

User Manual for AmphoraVizu

**a chart visualization for metagenomic and genomic analysis tools
AMPHORA2 and AmphoraNet**

Csaba Kerepesi, Balázs Szalkai, Vince Grolmusz

[PIT Bioinformatics Group](#)

Last update: Apr 17, 2015.

Table of contents:

1. Introduction

- 1.1. About AmphoraVizu
- 1.2. How to Cite

2. AmphoraVizu → webservice version

- 2.1. Availability
- 2.2. Quick start
- 2.3. Options
- 2.3. Terms of Use
- 2.4. Privacy

3. AmphoraVizu → offline version

- 3.1. Availability
- 3.2. Quick start
- 3.3. Options
- 3.4. Terms of Use

4. Further notes

- 4.1. Be careful when you use "Archaea and Bacteria" option
- 4.2. Estimating relative abundance if you have few marker protein hits
- 4.3. AMPHORA2 marker protein databases are built at October, 2011.

1. Introduction

1.1. About AmphoraVizu

AmphoraVizu visualizes the outputs generated by the metagenomic and genomic analysis tool AMPHORA2 or its webserver implementation AmphoraNet.

1.2. How to Cite

If you publish anything using AmphoraVizu webserver or Amphoravizu.pl program, you are asked to cite the following publication, and you are also suggested to refer to the [link of our webservice](#):

Csaba Kerepesi, Balázs Szalkai, Vince Grolmusz. Visual Analysis of the Quantitative Composition of Metagenomic Communities: the AmphoraVizu Webserver, **Microbial Ecology**. 2014 Oct 10.

2. AmphoraVizu → webserver version

2.1. Availability

AmphoraVizu webserver is available at the page <http://pitgroup.org/amphoravizu/>.

2.2. Quick start

1. Go to the main page (<http://pitgroup.org/amphoravizu/>) and upload an AMPHORA2 or AmphoraNet output file (sample files are available here: http://www.pitgroup.org/static/amphora2_results/)
2. Select the lowest taxonomic rank which you are interested in.
3. Click the "Visualize!" button. You will see a chart immediately (if it is not too wide to display in miniview). Then click "Magnify" for the magnified chart.

2.3. Options

Main options:

- **AMPHORA2/AmphoraNet output:** Select an AMPHORA2/AmphoraNet output file: Click *Browse* button.
- **Lowest taxonomic rank:** Select the lowest taxonomic rank you are interested in: *Superkingdom/Phylum/Class/Order/Family/Genus/Species*. Default value: Phylum.
- **Chart type:**
 - *Column Chart* (default): shows the distribution of phlotypes for each marker gene.
 - *Pie Chart*: shows the consensus distribution (averages of the distributions for each phylotype)

Advanced options:

- **Display mode:**
 - *Proportion* (default): for each marker gene, the ratio of hits is displayed which correspond to the specific phlotype. That is, the total hit count is normalized to 1 for each marker gene.
 - *Amount*: the exact number of hits is displayed for each marker gene and phylotype.
- **Minimum confidence:** The minimum confidence of hits to consider Default value=0.1.
- **Minimum average to show:** Only those phlotypes will be displayed whose average share is at least the given value. Default value: 0.02.
- **Title of magnified chart:** The following parameter references can be used in the title: `$chart_type`, `$filename`, `$lowest_rank`, `$min_average`, `$min_confidence`. Default value=*Visualization of \$filename (lowest rank: \$lowest_rank, min. confidence: \$min_confidence, min. average: \$min_average)*

2.4. Terms of Use

You can use this service only if you accept the following terms: We do not guarantee anything about this service: We do not state anything about the usability of this service, and we do not state that the results that we may return can be used for any purpose. We cannot guarantee that this service will be available in the future, and we cannot guarantee that your query would generate any output at all.

2.4. Privacy

We will not give out your data to anyone.

3. AmphoraVizu.pl → offline version

3.1. Availability

The open-source version of the AmphoraVizu program is available for download at <http://pitgroup.org/apps/amphoravizu/AmphoraVizu.pl>.

3.2. Quick start

1. Open Source AmphoraVizu is a program written in Perl programming language. Perl is downloadable from <https://www.perl.org/get.html>
2. Use an AMPHORA2 or AmphoraNet output file as input file (sample files are available here: http://www.pitgroup.org/static/amphora2_results/)
3. A typical command:

```
perl AmphoraVizu.pl -input example_output.txt -min_confidence 0.1 -lowest_rank 2 -min_average 0.02 -chart_type column > out.html
```
4. Finally open the out.html file in a web browser.

3.3. Options

- **-input:** input file (AMPHORA2 output file), default: /dev/stdin
- **-min_confidence:** minimum confidence, default: 0.1
- **-lowest_rank:** lowest rank (1=Superkingdom, 7=Species), default: 2
- **-min_average:** minimum average to display, default: 0.02
- **-chart_title:** chart title (with optional HTML tags), to override default title
- **-chart_type:** "column": Column Chart output (default), "pie": Pie chart output with averages
- **-chart_width:** width of chart area, all CSS values are valid (default: 100%)
- **-chart_height:** height of chart area (default: 100%)
- **-chart_compact:** if 1, margins are decreased as much as possible (default: 0)
- **-html_inline:** if 1, tags like <html>, <head> and <body> are omitted (default: 0)
- **-miniview:** 1: miniview on, 0: miniview off (default: 0)
- **-sort:** "avg": sorted by averages (default), "abc" sorted by alphabet
- **-quantify:** "prop": shows the proportions of marker genes (default), "amounts": shows the exact amounts of marker genes

3.4. Terms of Use

AmphoraVizu is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS

FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details:
<<http://www.gnu.org/licenses/>>.

For any other inquiries send an Email to Csaba Kerepesi: kerepesi@caesar.elte.hu

4. Further notes

4.1. Be careful when using "Archaea and Bacteria" option

Be careful when using "Archaea and Bacteria" option of AmphoraNet, because the more archaeal marker protein (104 vs 31) causes overestimating of relative abundance of archaea.

4.2. Estimating relative abundance if you have few marker protein hits

Be careful estimating relative abundance if your results have few marker protein hits. Our tests on simulated metagenomes show it should be at least 100 marker protein hits per taxon for accurate relative abundance estimation. We are working on the exact statistical analysis on this interesting problem.

4.3. AMPHORA2 marker protein databases are built at October, 2011

AMPHORA2 marker protein databases are built at October, 2011 (according the hmm files: [dnaG.hmm](#) using 1324 bacterial genomes and 114 archaeal genomes. At this moment there are 180 archaeal 3272 bacterial genomes and 180 archaeal genomes (according to [EBI complete genomes](#)).