Microsynth

Amplicon Metagenomics – 16S/ITS

Explore the microbial community composition of your samples
Compare taxonomic shifts within a given experimental setup

Introduction

Microbial communities are present in almost every environment. They are important drivers of biogeochemical processes, have a large influence on human health and harbor enormous potential for biotechnological applications. A major goal in the analysis of a microbial community is to identify its taxonomic composition and diversity. Sequencing of the ribosomal RNA (rRNA) gene and its internal spacer regions has become the gold standard for the identification of microorganisms such as bacteria, archaea and fungi. Its presence in all living organism, the combination of conserved and hypervariable genetic regions and the availability of curated reference databases make the rRNA gene an ideal phylogenetic marker. The combination of next generation sequencing (NGS) of partial regions of the rRNA gene combined with adequate bioinformatics analysis is a powerful tool to explore the composition of microbial communities.

Choice of the Primer System

Amplicon metagenomics is based on NGS sequencing of the microbial rRNA gene. Since NGS read lengths are limited, only parts of the rRNA gene can be amplified and sequenced. For prokaryotes, the analysis targets hypervariable regions (V1-9) on the 16S rRNA gene, while for fungi the internal transcribed spacer regions (ITS) are used for taxonomic profiling (see ***Figure 1***). An ideal primer system should be universal enough to cover a broad range of taxonomic groups, while the resulting amplicon must provide enough taxonomic information for a reliable taxonomic classification. Based on our experience and the validation of our 16S/ITS analysis pipeline, we recommend the primer systems displayed in ***Table 1***. Our service is not restricted to the displayed marker genes and primer systems, also other phylogenetic marker genes (e.g. cytochrome c oxidase I) and primer systems can be used. A pilot study can be very helpful to find the best primer system for your specific research question.

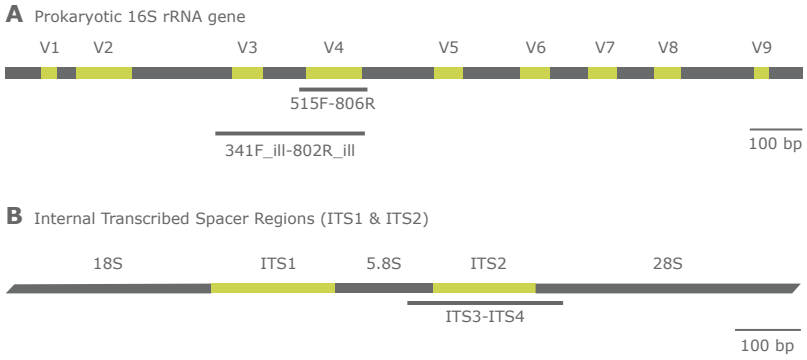


Figure 1. *Overview of loci of ribosomal gene loci commonly used for the taxonomic analysis of microbial communities. Hypervariable regions are marked in green while conserved regions are marked in grey. A. Structure of the prokaryotic 16S rRNA gene showing the nine hypervariable regions (V1-V9) and the regions targeted by the commonly used primer systems. B. Organization of the fungal rRNA gene operon showing two internal transcribed spacer regions (ITS). ITS2 is most often used for profiling fungal communities.*

Table 1. *Recommended and commonly used universal primer systems for prokaryotic (515F/806R, 341F_ill/802R_ill) and fungal (ITS3/ITS4) microbiome profiling. Only template specific sequences are shown. For Illumina-Sequencing the primer sequences must be combined with Illumina adaptor sequences in a two-step PCR approach. For further information consult our application note for Amplicon Deep Sequencing (https://www.microsynth.ch/amplicon-deep-sequencing.html).*

Primer Name	Sequence (5′-3′)	Region	Size (bp)	Source
515F 806R	GTGCCAGCMGCCGCGGTAA GGACTACHVGGGTWTCTAAT	V4	~300	[1]
341F_ill 802R_ill	CCTACGGGNGGCWGCAG GACTACHVGGGTATCTAATCC	V34	~460	[2]
ITS3 ITS4	GCATCGATGAAGAACGCAGC TCCTCCGCTTATTGATATGC	ITS2	300-400	[3]

Microsynth AG, Switzerland
Schützenstrasse 15 · P.O. Box · CH-9436 Balgach · Phone +41 71 722 83 33 · Fax +41 71 722 87 58 · info@microsynth.ch · www.microsynth.ch

Microsynth Competences and Services

Microbial community profiling is one of Microsynth's core competences. Based on our extensive experience, we provide a one-stop service from experimental design to bioinformatics analysis (see ***Figure 2***). You can either outsource the whole process or only single steps to us. We validated our whole amplicon metagenomics pipeline and the including methods by sequencing and analysis of two well characterized mock communities. For further information on our validation process and possibilities to validate your own study, please contact us.

