Deliverable

2.2 Report on standard protocol for sample processing

Version: 1

Due: Month 24

Completed: Month 24

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Contents

Deliverable Description

The overall aim of Deliverable D2.2 (*Standard protocol for sample processing*) is to summarize sample processing experiences for metagenomics and to provide a collection of well-working protocols (fit-for-purpose). The context of all deliverables in work package 2, and the overall workflow as well as the scope of deliverable 2.2 is given in Figure 1. Deliverable 2.1 was already addressing *Harmonized standards for sample handling* including sampling, transport, and storage conditions. Now we provide with the present document current protocols of sample homogenization, nucleic acid extraction, and subsequent processing like cDNA preparation from RNA.

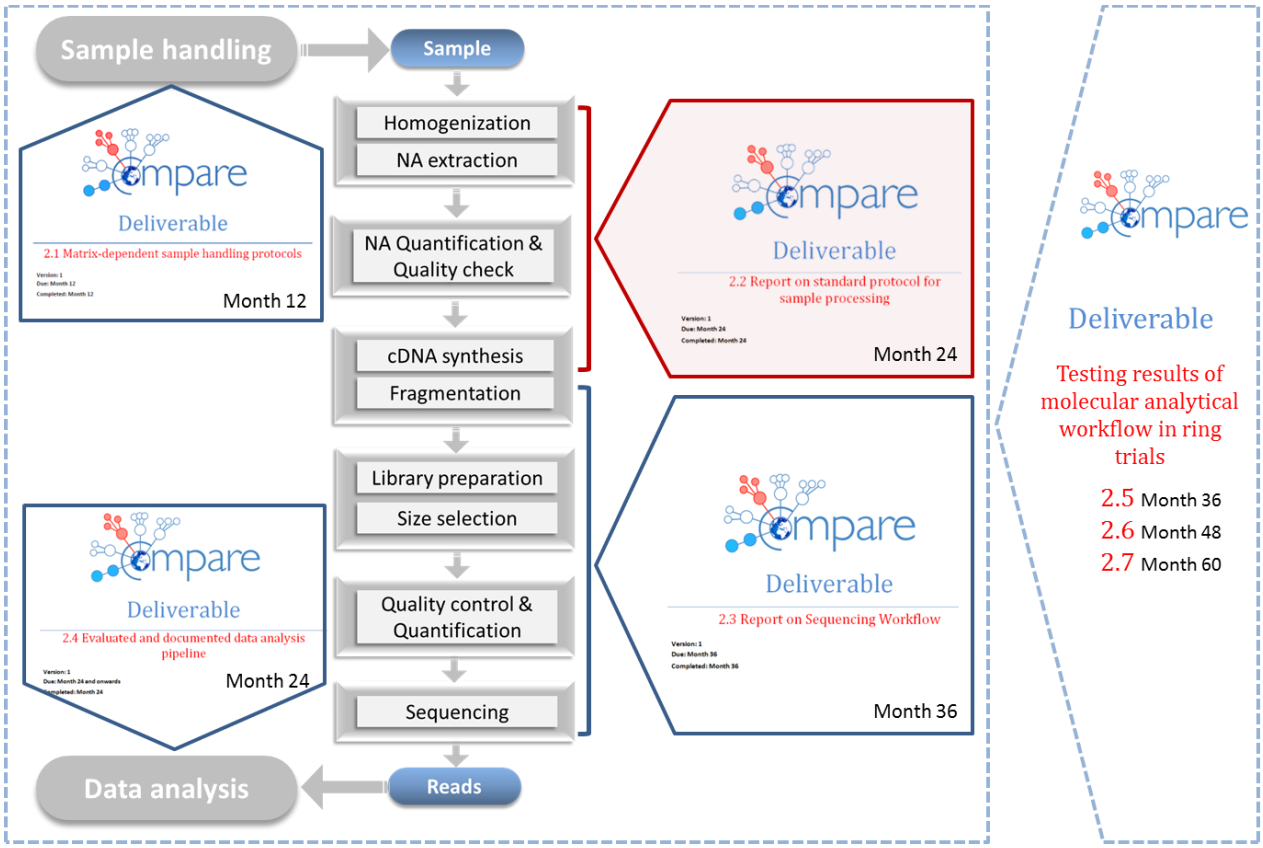

FIGURE 1: OVERVIEW OF DELIVERABLES (2.1 – 2.7) RELATED TO COMPARE WORK PACKAGE 2 WITH THEIR DUE-DATES AND SCOPES AS OUTLINED IN THE GRANT AGREEMENT. THE PRESENT DOCUMENT DEALS WITH D2.2 (MARKED RED).

