

Visibiome Supplementary Figures

Syafiq Kamarul Azman, Andreas Henschel, Zohaib Anwar

July 2017

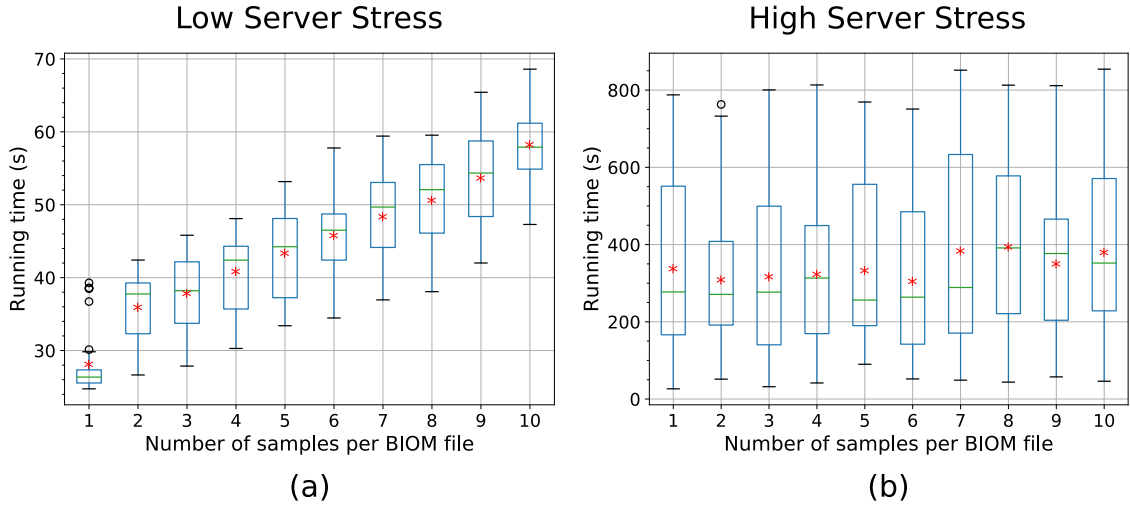


Figure 1: Barplot of Bray-Curtis analysis running time with different sample sizes in low and high stress situation on an EC2 `t2.medium` machine.

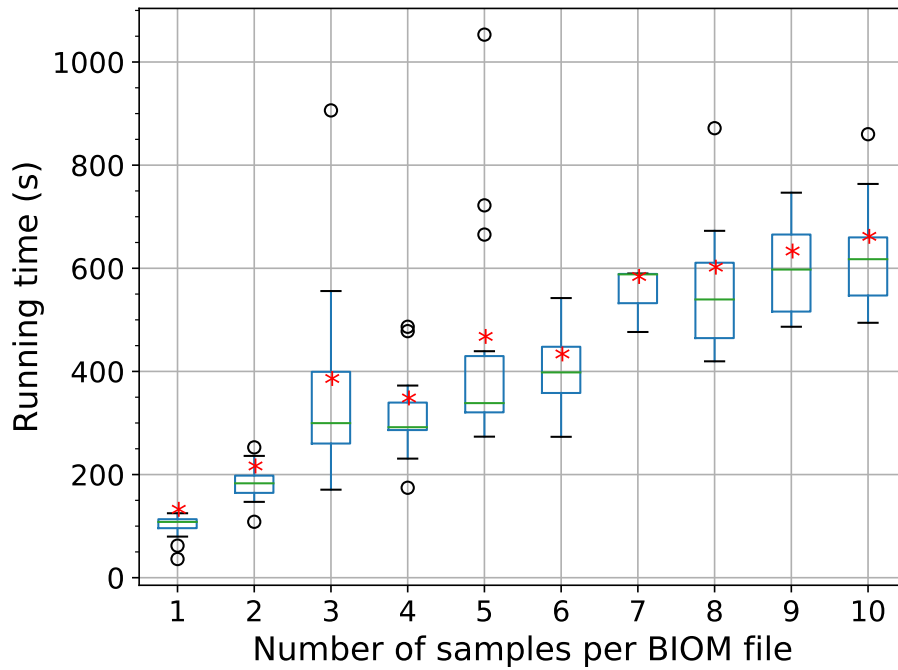


Figure 2: Barplot of GNAT/UniFrac running time with different sample sizes in low stress situation on EC2 `t2.medium` machine.

