92 | AnnotIEM: A novel tool for microbiome species-level annotation of 16S gene based microbial sequencing

M. Bhattacharyya[1,2], M. Reiger[1,2], L. Rauer[1,2,3], C. Huelpuesch[1,2], C. Traidl-Hoffmann[1,2,3,4], A. Neumann[1,2,3]
[1]*Chair of Environmental Medicine, Technical University Munich, Augsburg, Germany*
[2]*Department of Environmental Medicine, Faculty of Medicine, University of Augsburg, Augsburg, Germany*
[3]*Institute of Environmental Medicine, Helmholtz Zentrum Munich, Augsburg, Germany*
[4]*Christine Kühne – Center for Allergy Research and Education (CK-Care), Davos, Switzerland*

In order to generate meaningful results from microbiome analysis, identification of microbial species that are relevant to various clinical conditions (eg, Staphylococcus aureus in atopic dermatitis) is needed. However, annotation of microbial taxonomy down to the species level presents a unique challenge because the main databases available (eg, Silva, RDP and Greengenes) are not curated and have many errors and missing information. Here we present AnnotIEM - a new tool for species level annotation of sequences (OTUs or ASVs) derived from 16S rDNA sequencing. A novel hit selection algorithm is used to combine the annotation output from multiple databases to ensure accurate identification of a species and reducing the potential for errors from each of them alone. When tested using mock community control datasets, the precision of annotation with AnnotIEM is greater than 80% on the species level and greater than 90% on the genus level, better than any currently available annotation tool. Benchmarking using a number of real case studies shows that annotation by AnnotIEM gave rise to a higher fraction of sequences annotated on the species level (60-80%) as compared to any other currently available annotation tool (eg, IMNGS, SILVA, RDP Classifier and q2classifier from QIIME2). In conclusion, the AnnotIEM tool shows that using a number of databases combined with a novel hit selection algorithm results in a significantly better annotation on the species level, which can greatly enhance the results of any microbiome analysis study.