Work package 2 (WP2) is addressing the harmonization of standards for sample handling as a basis for other tasks of the COMPARE Project. D2.2 is closely connected to WP2 task 2. This task deals with the development of standardized protocols for sample processing as well as evaluation and comparison of available procedures with regard to their performance, yield, and quality. Different sample types (e.g. tissue, ticks, fecal samples, sewage, bacterial suspensions, food samples) and pathogens (viruses, bacteria and parasites) have been tested so far. Based on that, an inventory of methods and protocols has been prepared on the COMPARE Share Site. The as-yet used and to some extent intensively tested protocols are available in unified "COMPARE-branded" Laboratory Operating Procedures (LOPs). As a next step, these protocols will be validated in COMPARE e.g. with

organized ring trials, pilot studies and by the use of the different partner institutes; afterwards LOPs will be transformed into "official" COMPARE Standard Operation Procedures (SOPs).

Notwithstanding that the protocol collection is nearly completed, as outlined in the project planning, WP2 partners intend to continue workflow improvement to test further matrices for example. This will be implemented in close connection with ring trials and future pilot studies (WP2 task 7, *Quality management and ring trials*) as well as in collaboration with other work packages.

Metagenomic Workflow – from sample to cDNA

An overview of the protocols for sample homogenization, nucleic extraction and quantification and cDNA synthesis is given in Figure 2.

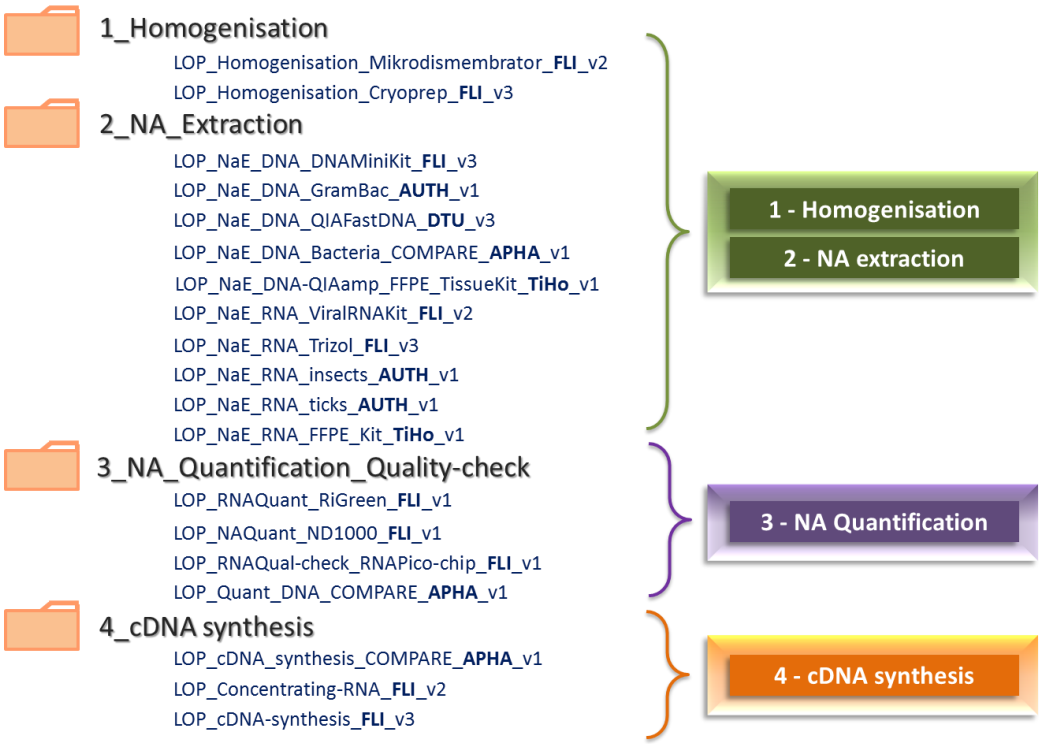

FIGURE 2: PROTOCOLS AVAILABLE IN LOP FORMAT FOR THE RESPECTIVE WORKFLOW STEPS AND DEPOSITED ON COMPARE SHARE SITE. FURTHER PROTOCOLS NOT IN LOP FORMAT ARE AVAILABLE IN A SEPARATE FOLDER.

