Name of the service: Service for metatranscriptome sequencing

Required activities - their description, scope (e.g. the number of samples, forest area), time (requirements) etc.:

- Estimated number of samples: **90 pcs (9Gb data per sample) altogether in 5 shipments,** (price for 5 shipments is included in the unit price stated in the Attachment No. 2 – Tender price, all other shipments will be charged)
- Sample of sufficient quality mean samples with the following minimal parameters: the amount of sample ≥ 2µg, concentration RNA c ≥ 20 ng/µl, cleanness of the sample OD 260/280≥1.8, OD 260/230 ≥1.8; RNA 28S:18S ≥ 1.0, RIN ≥ 6.5;

Requirements on the methodological procedure:

- Quality check of the delivered samples RNA before the commencement of the analysis
- Enhancement of **RNA fraction** (ribosome depletion for both prokaryotic and eukaryotic species) before the preparation of cDNA library
- Preparation of 250-300bp insert cDNA library (type of the library random-primed).
- Usage of the newest platforms of RNA sequencing, Illumina Novaseq6000 or any other platforms with similar output as mentioned above.
- **150 bp pair-end reads** (bilateral reading in the section length 150 bp)
- At least **30 million reads** (reading frames) per one sample.
- Requirement on the quality of data: Phred value min. **Q30 for at least 85%** reads (reading frames) and at the same time Phred value min. Q20 for at least 90% reads (reading frames)
- Delivery of the sequencing results electronically via server / shared disc in the format fasta and fastq, generally compatible data with databases of referential genome regardless of the used technology of sequencing.

The Client requires a description of the methodology of the preparation of the library, together with the delivery of sequencing results.

- Standard Analysis Metatranscriptome sequencing requirements (obligatory)
    - Data Quality Control: filtering reads containing adapter or with low quality
    - Statistics Analysis of Data Production and Quality

- Making custom database for potential host genomes and filtering host genomes sequences
- De novo Assembly
- Gene Functional Annotation (Blast2GO, eggNOG/COG, KEGG, CAZy, CARD annotation)
- rRNA& mRNA Taxonomic Analysis
- Gene Expression Analysis
- Differential Expression Analysis (two or more groups of samples) identifying genes (name and accession) that are differentially expressed.
- GO Enrichment Analysis of Differentially Expressed Genes (DEGs) (two or more groups of samples)
- Biochemical Pathway Enrichment Analysis (i.e., using KEGG) of Differentially Expressed Genes (DEGs) (two or more groups of samples)
- Comparative Analysis between Various Samples (3 or more samples, including eggNOG/COG functional comparisons, cluster analysis and Principle coordinate analysis)
- Identifying the diterpene gene cluster from sequence information supplied by the Client. (customized analysis)