DSMZ

Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH

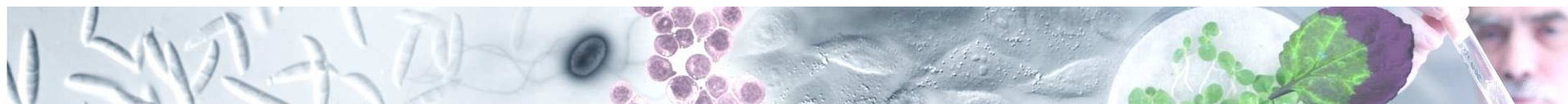# 'opm' - an R package for analysing OmniLog(R) Phenotype MicroArray data

Data import	Data management	Graphical analysis	Communication with other sofware	Results tabulation
☀ import of raw kinetic data from the OmniLog® software in CSV format	☀ many querying and subsetting options, which are	☀ for kinetic raw data (one to many plates):	☀ for applying other R packages:	☀ automated discretization (user-defined or by *k-means*) of curve heights into positive (+), weak (w) and negative (-)
☀ aggregation into curve parameters (lag phase, steepness of slope, maximum curve height, area under the curve)	☀ applicable, in part or in combination, to: ■ specific wells ■ specific time points ■ specific curve parameters	■ levelplots ■ xy-plots ☀ for (bootstrapped) curve parameters:	■ conversion into customized data frames ☀ for applying phylogeny software (TNT, RAxML, etc.):	☀ subsequent automated generation of reports in HTML for taxonomic journals such as IJSEM
☀ appending extensive organismic and/or experimental metadata (limited only by computer memory)	■ any type of organismic and/or experimental metadata	■ heatmaps ■ radial plots ■ confidence-interval plots	■ export of curve parameters in NEXUS, (extended) PHYLIP and Hennig86 format ☀ exchange with third-party software: ■ export in YAML format	

Application: 'opm' is designed to explore OmniLog® Phenotype MicroArray (PM) respiration kinetics into unlimited directions of analysis frameworks. The strength of 'opm' is the analysis of quantitative curve-parameter differences along with high resolutive graphical display. Therefore, 'opm' is highly suited for geneticists, biochemists or system biologists who wish to quantify significant differences among similar single or multiple curve kinetics. Still, the typical usage of OmniLog® PM data, the qualitative determination of positive, weak, or negative reactions, is facilitated in 'opm' by the automated discretization of curve heights and the generation of HTML reports, e.g., for taxonomic journals.

Functional Features: 'opm' imports raw kinetic data from numerous plates, aggregates them into the respective curve parameters and appends a nearly unlimited amount of metadata. All information can be stored in a single object (or exported file). Extensive data-management functions allow querying and subsetting according to all user-defined interests. Full or subset data, either raw kinetics or aggregated curve parameters, can be graphically displayed in different ways. User-friendly conversion tools allow the user to export raw kinetics, curve parameters and metadata for the full range of applications within R, with external phylogeny software, and principally any third-party software using the YAML format.
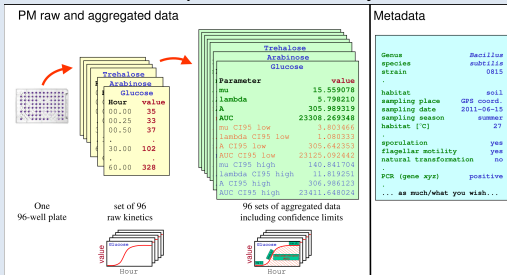
Technical Features: 'opm' is based on R, a popular, free software environment for statistical computing and graphics. 'opm' uses S4 objects, classes and methods. 'opm' utilizes functions from the R packages *boot*, *grofit*, *hwriter* and *lattice* as well as optionally *Ckmeans.1d.dp*, *gplots*, *multicore*, *optparse* and *yaml*. 'opm' is able to split time-consuming calculations over several cores on a multi-core machine.

More Information: http://www.dsmz.de/research/microorganisms/projects/analysis-of-omnilog-phenotype-microarray-data.html
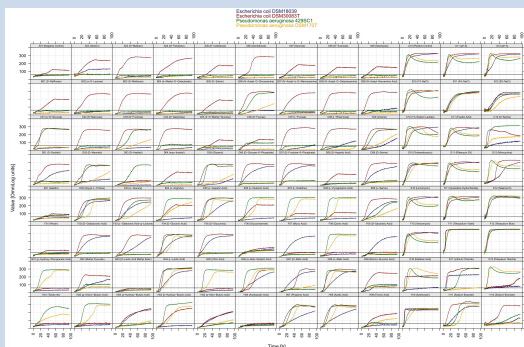
Some exemplary graphics generated by 'opm'

Data stored in 'opm' as OPMS objects

PM raw and aggregated data

Metadata

One 96-well plate

set of 96 raw kinetics

96 sets of aggregated data including confidence limits
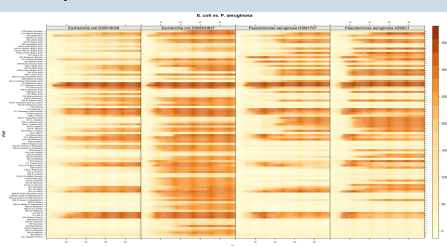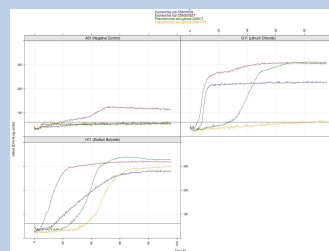
xy-plot of four strains in 12x8 layout

-> raw data display

levelplot of four strains

-> raw data display

three selected wells of four strains
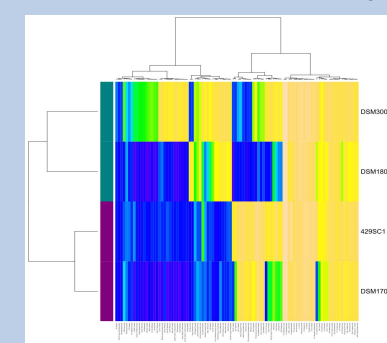
-> raw data display

heatmap of four strains
(curve parameter: maximum height)

-> overall similarity accross all curves

95% confidence interval of 5 curves (curve parameter: area under the curve)

A07 (Sucrose)

G11 (Lithium Chloride)
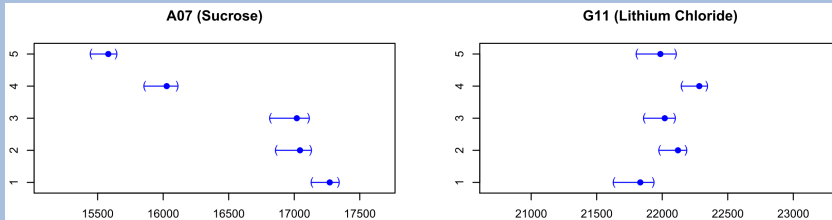
-> substantially different curves

-> curves largely identical