In **metaB**, an additional Polymerase Chain Reaction (PCR) step is performed prior to sequencing (Abdelfattah *et al.*, 2018). It allows to select and amplify a marker gene within the whole DNA (or cDNA, in the case of metaB performed on RNA) by using dedicated primer pairs. Marker genes are chosen according to the taxonomic resolution required: some genes are very well conserved and are used to describe diversity in the whole prokaryotic or eukaryotic community at high taxonomic levels (*e.g.*, the hypervariable regions of respectively 16S and 18S rDNA), while others allow to describe diversity at lower taxonomic levels but the targets are often restricted to a smaller group, targeted with a dedicated primer pair (*e.g.*, species level within a single kingdom) (*e.g.*, the Internal Transcribed Spacers (ITS) of the rDNA) (Breyer and Baltar, 2023). After sequencing, short reads are grouped by similarity into Operational Taxonomy Units (OTUs) or Amplicon Sequence Variants (ASVs) (Jeske and Gallert, 2022) that are used as proxies of families, genera or species depending on the percent of similarity used for clustering and on the marker gene, OTUs or ASVs are in turn given a taxonomy by comparing them to dedicated databases (*e.g.*, SILVA for 16S rDNA (Quast *et al.*, 2013), PR^2 for 18S rDNA (Guillou *et al.*, 2013), UNITE for ITS). For investigation on Fungi, 18S rDNA and ITS are the most commonly used marker genes (Breyer and Baltar, 2023).

While metaG and metaT focus on traits and functions, metaB is usually used to describe the alpha diversity within a sample and the beta diversity between samples. Dedicated programmes allow to infer a function to the taxa identified with metabarcoding (*e.g.*, FUNGuild for Fungi (Nguyen *et al.*, 2016)). As the sequencing depth (*i.e.*, the number of short reads) produced per sample is usually much lower in metaB (~ 10^3-10^4 reads) compared to metaG and metaT (~ 10^6-10^7 reads), metaB is much cheaper and hence more widely used than metaG and metaT. Due to the very high sensibility of RNA to degradation, metaT requires more sophisticated equipments and protocols from sampling to sequencing in order to preserve RNA. This makes metaT harder to implement, notably for sampling in remote places (*e.g.*, open ocean).

HTS data come with two important caveats:

- Sequencers produce a roughly fixed number of reads per run. As a consequence, HTS data are compositional, that is, counts of OTUs, ASVs or contigs must be interpreted as relative abundances of the total number of sequences within the sample and not as absolute counts that would reflect the true abundance in the sample of the organism they originate from (Gloor *et al.*, 2017). As a consequence, HTS data are complementary to but do not replace quantitative, more low-throughput methods (*e.g.*, cell counting with microscopy, flow cytometry).

- Several steps during the processing of samples dedicated to HTS may introduce biases towards specific groups (*e.g.,* against cells with resistant cell walls during DNA/RNA extraction) and distord the true microbial diversity of samples. This is particularly true for metaB, as the PCR step preferentially amplifies the DNA belonging to specific groups (*e.g.*, depending on the primer pair used) (Nichols *et al.*, 2018).

Figure adapted from Martini *et al.*, 2020, by Emile Faure.