Journal of Biotechnology 250 (2017) 51–55

Contents lists available at ScienceDirect

Journal of Biotechnology

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

# Impact of sample processing on human airways microbial metagenomes

Lutz Wiehlmann[a,b,c,*,1], Katarzyna Pienkowska[a,1], Silke Hedtfeld[a], Marie Dorda[a,b], Burkhard Tümmler[a,d]

[a] Clinical Research Group, OE 6711, Medizinische Hochschule Hannover, D-30625 Hannover, Germany
[b] Core Unit 'Next Generation Sequencing', Medizinische Hochschule Hannover, D-30625 Hannover, Germany
[c] Institute for Human Genetics, Medizinische Hochschule Hannover, D-30625 Hannover, Germany
[d] Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), Member of the German Center for Lung Research, D-30625 Hannover, Germany

ARTICLE INFO

*Article history:*
Received 6 September 2016
Received in revised form 4 January 2017
Accepted 10 January 2017
Available online 23 January 2017

*Keywords:*
Airways
Metagenome
Microbiome
Next generation sequencing

ABSTRACT

Whole metagenome shotgun sequencing provides information about the gene content and the composition of microbial communities provided that the processing of the samples does not introduce a methodology-driven bias. We tested the impact of DNA isolation and storage period on the metagenome profile. Deep throat swabs were collected from healthy adults and an infected infant. DNA was isolated by sonification or enzymatic lysis either immediately or after 24 h storage in agar gel Amies transport medium at room temperature. Disruption of cells and subsequent fragmentation of DNA by sonification was as suitable as the common enzymatic lysis to generate high-quality metagenomes particularly for low total DNA input of less than ten nanograms. Conversely, storage of samples for 24 h produced severely distorted metagenomes. The majority of species became less abundant or even extinct, whereas a few *Streptococcus*, *Neisseria* and *Haemophilus* spp. proliferated so that the total number of bacterial reads increased at the expense of human reads. We recommend that samples for metagenome analysis should be immediately processed or frozen at −80 °C.



## 1. Introduction

Microbial metagenomics is coming-of-age (Thomas et al., 2012; Nelson, 2015; Nayfach and Pollard, 2016). Next generation sequencing technologies provide sufficient number of reads, sequencing becomes less and less expensive, both soft- and hardware can meanwhile cope with the avalanche of data within a reasonable amount of time and visualization tools have been developed to present the major findings for the non-proficient user in an educational and appealing way (Nayfach and Pollard, 2016; Huson et al., 2016; Segata et al., 2013). The metagenome reads are either assembled into longer contigs to identify and annotate genes (Huson et al., 2016) or are mapped onto nucleotide or amino acid reference sequences deposited in the databases (Losada et al., 2016). The bottleneck of a mere mapping approach has been the limited number of reference sequences. This is still true for archaea, moulds, fungi and viruses, but the encyclopedia of completely sequenced bacterial genomes, particularly pathogens, has substantially increased during the last five years (Wattam et al., 2014) so that at least the bacterial communities in habitats of medical interest can be satisfactorily covered by an aligned whole metagenome shotgun sequence.

Value and significance of a metagenome analysis sensitively depend on adequate sampling and processing, for example, on the amount and microbial composition of available template (Morgan et al., 2010; Bowers et al., 2015), particularly in case of minute starting material of less than one nanogram of DNA (Hall et al., 2014; Rinke et al., 2016). The metagenome data output is moreover sensitively influenced by the employed DNA extraction method (Jones et al., 2015; Anderson et al., 2016; Knudsen et al., 2016), the protocol for library preparation (Bowers et al., 2015; Jones et al., 2015) and the applied sequencing technology (Clooney et al., 2016). If the site of sampling is distant from the laboratory, the collected specimen is typically stored in preservation medium and then shipped to the institution which is capable to process the sample from DNA

