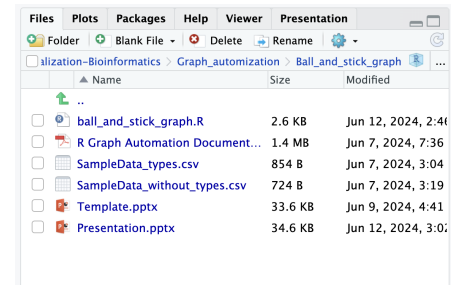
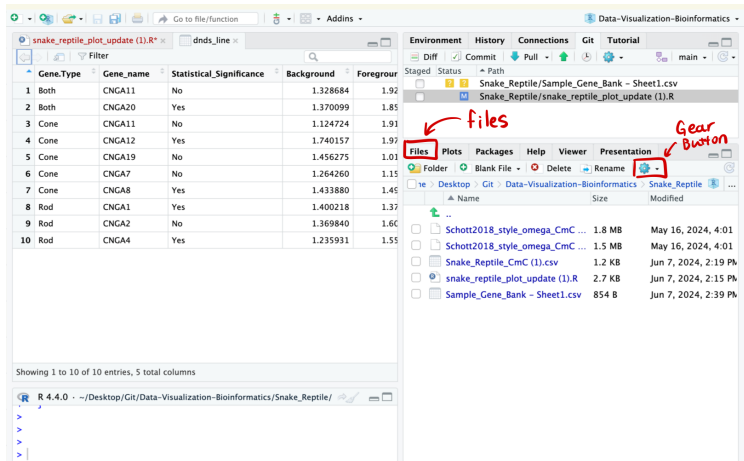
Done



This R script takes dN/dS values from PAML's clade model output, and graphs it in a ball and stick format. It can also group different genes associated with cellular pathways, or organismal ecologies, etc. Statistical significance is determined by comparison with a null model (M2a_rel, $p < 0.05$), and graphed as a dark grey; non significant as a light grey.

1. R settings

- a. Click the files button on the right hand side of the screen
- b. Find the Gears button just below the files bottom on the right side of the screen
- c. Click "Set As Working Directory"
 - i. Otherwise R will show the error "cannot open file 'SampleData_without_types.csv': No such file or directory"
- d. Open the R script
- e. Change line 16 to the name of the file that contains the data you want to generate a graph from.
 - i. Note: You can find the place to change the file name near "read csv" function at the beginning of the file
- f. Use Command+Shift+Enter (or Control+Shift+Enter on Windows) to run the entire file
- g. Then you should find a file named "Presentation.pptx" under files, which is a power point presentation that you can edit.



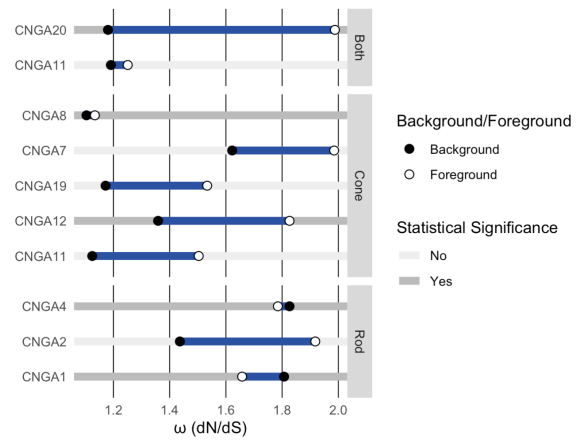
This is what will happen after you run the code. Note the new file "Presentation.pptx"

2. CSV/Excel Formatting

- Columns (Ordered doesn't matter, R is just picky about names):
Gene_type, Gene_name, Background_Foreground, dNdS, Statistical_Significance
- Gene Type: Please enter the name of the string, or enter "na"
- Gene_name: Please enter the name of the gene
- Background_Foreground: Please enter either "Background" or "Foreground"
- Statistical_Significance: Please enter either "Yes" or "No"
- dNdS: Numeric Value

3. Sample data with Gene Types

Gene Type	Gene_name	dNdS	Statistical_Significance	Background_Foreground
Rod	CNGA1	1.420898271	Yes	Background
Rod	CNGA1	1.399120233	Yes	Foreground
Rod	CNGA2	1.798042993	No	Background
Rod	CNGA2	1.287029638	No	Foreground
Rod	CNGA4	1.538076909	Yes	Background
Rod	CNGA4	1.736751058	Yes	Foreground
Cone	CNGA7	1.872901679	No	Background
Cone	CNGA7	1.862169926	No	Foreground
Cone	CNGA8	1.311531231	Yes	Background
Cone	CNGA8	1.0391834	Yes	Foreground
Cone	CNGA11	1.869341625	No	Background
Cone	CNGA11	1.010740569	No	Foreground
Cone	CNGA12	1.201309048	Yes	Background
Cone	CNGA12	1.089178849	Yes	Foreground
Cone	CNGA19	1.748352827	No	Background
Cone	CNGA19	1.752109061	No	Foreground
Both	CNGA20	1.90966686	Yes	Background
Both	CNGA20	1.466176017	Yes	Foreground
Both	CNGA11	1.222447703	No	Background
Both	CNGA11	1.691968101	No	Foreground



4. Sample data without Gene Types

Gene Type	Gene_name	dNdS	Statistical_Significance	Background_Foreground
na	CNGA1	1.672654992	Yes	Background
na	CNGA1	1.712828314	Yes	Foreground
na	CNGA2	1.063048757	No	Background
na	CNGA2	1.302023844	No	Foreground
na	CNGA4	1.401174259	Yes	Background
na	CNGA4	1.068695445	Yes	Foreground
na	CNGA7	1.740858486	No	Background
na	CNGA7	1.160488964	No	Foreground
na	CNGA8	1.492000778	Yes	Background
na	CNGA8	1.291456166	Yes	Foreground
na	CNGA11	1.258751012	No	Background
na	CNGA11	1.406918501	No	Foreground
na	CNGA12	1.293412095	Yes	Background
na	CNGA12	1.472288487	Yes	Foreground
na	CNGA19	1.265750147	No	Background
na	CNGA19	1.542183529	No	Foreground
na	CNGA20	1.48079509	Yes	Background
na	CNGA20	1.275036697	Yes	Foreground
na	CNGA11	1.314555225	No	Background
na	CNGA11	1.907426764	No	Foreground

