

LONDON  
SCHOOL of  
HYGIENE  
& TROPICAL  
MEDICINE



Vehkala, M; Shubin, M; Connor, TR; Thomson, NR; Corander, J  
(2015) Novel R pipeline for analyzing biolog phenotypic microar-  
ray data. PLoS One, 10 (3). e0118392. ISSN 1932-6203 DOI:  
10.1371/journal.pone.0118392

Downloaded from: <http://researchonline.lshtm.ac.uk/2131866/>

DOI: [10.1371/journal.pone.0118392](https://doi.org/10.1371/journal.pone.0118392)

#### Usage Guidelines

Please refer to usage guidelines at <http://researchonline.lshtm.ac.uk/policies.html> or alternatively contact [researchonline@lshtm.ac.uk](mailto:researchonline@lshtm.ac.uk).

Available under license: <http://creativecommons.org/licenses/by/2.5/>

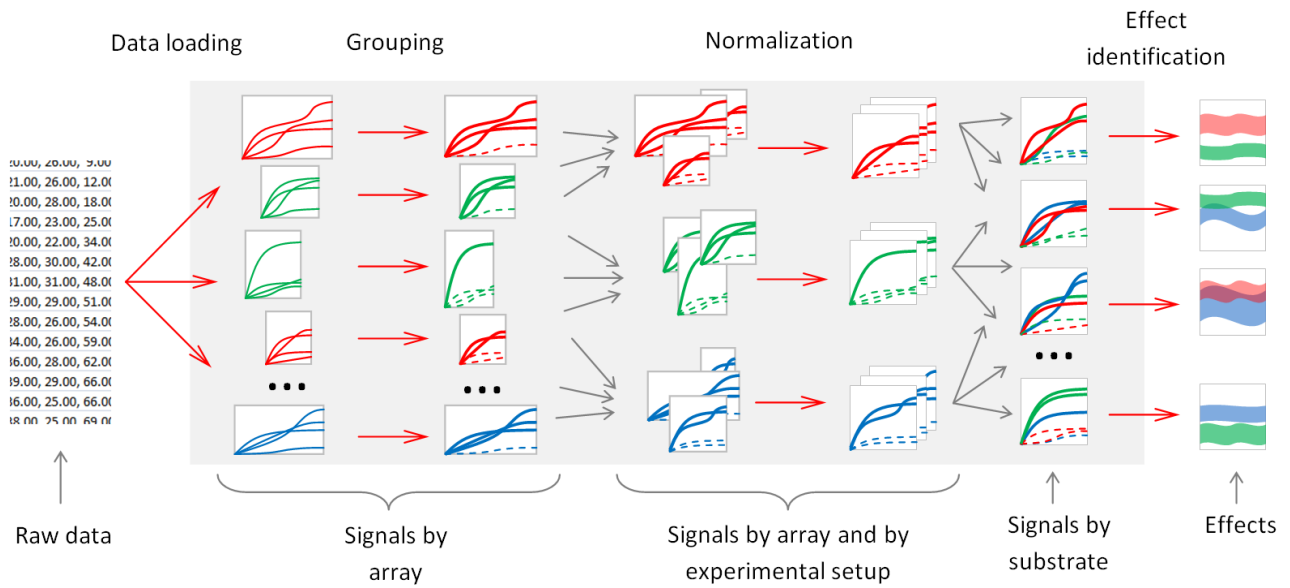


Figure S1. Graph of the analysis pipeline. The pipeline consists of three main steps: grouping, normalization and effect identification. Before these steps the raw numeric data is loaded into R workspace by using an R package called opm. Data plotted array-wise expose metabolic profiles over time. Different colours represent replicated measurements of an experimental setup. In the grouping step, the raw profiles are categorized by an EM algorithm into two groups which in the panels are illustrated by continuous and dashed lines. Next, the data is normalized. Normalization is performed for each experimental setup separately. Normalization makes the replicates of the same setup comparable with each other. Finally, effect identification reveals metabolic profiles differing between experimental conditions. This step is performed substrate-wise.