* Corresponding author at: Klinische Forschergruppe, Klinik für Pädiatrische Pneumologie, Allergologie und Neonatologie, OE 6710, Medizinische Hochschule Hannover, Carl–Neuberg–Str. 1, D–30625 Hannover, Germany.
*E-mail addresses:* Wiehlmann.Lutz@mh-hannover.de, wiehlmann@web.de (L. Wiehlmann), Pienkowska.Katarzyna@mh-hannover.de (K. Pienkowska), Jansen.Silke@mh-hannover.de (S. Hedtfeld), Boehm.Marie@mh-hannover.de (M. Dorda), Tuemmler.Burkhard@mh-hannover.de (B. Tümmler).
[1] These authors contributed equally to the work.

http://dx.doi.org/10.1016/j.jbiotec.2017.01.001
0168-1656/

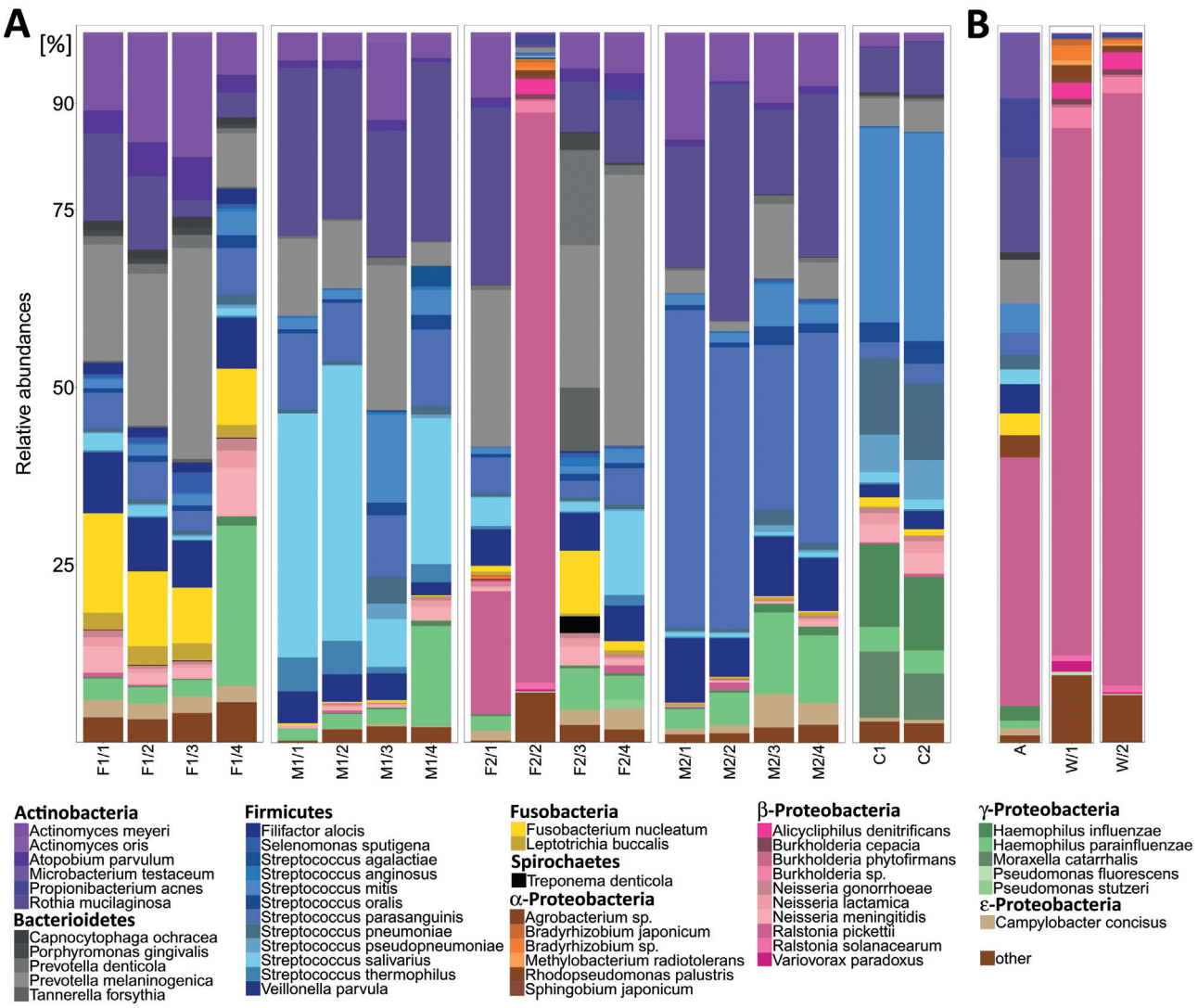


**Fig. 1.** A (left). Composition of bacterial communities in deep throat swabs of two healthy females F1, F2 and two healthy males M1, M2 and an infant C with an acute upper airway infection. Data are presented in a stacked bar chart of relative abundance as fractions of total reads. Nomenclature of samples: 1,2: processed by enzymatic lysis; 3,4: processed by sonification; 1,3: processed immediately after sampling; 2,4; processed after a storage period of 24 h at room temperature. B (right). Composition of bacterial communities in two samples of autoclaved deionized water used for the set-up of solutions in the laboratory and in an agar gel with Amies medium classified as sterile by the manufacturer.

extraction to final data output (Anderson et al., 2016). A shipping time of less than 4 h at 4 °C is recommended (Marre et al., 2008), but this request is often not manageable in environments with poor infrastructure and logistics (Anderson et al., 2016). But even in the context of collaborative field studies or multicenter clinical trials the shipping time to the core laboratory may take more than 24 h.