Workflow steps and tests

Sample Homogenization

FLI tested and compared different homogenization methods: Tissuelyser (Qiagen), CryoPrep CP02-system (Covaris) and Micro-Dismembrator (Sartorius).

The CryoPrep allows a rapid and efficient sample homogenisation of samples deeply frozen in liquid nitrogen in order to avoid RNA degradation during the homogenisation process. The system delivers repeatable high impact mechanical force in order to cryofracture sample material for improved lysis with extraction buffers. The sample homogenate is then resuspended in lysis buffer AL (Qiagen).

The Mikro-Dismembrator is a laboratory grinding mill which can be used for efficient and reproducible fine grinding and homogenisation of samples for the isolation of DNA and RNA. The homogenisation is performed in liquid nitrogen within PTFE shaking flasks and chrome steel or tungsten carbide grinding balls enabling a continuously cooling chain of the samples and thus prevents degradation of the RNA. The Mikro-Dismembrator is operated with a selected shaking frequency for a defined time. The homogenate is resuspended in lysis buffer AL (Qiagen).

The Tissuelyser disrupts samples within lysis buffer (e.g. AL buffer) through high-speed shaking in plastic tubes with stainless steel, tungsten carbide, or glass beads. The Tissuelyser can be used for several samples in parallel (high-throughput applications). However, this treatment works without cooling the sample and causes high sheering forces leading to fragmentation of the nucleic acids.

Comparable high quality results were obtained in most cases using CryoPrep and Micro-Dismembrator. Both homogenization techniques are combined with cooling steps based on liquid nitrogen shortly before (CrypPrep) or while homogenization (Micro-Dismembrator). In contrast, the Tissuelyser treatment without cooling step results in reduced RNA integrity at least for bacteria suspensions and faeces samples (Fig. 3) as well as highly reduced DNA amount extracted from Echinococcus egg samples (not shown).

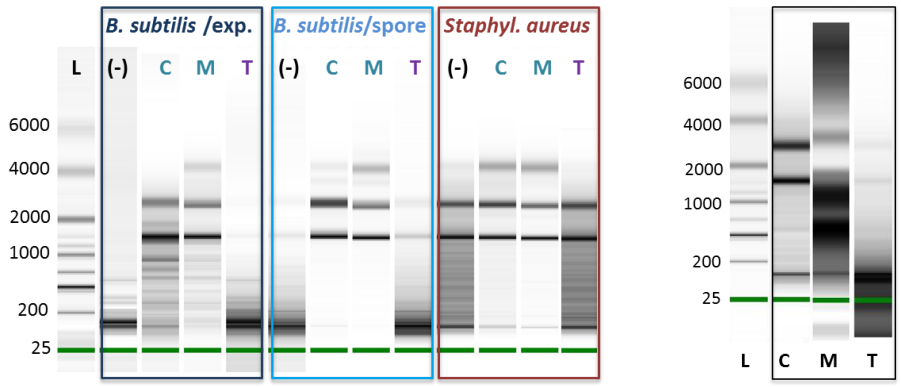

FIGURE 3: COMPARISON OF DIFFERENT HOMOGENIZATION METHODS. SHOWN IS THE RNA QUALITY USING RNA 6000 PICO CHIP (BIOANALYZER, AGILENT). BACTERIAL SAMPLES ARE SHOWN ON THE LEFT SIDE; FECAL SAMPLES ARE SHOWN ON THE RIGHT SIDE. LEGEND: EXP., EXPONENTIAL GROWTH PHASE; **L**, LADDER; **(-)**, WITHOUT HOMOGENIZATION STEP; **C**, CRYOPREP; **M**, MICRODISMEMBRATOR; **T**, TISSUE LYSER. THE LABELLING FOR THE LADDER IS GIVEN ON THE LEFT SIDE.

The cooling step seems to be crucial to obtain high quality RNA. Interestingly, when midges were used as sample material (not shown), CryoPrep and Tissuelyser produced high quality RNA and not the Micro-Dismembrator. This may be due to the application of lysis buffer during the homogenization process as used with the Tissuelyser. Taken together, the CryoPrep technique worked well for most samples, whereas Micro-Dismembrator and Tissuelyser showed some weaknesses. At the moment, we provide LOPs for CryoPrep and Micro-Dismembrator and not for Tissuelyser. The latter device can be useful for high-throughput applications and sample preparation for standard diagnostics via real-time PCR.

