241

Extraction bias and chimera formation predicted by bacterial morphology and cell number in microbiome sequencing data

L Rauer[1,2,3], A De Tomassi[2], CL Müller[4,5], C Hülpüsch[2,6], C Traidl-Hoffmann[1,2,3,6,7], M Reiger[2] and AU Neumann[1,2] *1 Institute of Environmental Medicine, Helmholtz Munich, Augsburg, Germany, 2 Environmental Medicine, Faculty of Medicine, University of Augsburg, Augsburg, Germany, 3 Chair of Environmental Medicine, Technical University of Munich, Munich, Germany, 4 Institute of Computational Biology, Helmholtz Munich, Neuherberg, Germany, 5 Department of Statistics, Ludwig-Maximilian University of Munich, Germany, 6 CK CARE, Christine Kühne Center for Allergy Research and Education, Davos, Switzerland and 7 ZIEL — Institute for Food & Health, Technical University of Munich, Freising, Germany*

Microbiome next-generation sequencing data are distorted by multiple laboratory and bioinformatic biases. Extraction bias, sequence errors and contamination are major factors blurring true biological signals, and could potentially be corrected by jointly optimizing experimental and computational workflows. We compared dilution series (10^8-10^4 bacteria) of 3 mock communities with an even or staggered composition. DNA was extracted with 8 different extraction protocols (2 buffers, 2 extraction kits, 2 lysis conditions). Extracted DNA was sequenced (V1-V3 16S) together with corresponding DNA mocks. Sequences were denoised using DADA2, and annotated by exact matching against reference genomes. Independent of the extraction protocol, contamination increased with less input cells, but interestingly, chimera formation increased with higher input cells. Microbiome composition was significantly different between extraction kits and lysis conditions, but not between buffers. Differences in the skin microbiome between two participants were more pronounced than any difference between extraction protocols. Bias in microbiome composition compared to corresponding DNA mocks revealed that extraction protocols favored specific groups of bacteria. Strikingly, this extraction bias per groups of species was predictable by bacterial cell morphology. We provide novel explanations that higher DNA density increases chimera formation during PCR amplification, and present a robust link between cell morphology and extraction bias. These findings pave the road for bioinformatic correction of biases in microbiome data.