We wanted to test the impact of storage and processing on the outcome of metagenome analysis. Having established a pipeline for human airway metagenomics (Losada et al., 2016) applicable to samples with cystic fibrosis (Lim et al., 2014) or chronic obstructive lung disease (Cameron et al., 2016), we performed a comparative study with throat swabs that are routinely applied for the diagnostics of airway infections. For the first time the distorting influence of a common shipping time of one day on the abundance of bacterial species in the metagenome data output is demonstrated.

## 2. Experimental procedures

Four serial deep throat swabs were collected within five minutes from an infant with an acute airway infection (disease control) and four healthy adults. Swab tips were inserted into agar gel with Amies medium of the 108c transystem (Copan Italia, Brescia) which according to the manufacturer and the literature (Citron et al., 2000) is an adequate transport system for aerobic and anaerobic bacteria and maintains 100% survival of most test bacteria for at least 24 h.

Half of the samples were processed immediately and the other half was stored in the dark at room temperature and processed 24 h after sampling. DNA was isolated from the swabs by either mechanical or chemical lysis. In case of the latter, the cells were removed from the swab tips by 15 min incubation in SE-buffer with intermittent vortexing. After centrifugation (15 min; 14,000*g*) the DNA was extracted from the dissolved pellet by digestion with proteinase K and the NucleoSpin Tissue Kit (MACHERY-NAGEL, Düren) following the 'hard-to-lyse-bacteria' protocol. Alternatively, the swab tip was vortexed in 0.3 ml low-TE buffer. Suspended cells were sonicated in a Covaris S220 instrument Step 1: 5s, Peak Incident Power (PIP: 200W, Duty Factor (DF): 2, Cycles/Burst (CB): 100, Average Power (AP): 4 W; Step 2: 30 s, PIP: 275W, DF: 5, CB: 100, AP: 13.8 W). Step 1 and 2 were repeated three times. An aliquot of 130 μl was transferred to a shear tube (micro Tube AFA Fibre SnapCap 6 × 16 mm, PN 520045) and sheared in the Covaris instrument (360 s, PIP: 175 W, DF: 10, CB: 200, AP: 17,5 W). The sheared DNA with an average