Hands-on protocol to create heatmap with Test data

1. Download the Heatmap_Test_data.zip via the link from GitHub (https://github.com/zx0223winner/HSDFinder).

zx0223winner Add files via upload	0409
🗅 Example.fasta	Add files via upload
HSDFinder.py	Add files via upload
B HSDFinder_workflow.png	Add files via upload
HSD_to_KEGG.py	Add files via upload
Heatmap_Test_data.zip	Add files via upload
KO database and its category.keg	Add files via upload
NoBadWordsCombiner.py	Add files via upload
Readme.txt	Update Readme.txt

- 2. We provided the data from eight algae species (*Chlamydomonas* sp. UWO241, *Chlamydomonas* sp. ICE-L, *Chlamydomonas reinhardtii*, Chlamydomonas *eustigma*, *Gonium pectorale*, *Dunaliella salina*, *Volvox carteri* and *Fragilariopsis cylindrus*), user can create the heatmap by selecting some of them.
 - Each folder represents one species, there are two files in each folder. Take Chlamy genome as example, The file named "HSD_Chlamy_90pct_10aa.txt" is the result of highly similar duplicates (HSDs) in Chlamy genome with the filtering option more than 90% amino acid pairwise identity and within 10 amino acid difference. (Check the tutorial for more details about how to acquire this file on your own species)

Back/Forward	Tutorial Creating heat	tmaps	□□ ✔ View	Group	ப Share	∽ Edit Tags	₩¥ Action
UWO241	ICE-L	Chlamy		Ceustigma			
Fragilariops	sis Volvox	Dsalina		Gonium			

- Another file named "Chlamy_KO_annotation.txt" is retrieved from the KEGG database documented the correlation of KO accession with each gene model identifier.
- User can upload the respective files to the web server to create the Heatmap (Check the screenshot below). Make sure you have an organism name for the files you chose to upload.

। Creat Hea	each species fits	the options.	
HSD File 😧 HSD_File_exan Choose File	nple.txt HSD_UWO241t_10aatxt	Gene list with KO annotation @ Genelist_KO_annotation_example.txt Choose File_UWO241_KO_aotation.txt	Organism name 🕢 UWO241
HSD File		Gene list with KO annotation	Organism name
Choose File	HSD_ICE-L_90pct_10aa.txt	Choose File ICE-L_KO_annotation.txt	ICE-L
HSD File		Gene list with KO annotation	Organism name
Choose File	HSD_Chlamy_9ct_10aa.txt	Choose File Chlamy_KO_annotation.txt	Chlamy
+ add specie	S		

 Once the user finished uploading the files, there is an option below to design the figure size at user's preference. The 'row' option can change the width of the heatmap image. The 'col' option can change the length of the heatmap image.



4. Tap the Create Heatmap button, the following gif. Will jump out. It usually takes around less than one minute to run with three to five species.



5. Once heatmap created, the HSDs numbers for each species will be displayed in a heatmap under different KEGG function category. On the left side, the color bar indicates a broad category of HSDs who have pathway function matches, such as carbohydrate metabolism, energy metabolism, translation etc. The color for the matrix indicates the number of HSDs across species.



6. User can also find the option below the image where to download the high-resolution version and the text file for user to use in future analysis.



7. Just in case you are not satisfied with the heatmap figure size. E.g., the image texts are overlapped. You can always rerun with more appropriate 'width and length' options.

Figure Size: rov	v 10	col	25
Creat H	leatma	p	

8. The new heatmap after changing the size of the image.

