Analytica Chimica Acta 963 (2017) 1–16

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Combining ANOVA-PCA with POCHEMON to analyse micro-organism development in a polymicrobial environment☆

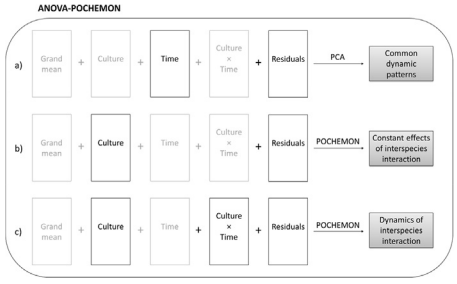
Brigitte P. Geurts [a], Anne H. Neerincx [b], Samuel Bertrand [c, d], Manja A.A.P. Leemans [a], Geert J. Postma [a], Jean-Luc Wolfender [d], Simona M. Cristescu [b], Lutgarde M.C. Buydens [a], Jeroen J. Jansen [a, *]

[a] Department of Analytical Chemistry, Institute for Molecules and Materials, Radboud University, Nijmegen, The Netherlands
[b] Department of Molecular and Laser Physics, Institute for Molecules and Materials, Radboud University, Nijmegen, The Netherlands
[c] Faculty of Pharmacy, University of Nantes, EA 2160-Mer Molécules Santé, 9 Rue Bias BP 53508, Nantes-cedex 1 44035, France
[d] School of Pharmaceutical Sciences, University of Geneva, University of Lausanne, CMU – Rue Michel Servet 1, 1211 Geneva 11, Switzerland

HIGHLIGHTS

- ANOVA-POCHEMON disentangles different information sources to study micro-organism development in a polymicrobial environment.
- It combines ANOVA with PCA of the isolated interspecies interaction-related chemistry in pathogen development.
- Two complementary co-culture studies show how it provides novel metabolic insight into interspecies interactions.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 28 September 2016
Received in revised form
26 January 2017
Accepted 31 January 2017
Available online 13 February 2017

Keywords:
ANOVA-PCA
POCHEMON
Micro-organism
Co-culture
Interspecies interaction

ABSTRACT

Revealing the biochemistry associated to micro-organismal interspecies interactions is highly relevant for many purposes. Each pathogen has a characteristic metabolic fingerprint that allows identification based on their unique multivariate biochemistry. When pathogen species come into mutual contact, their co-culture will display a chemistry that may be attributed both to mixing of the characteristic chemistries of the mono-cultures and to competition between the pathogens. Therefore, investigating pathogen development in a polymicrobial environment requires dedicated chemometric methods to untangle and focus upon these sources of variation. The multivariate data analysis method Projected Orthogonalised Chemical Encounter Monitoring (POCHEMON) is dedicated to highlight metabolites characteristic for the interaction of two micro-organisms in co-culture. However, this approach is currently limited to a single time-point, while development of polymicrobial interactions may be highly dynamic. A well-known multivariate implementation of Analysis of Variance (ANOVA) uses Principal Component Analysis (ANOVA-PCA). This allows the overall dynamics to be separated from the pathogen-specific chemistry to analyse the contributions of both aspects separately. For this reason, we propose to integrate ANOVA-PCA with the POCHEMON approach to disentangle the pathogen dynamics and the specific biochemistry in interspecies interactions. Two complementary case studies show great potential for both liquid and gas

☆ Selected paper from XVI Chemometrics in Analytical Chemistry, 6-10 June 2016, Barcelona, Spain.

* Corresponding author.
E-mail address: chemometrics@science.ru.nl (J.J. Jansen).

http://dx.doi.org/10.1016/j.aca.2017.01.064

chromatography - mass spectrometry to reveal novel information on chemistry specific to interspecies interaction during pathogen development.

1. Introduction

The interaction between different micro-organisms is important in many scientific fields. It may for one be a serious health problem: in patients with cystic fibrosis (CF), respiratory co-infections can lead to a higher exacerbation and hospitalization rate compared to patients only infected with one pathogen [1]. On the other hand, the unique biochemistry of co-occurring micro-organism may also be a way to enhance chemical diversity for drug discovery [2]. Co-occurring micro-organisms may also influence water quality [3,4] and be explicitly used in industrial fermentation processes [5,6].

Co-occurrence of micro-organisms can lead to interaction between the species. Interaction-related Metabolites may be 1) *de novo* produced, or 2) upregulated, or downregulated compared to the metabolites that the individual species produce [2,7]. These metabolite changes of interspecies interaction can be either beneficial or detrimental for both or for one of the species and, if there is any, to their human host (e.g. a human with respiratory infections) [8,9].