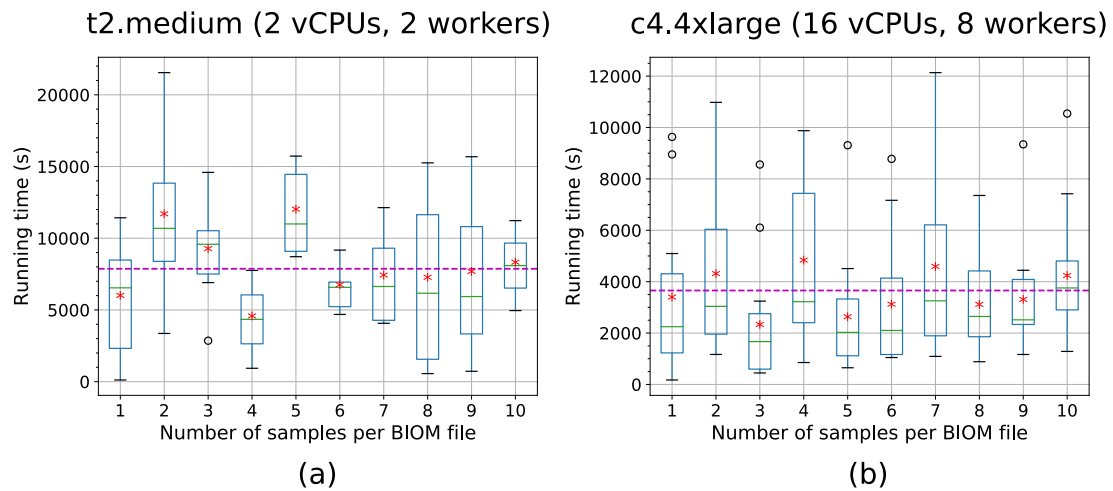


Figure 3: Barplot of GNAT/UniFrac running time with different sample sizes in high stress situations on EC2 t2.medium and c4.xlarge machines. The purple dotted line shows the average running time for all jobs.

Figure 4 is a screenshot of the VisiBiome (alpha release) web interface. The interface includes a navigation bar with links to Dashboard, New Search, Contact, Help, and Test Example. The main content area is titled "VisiBiome (alpha release)" and describes the tool as a webserver to visualize diversity against annotated 16S rRNA microbial profiles. It provides instructions on how to analyze 16S rRNA samples against a reference database.

The form includes the following fields and options:

- Name:** Unnamed Job
- Paste your OTU table, GreenGenes 13.5 OTUs only!** (Text input field)
- or upload your BIOM file, GreenGenes 13.5 OTUs only!** (Choose File button, No file chosen)
- Select the ecosystem(s):**
 - ☒ All
 - ☐ Animal/Human
 - ☐ Anthropogenic
 - ☐ Freshwater
 - ☐ Marine
 - ☐ Soil
 - ☐ Plant
 - ☐ Geothermal
 - ☐ Biofilm
- Analysis type:** GNAT/UniFrac
- Range query value:** 0.3
- I have normalized the 16S OTU copy numbers for this BIOM table:** ☐
- Perform adaptive rarefaction on my samples:** ☐ (More info link)
- Taxonomy ranks:**
 - ☒ Phylum
 - ☐ Class
 - ☐ Order
 - ☒ Family
 - ☒ Genus
- Analyze** (Green button)

Figure 4: Screenshot of the search form for VisiBiome with all combinations of settings for the search.

