Terminology

eDNA – DNA extracted from bulk environmental samples (soil, water, air) w/o isolating target organisms or their parts for the sample.

eDNA assay – investigative procedure in lab to measure presence, amount, & function activity of target entity (analyte)

replication - DNA replication is the process by which DNA makes a copy of itself during cell division. The first step in DNA replication is to 'unzip' the double helix structure of the DNA molecule. This is carried out by an enzyme called helicase which breaks the hydrogen bonds holding the complementary bases of DNA together (A with T, C with G).

PCR - The polymerase chain reaction (PCR) is a technology in molecular biology used to amplify a single copy or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence.

PCR inhibition – interference w/the polymerase reaction causes by an excess of non-target DNA molecules or by non-DNA substances inadvertently extracted w/ the DNA sample. Complete PCR inhibition causes failure to detect target DNA and partial inhibition biases quantification of target DNA

qPCR – **(quantitative real-time polymerase chain reaction)** is a laboratory technique of molecular biology based on the polymerase chain reaction (PCR). It monitors the amplification of a targeted DNA molecule during the PCR, i.e. in real-time, and not at its end, as in conventional PCR. Real-time PCR can be used quantitatively (quantitative real-time PCR), and semi-quantitatively, i.e. above/below a certain amount of DNA molecules (semi quantitative real-time PCR). Involves thermocycled chemical reaction used for targeted detection and quantification of specific nucleic acids during the reaction (in 'real time') based on their nucleotide sequence and a fluorescent reporter dye.

ddPCR - Droplet Digital PCR (ddPCR) is a method for performing digital PCR that is based on water-oil emulsion droplet technology. A sample is fractionated into 20,000 droplets, and PCR amplification of the template molecules occurs in each individual droplet. ddPCR technology uses reagents and workflows like those used for most standard TaqMan probe-based assays.

probe based qPCR - Probe-based qPCR functions by recognition of a specific sequence on the desired PCR product. Unlike SYBR® Green qPCR methods, that use an intercalating dye to bind all double-stranded DNA, probe-based qPCR uses fluorescent-labeled target-specific probes. This technique yields increased specificity and sensitivity since only specific DNA molecules will be labeled. Other fluorescent dyes can also be combined with probe-based qPCR to label, identify, amplify, and quantify various sequences.

Barcoding - is a taxonomic method that uses a short genetic marker in an organism's DNA to identify it as belonging to a particular species. It differs from molecular phylogeny in that the main goal is not to determine patterns of relationship but to identify an unknown sample in terms of a preexisting classification.

Metabarcoding - Metabarcoding is a method of sequencing the DNA barcodes of many different organisms in parallel.

Classic DNA barcoding is of limited use for biodiversity surveys for two main reasons. First, it is too slow and expensive for generating large-scale data across diverse groups (e.g. arthropods other invertebrates) because it requires a separate sequencing reaction for each specimen. Second, it can’t be used on samples that contain DNA from a mixture of related species (e.g. environmental samples).

Metabarcoding solves this problem by using High Throughput Sequencing on community DNA to identify diverse taxa in a single reaction. When metabarcoding is applied to environmental samples, such as water, it is known as **'eDNA metabarcoding'**.

High Throughput Sequencing (HTS) - Thousands to millions of sequencing reactions can be run at once with these approaches, but there are a few drawbacks. HTS has a high error rate, and as HTS generates sequence fragments only, a template must be used - i.e., a known genome sequence to align the sequence fragments to.

However, the benefits of HTS are outstanding. HTS is fast, cheaper than traditional DNA sequencing methods, and generates a massive amount of data. For example, the first human genome sequence generated by the Human Genome Project using Sanger sequencing took over a decade and cost billions of dollars. Sequencing a human genome with HTS methods can cost as little as $1,000 and take only a few

months. This means the number of genomes that can be sequenced has dramatically increased.

cross amplification - is a class of isothermal nucleic acid amplification reactions that utilize multiple primers and probes, one or more of which is a cross primer.

false positive - A result that indicates that a species is present when it is not.

false negative - A result that indicates that a species is not present when it is.

Demultiplexing - Pooling multiple samples increases the efficiency and lowers the cost of **DNA** sequencing. One approach to multiplexing is to use short **DNA** indices to uniquely identify each sample. After sequencing, reads must be assigned in silico to the sample of origin, a process referred to as demultiplexing.

Bioinformatic - the science of collecting and analyzing complex biological data such as genetic codes

internal positive control (IPC) - are simultaneously extracted and/ or amplified in the same tube with the pathogen target and, combined with a positive control, prove the functionality of the reaction mix for correct amplification of the pathogen target

internal amplification control (IAC) - an IAC is a nontarget DNA sequence present in the very same sample tube, which is coamplified simultaneously with the target sequence

SOC – species of concern (rare, invasive, protected)

SMAP-SOC – System for mapping and predicting species of concern

Genome - genetic material of an organism. It consists of DNA (or RNA in RNA viruses). The genome includes both the genes and the non-coding sequences of the DNA/RNA.

Genomics - discipline in genetics that applies recombinant DNA, DNA sequencing methods, and bioinformatics to sequence, assemble, and analyze the function and structure of genomes (the complete set of DNA within a single cell of an organism). Advances in genomics have triggered a revolution in discovery-based research to understand even the most complex biological systems such as the brain.

Intraorganismal eDNA – eDNA contained in living organisms such as microbes, meiofauna (small benthic invertebrates), or microbial larvae, which is protected, active, and can replicate.

Extraorganismal eDNA – eDNA outside of living organisms, such as cellular DNA in shed tissue, shed macrobial cells, and other 'free' DNA molecules from unicellular or multicellular organisms, where it is less protected, inactive, and cannot replicate.

Extramembranous DNA – DNA not bound by cellular, organellar, or viral membranes, synonymous w/the term extracellular DNA in microbial literature.

Massive parallel sequencing or massively parallel sequencing - any of several high-throughput approaches to DNA sequencing using the concept of massively parallel processing; it is also called next-generation sequencing (NGS) or second-generation sequencing. Some of these technologies emerged in 1994-1998 and have been commercially available since 2005. These technologies use miniaturized and parallelized platforms for sequencing of 1 million to 43 billion short reads (50-400 bases each) per instrument run.
Many NGS platforms differ in engineering configurations and sequencing chemistry. They share the technical paradigm of massive parallel sequencing via spatially separated, clonally amplified DNA templates or single DNA molecules in a flow cell. This design is very different from that of Sanger sequencing—also known as capillary sequencing or first-generation sequencing—that is based on electrophoretic separation of chain-termination products produced in individual sequencing reactions.

Gel electrophoresis - technique commonly used in laboratories to separate charged molecules like DNA, RNA and proteins according to their size.; Charged molecules move through a gel when an electric current is passed across it.