The complexity of the microbiome makes studying the interaction between different species *in vivo* a very challenging task. The metabolite production by pathogens can be different in the presence of other pathogens, and it may also be highly dynamic with regards to pathogen growth [7,10,11]. Therefore, *in vitro* studies are necessary to understand these complex biochemical interactions. Microorganism co-cultures (multiple microorganism species grown within a single confined environment) can be used to study how pathogens develop over time *in vitro*, as well as how they behave in close proximity of other micro-organisms [12]. The *de novo* produced compounds may exhibit interesting biological activities, such as antimicrobial and anticancer activities [2]. This makes microorganism co-cultures a promising approach to discover new natural bioactive compounds that can be used e.g. for medicinal purposes [7].

Detecting the induction of metabolite biosynthesis in microorganism co-culture requires sensitive metabolomic techniques mainly based on mass spectrometry [13]. Both liquid and gas chromatography coupled to mass spectrometry (LC-MS, GC-MS) provide efficient determination of metabolites produced by the pathogen(s) under study. Data analysis to find those metabolites characteristic for interspecies interaction is often done in a univariate manner [7,14,15]. However, a pathogen can often not be identified based on one characteristic metabolite, and a multivariate metabolite pattern is then required [16]. The metabolites are produced at different rates in different stages of the infection [7], which may provide invaluable information on the interaction dynamics. These patterns might be obscured by other natural variability in the data, such that a generic data analysis method may not detect them.

Dedicated chemometric methods may provide a comprehensive overview of the involved metabolites. Methods used for co-culture studies include Principal Component Analysis (PCA) [13], Analysis of Variance (ANOVA) [17,18], Self-Organizing Maps (SOM) [19], and multivariate Discriminant Analysis [10,13]. Although these methods provide insight in which aspects of the metabolic profiles are co-culture specific, they do not discriminate between the two different sources of co-culture biochemistry, i.e. mixing and interspecies interaction. This means that the biochemistry related to interspecies interaction remains convoluted. Recently, we presented Projected Ortogonalized CHemical Encounter MONitoring (POCHEMON) to specifically highlight these metabolic alteration in co-culture [7]. However, the dynamics of pathogen development in co-culture cannot be directly assessed with POCHEMON or any other of the above mentioned methods.

Analysis of Variance can be used to separate the data into contributions related to different factors of variation in the data and their interactions [20,21]. Multivariate Analysis of Variance (MANOVA) is the extension of ANOVA to multiple independent variables, which has several disadvantages making it less applicable [21]. Several of these drawbacks may be overcome by regularization, involving an additional meta-parameter [22]. Several other multivariate implementations of ANOVA exist, which vary in the way the effect matrices are analysed. The most widely used methods are ANOVA-Simultaneous Component Analysis (ASCA) [23,24], and ANOVA-PCA [20,21]. In ASCA, PCA is applied directly onto each effect matrix. In ANOVA-PCA, PCA is applied to the sum of an effect matrix and the matrix of residuals. Other methods have been developed to perform PCA on biologically more relevant partitions than those obtained from 'standard' ANOVA models, such as Principal Response Curves [25] and SMART analysis [26], that fit within a generic framework that combines ANOVA and PCA [27]. Also alternatives for PCA, used within the ANOVA framework have been described such as Parallel Factor Analysis (PARAFASCA) [28] and Target Projection (ANOVA-TP) [29].

We propose the combination of ANOVA-PCA with POCHEMON for dedicated analysis of dynamic co-culture studies. This strategy allows for the extraction of three types of information:

1) information on the dynamic patterns common to both pathogens and to their co-culture,
2) information on the constitutive effect of interspecies interaction on pathogen metabolism, present at all stages of infection, and
3) information on the interspecies interaction dynamics.

We demonstrate this strategy on two complementary time-resolved microbial co-culture studies: an LC-MS study on *Aspergillus clavatus* and *Fusarium* sp. at four different time points at day level, and a GC-MS study of *Pseudomonas aeruginosa* and *Aspergillus fumigatus* at three different time points at hour level. The LC-MS study involves a fungus-fungus interaction where the metabolites are detected in the growth medium. Since the method is destructive, each time point measured involves different culture samples. In contrast, the GC-MS study involves a bacterium-fungus interaction where volatile metabolites are detected in the culture headspace such that the same samples may be followed over time. To assess the added value of the information from ANOVA-POCHEMON, we compare its results with its two constituent methods POCHEMON and ANOVA-PCA.

2. Theory

2.1. PCA

In PCA, a data matrix **X** is decomposed into a score matrix **T** and