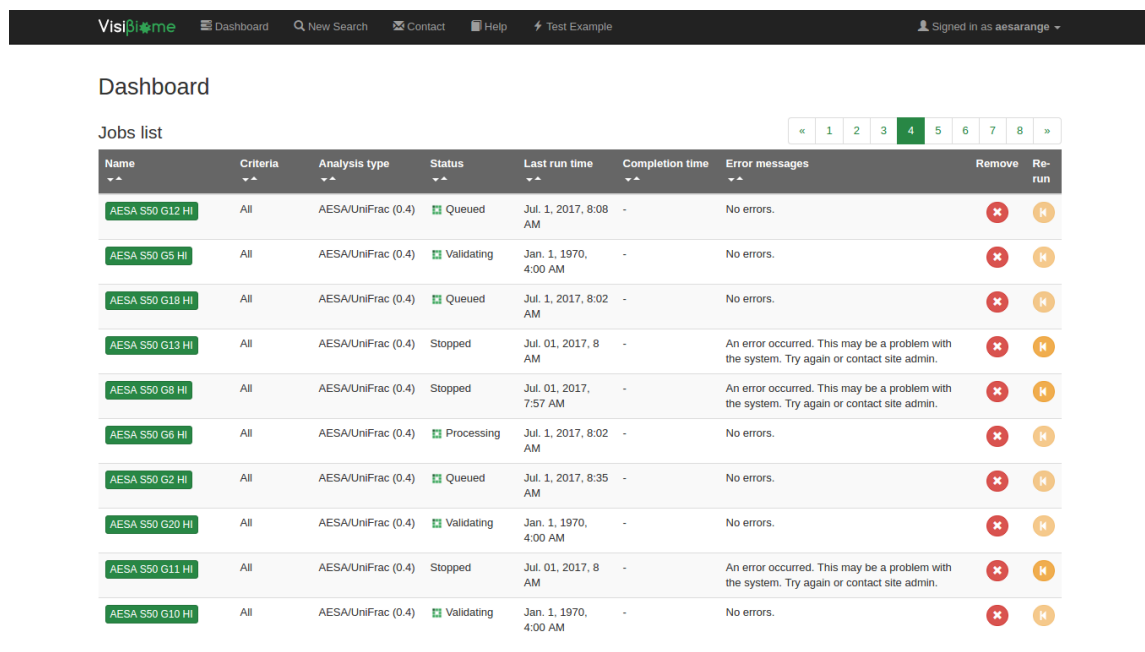


Figure 5: Screenshot of an example Dashboard for a user. The dashboard shows the running jobs and queued jobs as well as completed jobs and jobs that reached an error while processing. It features pagination, rerun and deleting.

