

## T15. Pattern recognition methods for profiling microbial communities

*Agnieszka Lemanska<sup>1</sup>, Karlheinz Trebesius<sup>2</sup>, Karl Grammer<sup>3</sup>, Dustin J Penn<sup>4</sup>, Richard G Brereton<sup>1</sup>*

<sup>1</sup>*Centre for Chemometrics, School of Chemistry, University of Bristol, United Kingdom*

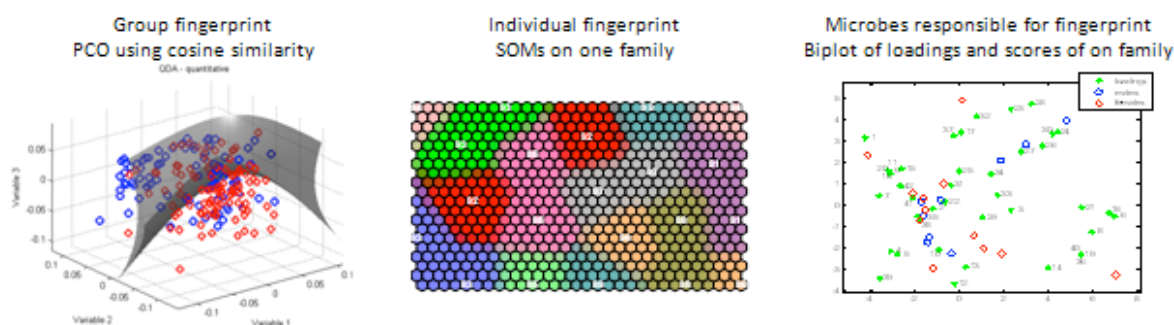
<sup>2</sup>*Vermicon AG, Munich, Germany*

<sup>3</sup>*Ludwig Boltzmann Institute for Urban Ethology, Department for Anthropology, Vienna, Austria*

<sup>4</sup>*Konrad Lorenz Institute for Ethology, Austrian Academy of Sciences, Vienna, Austria*

Human secretions contain large numbers of microbes. In this study carried out in a remote village in Austria microbial samples from armpits of almost 200 individual donors (88 males and 108 females) grouped into families were investigated. Several pattern recognition techniques were applied to reveal trends in microbial fingerprint of subjects.

DGGE (Denaturing Gradient Gel Electrophoresis) band tables that require accurate alignment of all the unique bands (corresponding to different microbes) among all the included samples may be difficult to produce. For smaller studies (one family) they are feasible. PCA (Principal Component Analysis) can be applied to band tables.



For larger studies automated pair-wise similarity measures are more feasible alternatives. In this approach both qualitative (presence or absence of unique bands using the Jaccard distance) and quantitative (presence and position using the cosine distance) measures combined with fuzzy matching of peak positions were applied. Patterns represented by dissimilarity matrices were revealed by PCO (Principal Co-ordinates Analysis).

Pair-wise dissimilarity and PCO analysis revealed considerable separation between genders. Rank analysis showed that the within individuals variation is significantly less than between individuals variation. This suggests that the individuals have unique microbial fingerprint. SOMs (Self Organising Maps) deliver information about specific microbes characteristic for individuals and demonstrate the existence of an individual fingerprint. The SOMDI (SOM Discrimination Index) together with supervised SOMs is described as a mechanism for determining which microbes are characteristic of individuals.

Group fingerprints (e.g. to differentiate males and females) can be determined using PCO components as input to supervised classifiers, including one class methods such as Quadratic Discriminant Analysis (QDA) and Support Vector Domain Description (SVDD) and two class methods such as PLS-DA (Partial Least Squares Discriminant Analysis) and Support Vector Machines (SVM). The validation of classification models is described.