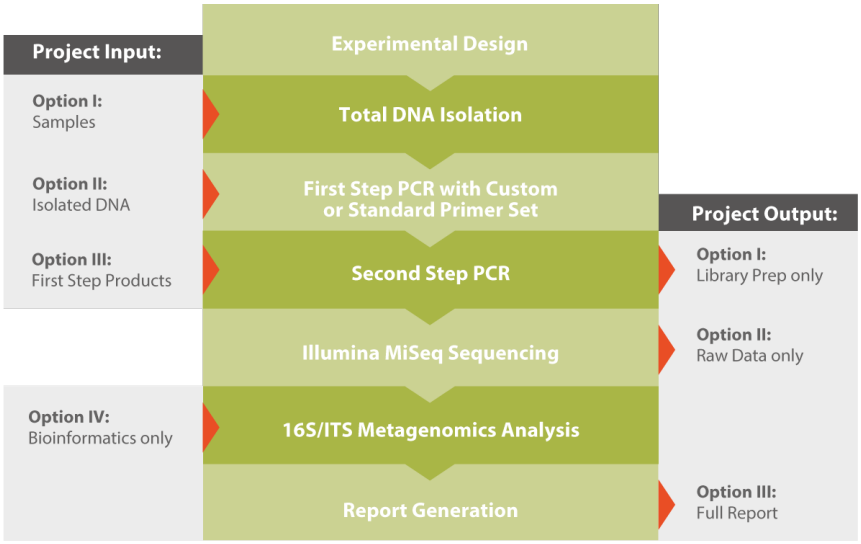


Figure 2. *Microsynths workflow for amplicon metagenomics projects. The workflow can be entered and exited at various steps dependent on the customers' requirements.*

Experimental Design

The gain and impact of a study is highly dependent on its experimental design. The use of replicates, controls and appropriate sampling methods are only a few examples of important points to consider. Make use of our experience; our NGS specialists will assist you from the start.

DNA Isolation

Microorganisms are highly diverse regarding their physiology and are often associated with surfaces like soil particles or are organized in stable agglomerations such as biofilms. DNA extraction of complex matrices is challenging. You can either perform the extraction yourself or outsource this critical step. Microsynth has extensive experience in DNA and RNA isolation from various matrices (e.g. soil or feces).

PCR Amplification

PCR amplification follows a two-step protocol. In a first step the locus-specific sequence is amplified, while in a second step the Illumina sequencing adaptors and indices are added. For projects with very low amounts of starting material we recommend our three-step PCR protocol including two subsequent locus-specific PCRs to increase the yield of sequenceable amplicons.

NGS Sequencing

Sequencing is performed on the Illumina MiSeq system. The MiSeq allows high throughput profiling at low cost, supporting read lengths up to 2x300 bp.

Bioinformatics Analysis

Microsynth's amplicon metagenomics analysis is based on up-to-date and published bioinformatics software such as USEARCH [4] to get meaningful and reliable results for your sequencing data. 16S/ITS analysis includes extensive quality filtering, denoising of operational taxonomic units (OTU), taxonomic classification based on various databases and alpha diversity analysis (i.e. OTU diversity within a given sample). Depending on the aim of the study and the experimental design (conditions, replicates), Microsynth also offers a complementary module for comparative statistics. This module includes beta diversity analysis (i.e. OTU diversity across samples) and determination of differentially abundant OTUs including appropriate statistical measures. The results are presented in an interactive and understandable format, including guidance through the results by our bioinformatics experts.

Example Results of the 16S/ITS Analysis

Microsynth's analysis pipeline provides you with a full characterization of the microbial communities in your samples. Alpha diversity measures provide an overview on OTU richness and diversity within the samples (see ***Figure 3***).