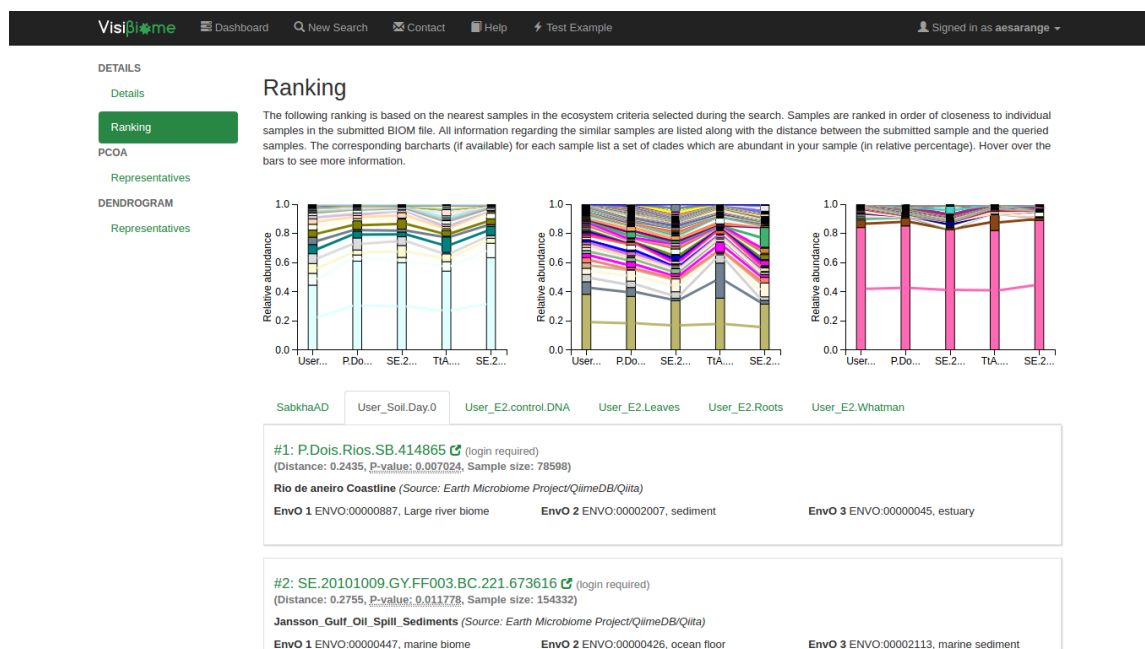


Figure 6: Screenshot of the barcharts visualization of a job. Barcharts are zoomable which enables users to delve into the culprits of the matches for their sample against the database samples.

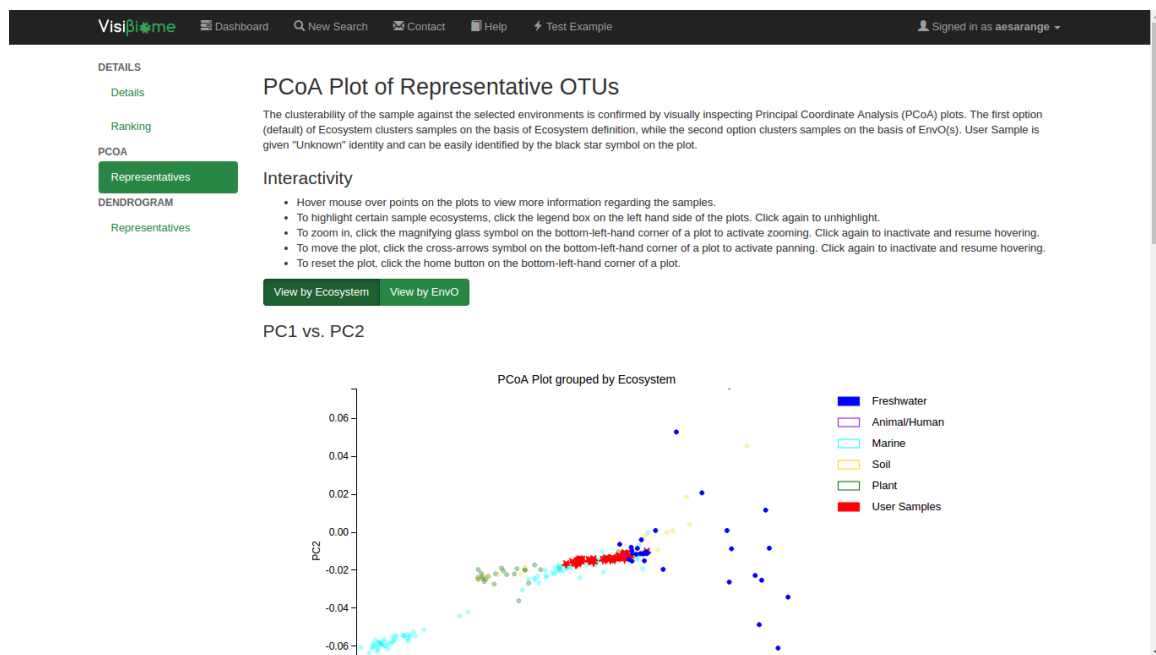


Figure 7: Screenshot of the PCoA visualization of a job. PCoA plots are 2-dimensional but the first three dimensions are enumerated. Plots are zoomable and hoverable to display more information for matched samples.