Nucleic acid extraction

Depending on the sample type, anticipated pathogen, and research question, one could extract either DNA or RNA. The use of genomic DNA should in principle allow detection of all organisms present in a certain sample but would neglect RNA viruses. Therefore, we recommend using RNA as template especially in cases where the nature of the pathogen is still unknown.

Many nucleic acid extraction protocols are available in LOP format (see Table 1). Their use especially depends on sample matrix (e.g., liquid or solid material), sample type (e.g., arthropod vectors, faeces, bacterial suspensions) or sample treatment (e.g., untreated, frozen, fixed, processed).

TABLE 1: OVERVIEW OF PROTOCOLS FOR NUCLEIC EXTRACTION (DNA OR RNA) FOR MANY DIFFERENT SAMPLE TYPES AVAILABLE IN LOP FORMAT ON THE COMPARE SHARE SITE.

Sample material	NA	Kit	Lysis buffer or additional reagents	Provided by
Bacterial culture	DNA	without	none	APHA
Liquid samples (whole blood, plasma, serum, buffy coat, other body fluids, bone marrow, lymphocytes, cultured cells) or homogenised and lysed samples (tissue and stool homogenates)	DNA	QIAamp DNA Mini Kit (Qiagen)	Buffer AL with Proteinase K	FLI
Cultures of Gram-positive bacteria	DNA	without	Lysostaphin, Proteinase K	AUTH
Fecal or sewage samples	DNA	QIAamp Fast DNA stool mini kit (Qiagen)	Buffer AL, Proteinase K	DTU
Formalin-fixed, paraffin-embedded material	DNA	QIAamp DNA FFPE Tissue Kit (Qiagen)	Buffer AL, Proteinase K	TiHo
Liquid samples (serum/plasma, cell culture supernatant, EDTA-blood diluted 1:2 in PBS, supernatant from swabs, liquor) or samples homogenised and lysed	RNA	QIAamp RNeasy Mini Kit (Qiagen)	Buffer AL, Trizol	FLI
Ticks	RNA	QIAamp RNeasy Mini Kit (Qiagen)	Buffer RLT with β-Mercaptoethanol	AUTH
Insects	RNA	QIAamp RNeasy Mini Kit (Qiagen)	Buffer RLT with β-Mercaptoethanol	AUTH
Liquid samples (serum/plasma, cell culture supernatant, EDTA-blood diluted in PBS, supernatant from swabs, liquor) or homogenised and lysed samples (tissue and stool homogenates)	RNA	QIAamp Viral RNA Mini Kit (Qiagen)	Buffer AVL	FLI
Formalin-fixed, paraffin-embedded material	RNA	RNeasy FFPE Kit (Qiagen)	Buffer PKD, Proteinase K	TiHo

A combined QIAmp RNeasy Mini Kit (Qiagen) and Trizol procedure provided best results for homogenised organ samples (Fig. 4). In comparison with classical phenol chloroform extractions, the combined QIAmp RNeasy Mini Kit / Trizol procedure lead to RNA of higher quality (Fig. 4).

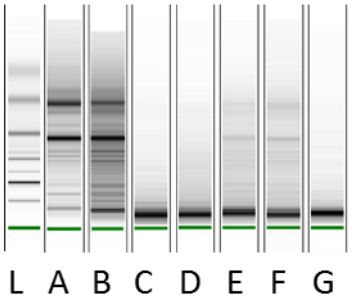

FIGURE 4: COMPARISON OF RNEASY KIT WITH CLASSICAL EXTRACTION AND PRECIPITATION METHODS. SHOWN IS THE RNA QUALITY USING RNA 6000 PICO CHIP. SAMPLE MATERIAL, POOL OF SPLEEN AND BRAIN. **L**, LADDER; **A**, CRYOPREP (CP) + TRIZOL + RNEASY; **B**, AS A, AFTER 13 WEEKS STORAGE OF THE HOMOGENATE AT 4°C; **C**, CP + ETHANOL/NEUTRAL SODIUM ACETATE PRECIPITATION (SAP); **D**, CP + ISOPROPANOL/NEUTRAL SAP; **E**, CP + PHENOL/CHLOROFORM EXTRACTION + ETHANOL/NEUTRAL SAP; **F**, CP + PHENOL/CHLOROFORM EXTRACTION + ISOPROPANOL/NEUTRAL SAP; **G**, CP + PHENOL/CHLOROFORM EXTRACTION + ETHANOL/ACIDIC SAP WITH GLYCOGEN.