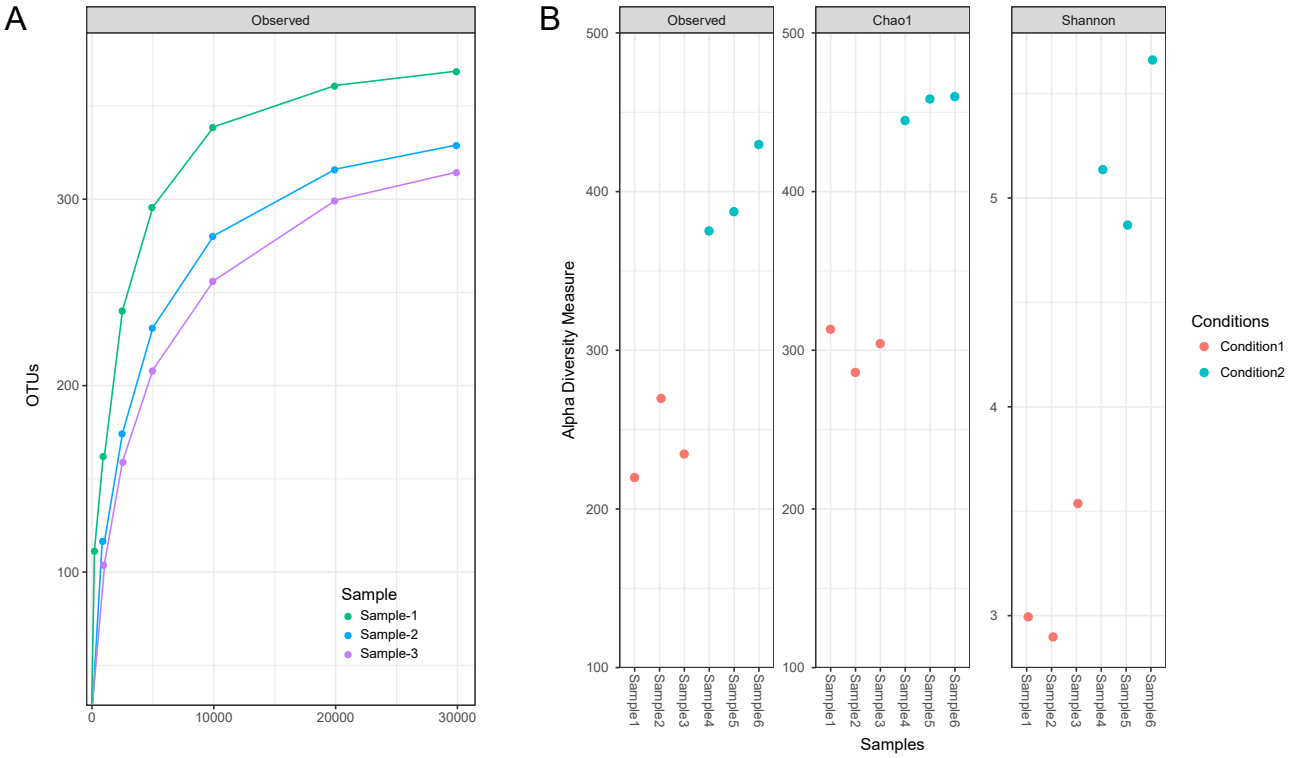


***Figure 3**. Examples for alpha diversity results. 3A. Rarefaction curves indicating whether sampling and sequencing covered the sample richness. 3B. Alpha diversity measures for the analyzed community including observed richness, Chao 1 indices representing the estimated richness and the Shannon diversity indices.*

Various reference databases are available for the taxonomic classification of OTUs. Classifications are only shown to the taxonomic level, where the classification fulfills certain confidence criteria in order to prevent false positive results. All observed OTUs are displayed together with their taxonomy and the relative abundance within the samples (see ***Table 2***).

***Table 2.** This detail of a sortable results table lists relative abundances and taxonomic identification of observed OTUs in the different samples. For taxonomic classification confidence values are calculated and OTUs are only classified up to a specific rank if its confidence value is above a preset threshold to avoid dubious classifications and false-positive results.*

Relative abundance OTU table

10 records per page Search all columns:

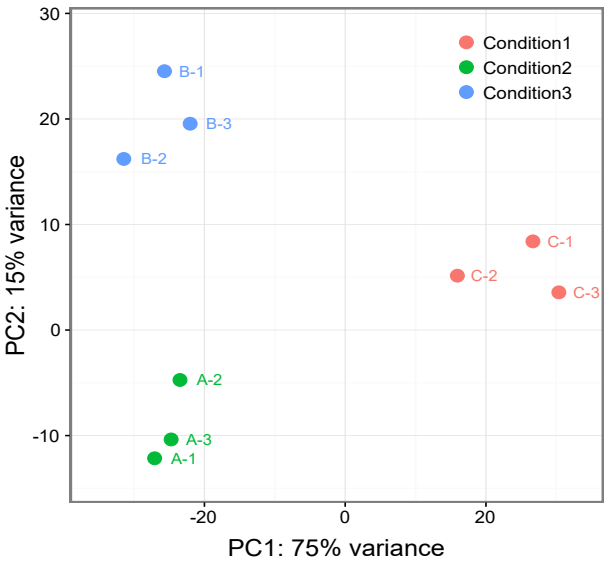
OTU_ID	Sample_1	Sample_2	Sample_3	Sample_4	Sample_5	Sample_6	Domain	Phylum	Class	Order	Family	Genus	OTU_reads	OTU_%
OTU1	16.20	16.50	16.40	19.50	20.30	18.70	Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	36000	17.90
OTU19	20.00	20.20	19.70	13.00	11.30	12.60	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae_1	Bacillus	32500	16.20
OTU2	15.70	15.50	15.70	15.00	14.10	15.30	Bacteria	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	30500	15.20
OTU3	12.50	12.40	12.10	14.50	15.00	14.80	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Salmonella	27200	13.60
OTU4	13.70	13.60	13.30	12.00	12.40	11.50	Bacteria	Firmicutes	Bacilli	Bacillales	Listeriaceae	Listeria	25600	12.80
OTU5	9.12	9.41	9.66	11.50	12.00	11.50	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	EscherichiaShigella	21100	10.50
OTU6	6.07	6.19	6.44	8.65	9.24	9.37	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	15300	7.64
OTU7	6.68	6.29	6.58	5.91	5.77	6.25	Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Enterococcus	12500	6.24

Showing 1 to 8 of 8 entries ← Previous 1 Next

Comparative Statistics

Microsynth's comparative statistical module requires an adequate experimental design including sample replicates and at least two conditions that can be compared (e.g. treatment versus control). To obtain statistically valid results, we recommend using at least three biological replicates per condition. The analysis includes beta diversity calculations which are presented in various graphs including principal component analysis plots (see ***Figure 4A***) and network or tree graphs displaying similarities and differences between the samples. The major outcome of this module is the identification of OTUs that are differentially abundant between the chosen conditions together with appropriate statistical measures. The differentially abundant OTUs are summarized in an interactive table, allowing an easy identification of significant changes in an OTUs abundance (see ***Figure 4B***).

A



B

Condition-1 vs Condition-2 OTU table

10 records per page

OTU_ID	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
OTU1	5960	-0.170	0.0308	-5.54	3.09e-08	4.11e-08
OTU19	5410	0.765	0.0668	11.40	2.45e-30	1.96e-29
OTU2	5080	0.165	0.0428	3.85	1.19e-04	1.19e-04
OTU3	4510	-0.176	0.0258	-6.80	1.06e-11	1.69e-11
OTU4	4260	0.263	0.0269	9.79	1.28e-22	3.41e-22
OTU5	3500	-0.228	0.0281	-8.11	5.25e-16	1.05e-15
OTU6	2540	-0.457	0.0435	-10.50	9.27e-26	3.71e-25
OTU7	2090	0.204	0.0432	4.72	2.34e-06	2.67e-06

Showing 1 to 8 of 8 entries

Figure 4. *Examples for beta diversity results. 4A. Principal component analysis plot to visualize sample clustering. 4B. This detail of a results table shows differential abundance of OTUs between two conditions, including statistical measures for differential abundance (log fold change) and significance (adjusted p-value). baseMean: mean OTU size; lfcSE: standard error; stat: Wald statistic; p-value: Wald test p-value; padj: p-value adjusted for multiple testing.*

Other Analysis Possibilities for Microbial Communities

For the analysis of the functional potential of a microbial community, shotgun metagenomics is the method of choice. We offer a full shotgun metagenomics service providing you with full flexibility regarding the analysis of your results. For further information see our shotgun metagenomics application note. In case you are interested in the actively expressed genes of a microbial community our shotgun metatranscriptomics service offers an appropriate solution.

References

[1] http://www.earthmicrobiome.org/protocols-and-standards/16s/, accessed November 2017

[2] Klindworth et al. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies, Nucleic Acid Res, 41, e1.

[3] White et al. (1990) Amplification and sequencing of fungal ribosomal RNA genes for phylogenetics, in PCR protocols. A guide to methods and applications, San Diego, Academic Press Inc., 315–322.

[4] Edgar (2016) UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon reads. bioRxiv, 081257, doi: https://doi.org/10.1101/081257.