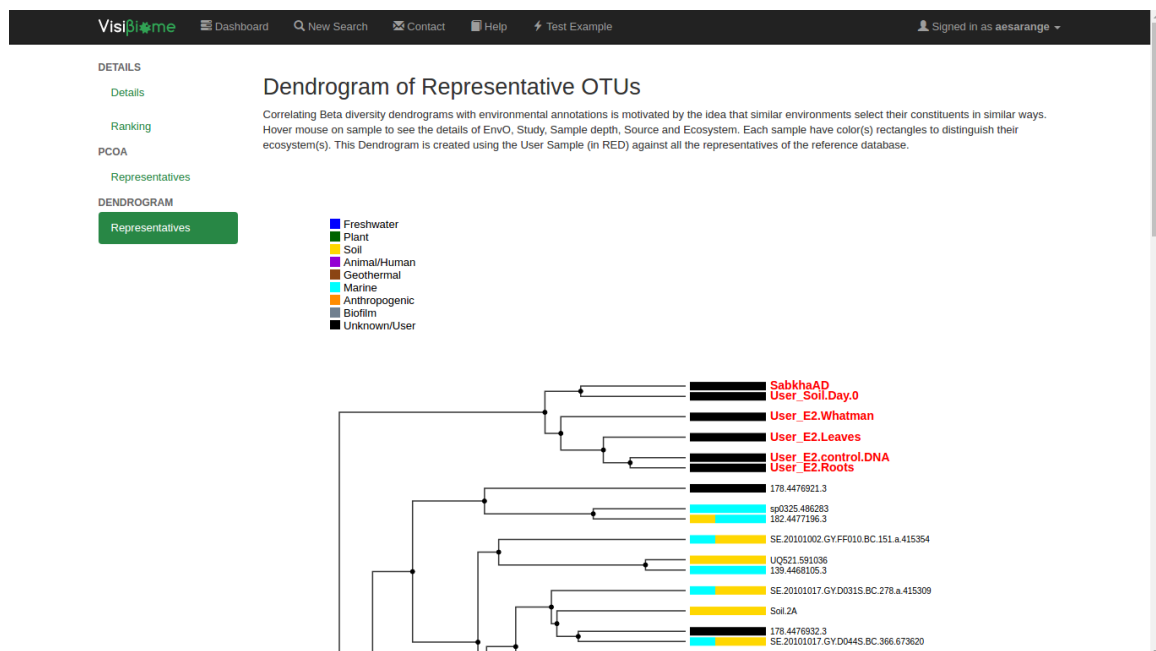


Figure 8: Screenshot of the dendrogram visualization of a job. Dendrogram plots show the relation between samples and the database sample's EnvO color. The samples can be hovered over to display more information.

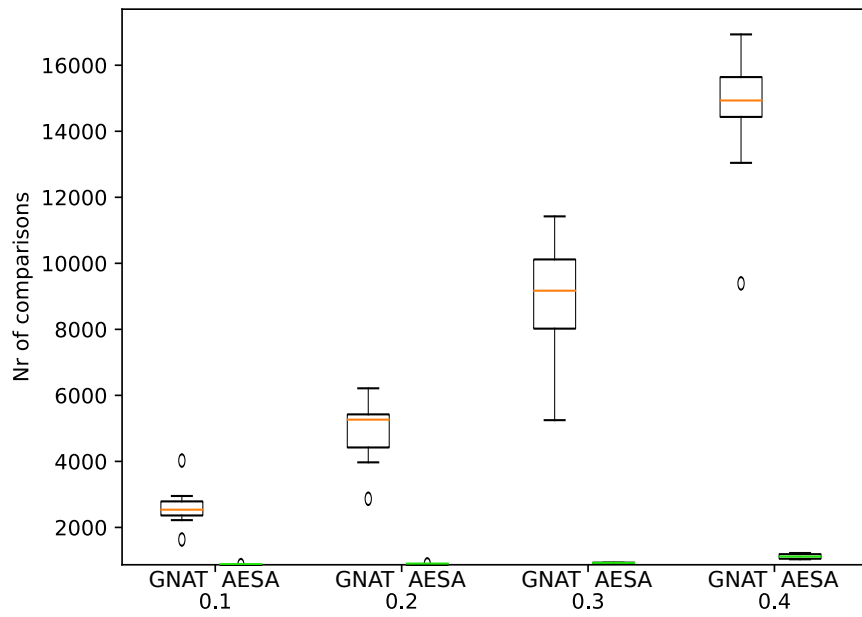


Figure 9: Barplot of the number of comparisons made for GNAT and AESA search on different ranges.