AUTH successfully used the QIAmp RNeasy Mini Kit for the RNA extraction from arthropod vectors (ticks and insects). In this case, lysis buffer RLT supplemented by β-Mercaptoethanol is recommended. For cockroaches, the CryoPrep and RNeasy/Trizol procedure worked very well.

RKI and DSMZ performed a study comparing different DNA extraction techniques and protocols and their influence on plasmid vs. chromosomal DNA recognition and sequence quality after Illumina sequencing. The Gram-negative bacterium *Klebsiella pneumoniae* were chosen as an example pathogen. Purification of DNA was based on matrix binding (silica or anion exchange resin) or differential precipitation (salting out), respectively. The choice of extraction kit had little effect on sequencing quality and coverage across different replicons (chromosome; large, low-copy number and small, high-copy number plasmids), except for an apparent depletion of small plasmids (< 5kb) during precipitation-based extractions. Sequencing coverage provided copy-number estimates of small plasmids that were consistently higher than those from quantitative real-time PCR (Becker et al., 2016).

DTU performed a study comparing a total of 8 commercially available DNA extraction kits, and 16 protocols based on these in a stepwise approach. The kits were tested using 3 sample matrices: human faeces, pig faeces, and hospital sewage. DNA extraction was evaluated using spike-in controls and different types for bead beating. DNA concentration, purity, and stability was assessed and the microbial community compositions compared using 16S rRNA gene profiling, and for selected samples using metagenomic sequencing. While there was no single approach among the 16 procedures tested that appeared to completely resolve all challenges, we found one procedure based on the QIAamp Fast DNA stool Minikit useful for a number of reasons, including the following: (i) DNA extracts contained large amounts of DNA (sufficient to permit PCR-free metagenomic sequencing) with high reproducibility, (ii) DNA extracts were of high quality in terms of DNA purity and stability, (iii) DNA from both Gram-positive and Gram-negative bacteria was reasonably well extracted (including from *Bifidobacteria*) as determined by 16S rRNA amplicon profiling and metagenomic sequencing of spiked and unspiked complex samples, (iv) the method worked well for all examined sample types based on the DNA quality assessment and inferred microbiota composition, (v) the reagents and materials required were cheaper, and (vi) the time needed for carrying out the DNA isolation was shorter than for several of the other procedures. (Knudsen et al. 2016, mSystems, DOI: 10.1128/mSystems.00095-16).

Artemis One Health conducted a preliminary study comparing two different RNA isolation methods, namely the Trizol method and the High Pure RNA isolation kit (Roche) for the extraction of viral RNA from Zika-spiked urine (diluted to a Ct value of approx. 31) that was stored at several different temperature conditions (RT, 4°C, -20°C and -134°C) over different time periods (0h, 24h, 48h and 72h). To date, viral RNA was isolated from the urine samples using the two different methods, and the RNA was quantified by RT-qPCR using primers targeting the NS5 of Zika H/PF2013. The Ct values for the two different RNA extraction methods were compared. It was found that the average Ct value of the RNA samples isolated using Trizol and the High Pure kit were both 33.5, which is equivalent to about 70000 RNA copies. The largest standard deviation between each time point when comparing the two methods was 0.61 Ct value. This indicates that the yield of the RNA extracted from Zika virus-spiked urine via the two different methods is highly comparable. Follow-up experiments will also focus on determining and comparing the quality of the isolated RNA for NGS.

Further trials testing the different nucleic acids in parallel were not yet conducted.

Nucleic acid quality check and quantification

High quality input RNA is essential for successful preparation of cDNA shotgun sequencing libraries. The Agilent Bioanalyzer (RNA 6000 Pico assay) is well-suited to analyse total RNA or mRNA with a size between 25 and 6000 bases using only 1 µl of sample. The dye-labelled nucleic acids are separated electrophoretically based on their size in a microfluidic system. Internal standards ensure assay accuracy and reproducibility. From the generated electropherogram, the software calculates an RNA Integrity Number (RIN) that can help in assessing RNA quality. According to our experience, however, the RIN value is not correlated to the sequencing success.