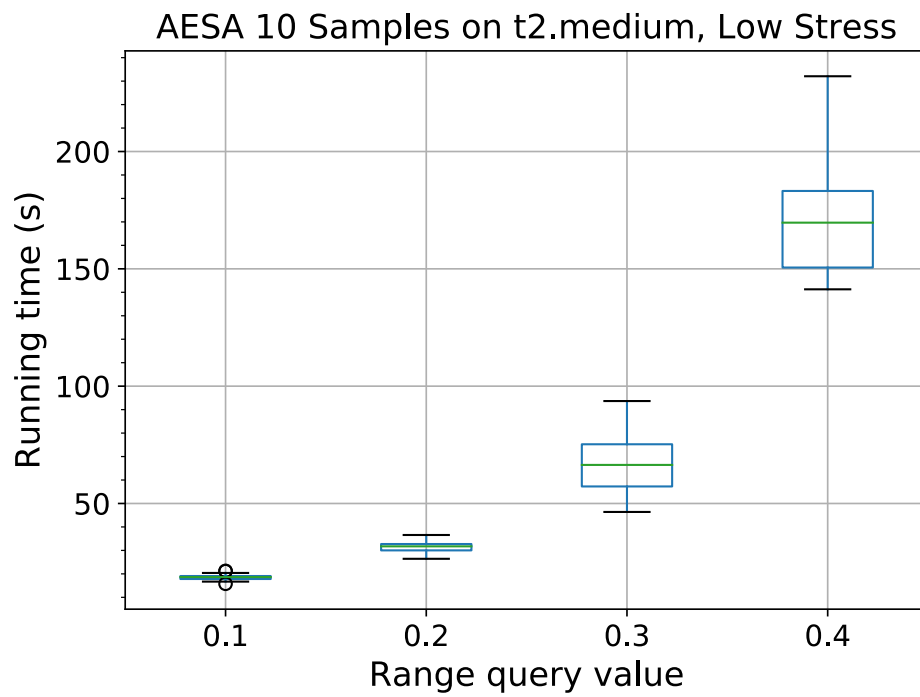


Figure 10: Barplot of AESA/UniFrac running time with different ranges in low stress situation on EC2 't2.medium' machine.

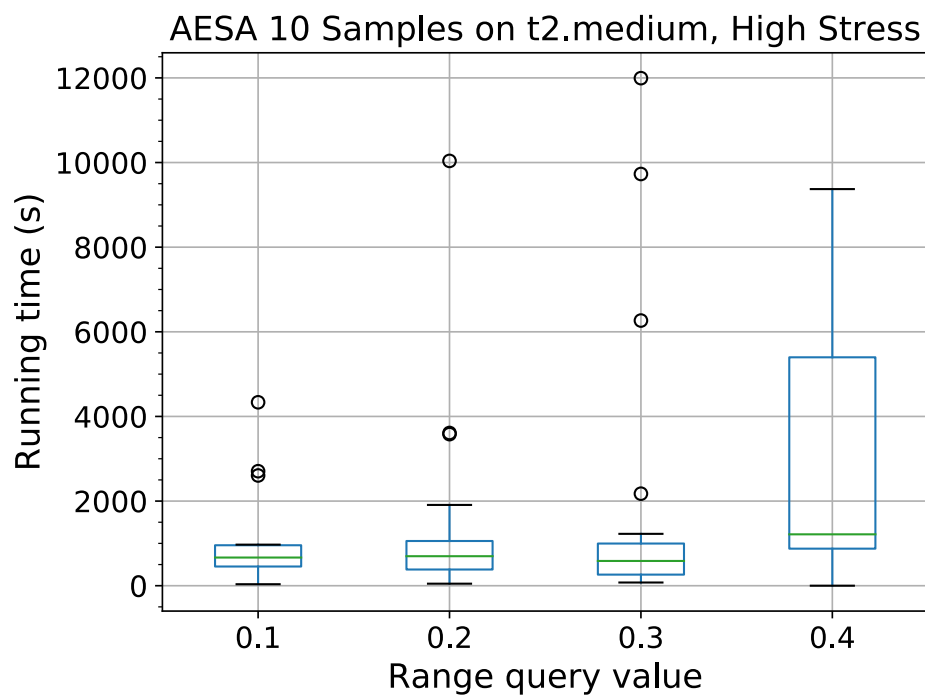


Figure 11: Barplot of AESA/UniFrac running time with different ranges in high stress situation on EC2 't2.medium' machine.