For nucleic acid quantification, different methods can be applied. The NanoDrop 1000 Spectrophotometer (Thermo Scientific) measures concentration and purity of nucleic acids in minute volumes (1-2 µl) with high accuracy and reproducibility. Due to the wide concentration range that can precisely be measured (lower limit 2 ng/µl, upper limits dsDNA 3700 ng/µl, RNA 3000 ng/µl, ssDNA 2400 ng/µl) prior dilution of the samples is not necessary. A more reliable quantification of RNA possible with the ultrasensitive fluorescent nucleic acid stain Quant-iT RiboGreen RNA reagent. It enables quantification of as little as 1 ng/ml RNA using a fluorescence microplate reader. The RiboGreen reagent is non-fluorescent when free in solution; upon binding to nucleic acids (single- or double stranded), the fluorescence of the RiboGreen reagent increases more than 1000-fold. The linearity of fluorescence is maintained in the presence of several compounds commonly found to contaminate nucleic acid preparations, including nucleotides, salts, urea, ethanol, chloroform, detergents, proteins, and agarose. Since RiboGreen reagent also binds to DNA, pre-treatment of mixed RNA/DNA samples with DNase has to be applied to generate an RNA-selective assay. For DNA samples, the Quantifluor dsDNA System (Promega) is recommended. It contains a fluorescent DNA-binding dye that enables sensitive quantitation of small amounts of double-stranded DNA (dsDNA) in solution. The dye shows minimal binding to single-stranded DNA (ssDNA) and RNA, allowing specific quantitation of dsDNA.

Qubit (Thermo Fisher) is also an alternative for RNA or DNA (ss DNA/ ds DNA), it use only 1µL of sample with quantification range 0,25 - 1000 ng/µL. The dye in each Assay emits fluorescence only when specifically bound to the target molecule, which minimizes the effects of contaminants on the result and improve accuracy and precision.

cDNA synthesis

For the synthesis of double stranded (ds) cDNA from purified total RNA the cDNA Synthesis System (Roche) is recommended. For that, up to 1000 ng high quality RNA in a maximum volume of 17 µl (FLI protocol) or 8.5 µl (APHA protocol) is needed. Low concentrated RNA solutions need to be concentrated with the Agencourt RNAClean XP Beads. Then, a first-strand synthesis reaction has to be performed and in order to generate ds cDNA, a subsequent second strand synthesis is done. The first and second strand syntheses are performed in the same tube, speeding up the synthesis procedure and maximizing recovery of cDNA. The last step in the cDNA synthesis is to ensure that the termini of the cDNA are blunt.

Summary and Conclusion

We provide here protocols for the crucial part of a sample processing workflow that has been often tested for metagenomics (e.g. at FLI). The first steps of the workflow, homogenization and nucleic extraction, are depending on sample type and research question. DNA is suitable for most purposes. However, when metagenomics are used for detecting unknown pathogens, that is generally the intention of the COMPARE project, it is recommendable to use RNA as initial template to avoid the a priori exclusion of RNA viruses. These first steps decide about the integrity of nucleic acids and their accessibility for further steps. We focused therefore on the improvement of these steps, homogenization and nucleic extraction, and tested different sample types and pathogens. The enthusiastic aim of these optimization steps is to develop and to provide one workflow for all samples for metagenomics with some few variations depending on sample type.
For metagenomics to detect unknown pathogens, but not for well-working standard diagnostics, we recommend the CryoPrep device for homogenisation in combination with the RNeasy Mini Kit to extract RNA (the AL+Trizol or RLT procedure depending on sample material). The sample processing workflow is currently tested with several tissue samples containing different loads of *Mycobacterium avium* subspecies.
Helpful discussions of important criteria to be considered for metagenomics to successful identify infectious pathogens can be found in Höper et al. (2016).

Becker, L., Steglich, M, Fuchs, S, Werner, G, Nübel, U. 2016. Comparison of six commercial kits to extract bacterial chromosome and plasmid DNA for MiSeq sequencing. Sci Rep 6:28063.

Höper D, Mettenleiter TC, Beer M. 2016. Metagenomic approaches to identifying infectious agents. Rev Sci Tech **35**:83-93.

Knudsen BE, Bergmark L, Munk P, Lukjancenko O, Priemé A, Aarestrup FM, Pamp SJ. 2016. Impact of Sample Type and DNA Isolation Procedure on Genomic Inference of Microbiome Composition. mSystems. 1(5). pii: e00095-16. DOI: 10.1128/mSystems.00095-16.