

Glossary: terms related to traditional and molecular biodiversity assessment

Created by SLU PhD course PNS0169 and the DNAqua-Net COST Action CA15219

Molecular / bioinformatic terms

BIN (Barcode Index Number)

The BIN combines a pre-checked registry of barcodes with an online database of specimens to identify and support taxonomic placement. By using graph-theory algorithms, specimen sequences are clustered into distinct OTUs that resemble biological species as closely as possible. These clusters are given a distinct number, i.e. BIN, that can then be linked to other studies as well as to other information in BOLD or other systems. BINs are not stable but can fuse as further sequence data is added.

Bioinformatics

Bioinformatics is the use of computational and statistical tools to process and interpret biological data. Bioinformatics typically uses mathematical models that explain the scenario in which the data were collected, how they could be interpreted and which factors may have affected the final fate of the experiment. Different software and pipelines have been developed to apply complex algorithms that allow analyzing and assembling data which is mainly founded in molecular studies such as high throughput sequencing. [1]

BLAST (Basic Local Alignment Search Tool)

BLAST is a fast bioinformatic method that can be run locally or online, to find similarities between a query DNA (or protein) sequence and other sequences that are registered in a database (e.g. NCBI, GenBank) and to align them. The BLAST algorithms include calculations of the statistical significance of the retrieved hit, to determine if the alignment is random or if it actually reflects conservative biology: BLAST indicates an expectation value (e-value) for each aligned sequence pair; if the e-value is extremely small (from 10^{-10} to 10^{-200}) the alignment is pertinent and unlikely to be random. [2]

BOLD Systems (Barcode of Life Data Systems)

BOLD (Barcoding of life data systems) is a database and workbench of DNA barcode data. This database also includes the Barcode Index Number (BIN) information for the deposited barcodes, which consists of similar barcode sequences clustered into units that closely approximate biological species. BOLD also stores further metadata, e.g. information about data preservation (i.e. primers, electropherograms, images and sequences) or collection (collector, habitat, locality, GPS coordinates, date). [3–5]

Clade

A clade or monophyletic group is a group of organisms that are lineal descendants of a common ancestor (e.g. the anthropoid apes and humans), where each of these descendants represents a single branch in the “tree of life”.

Clustering

Clustering is an algorithmic method that attempts to find patterns in a dataset and then groups these patterns into smaller partitions, e.g. based on similarity (OTU clustering, where barcode sequences are grouped according to similarity). Examples are K-Means clustering, Mean-Shift clustering, hierarchical clustering, OPTICS and DBSCAN clustering algorithms.

Community DNA

Community DNA is the whole DNA present in a community, therefore it represents the whole biological diversity within a sample. Community DNA profiles are used in forensic biology and soil investigations because the total microbiome in the sample can be revealed compared to standard culture techniques. [6, 7]

Droplet-digital PCR

Droplet-digital PCR (ddPCR) is a quantitative PCR method that partitions PCR reactions into droplets and PCR amplification is detected by using fluorescence probes. This so-called “third-generation” PCR detects the amplification in each of the thousands of droplets, allowing the direct quantification of the target DNA without a standard curve of the reference. ddPCR is more accurate than qPCR; it can detect differences as low as 1.25-fold, while qPCR can detect differences 2-fold and above. [8]

DNA barcoding

DNA barcoding is a system for species identification focused on the use of a short, standardized genetic region acting as a “barcode” in a similar way that Universal Product Codes (UPCs) are used by supermarket scanners to distinguish commercial products. [9]

DNA metabarcoding

The prefix ‘meta’ refers to the simultaneous collection of barcode sequences from multiple specimens (and species/taxa). DNA metabarcoding is the use of DNA barcoding methods applied to the identification of complex and different environmental samples (such as water, soil, feces or sediments) amplifying short regions of one or few informative genes obtained with HTS. It requires competence in PCR, bioinformatics and biostatistics to analyze sequencing results. [4, 10, 11]

Environmental DNA (eDNA)

Environmental DNA (eDNA) refers to DNA that can be extracted from environmental samples (such as soil, water or air), without first isolating any target organism. This DNA originates from cells of organisms or their waste products (e.g. saliva, urine, feces). In certain cases, DNA extracted from organismal tissue can be seen as eDNA (e.g. for parasite or bacterial communities). [12, 13]

Free DNA

Free DNA in ecology refers to free DNA molecules that can be found in various ecosystems, terrestrial or aquatic. Such molecules are extracellular DNA freely existing outside organisms, either bound to the sediments or carried by water over large distances. Also the term “ExDNA” (extracellular DNA) is used in this context. Free DNA can originate from organisms of any taxon and consequently can be used in detection and monitoring studies. However, when referring to microorganisms (such as bacteria, microalgae etc.), it is not straightforward to separate free DNA from cellular DNA. [14–17]

Haplotype

A haplotype (haploid genotype) is a group of alleles in an organism that are inherited together from a single parent. Haplotypes from the human diploid focus are non-recombining regions of one chromosome.

A haplotype also refers to haploid markers, e.g. mitochondrial or chloroplastid markers, or a set of unique linked single-nucleotide polymorphisms (SNPs) that tend to always occur together.

HPC (High-Performance Computing)

A HPC or supercomputer is particularly powerful mainframe computer with a high level of computational capacity, which uses parallel processing techniques for solving complex computational problems.

HTS (High-Throughput Sequencing)

High-Throughput Sequencing is a set of different sequencing technologies that can produce thousands to millions of DNA sequences from many samples in a short amount of time. They replace older, slower and more expensive (based on costs per base-pair) technologies like Sanger sequencing. An older term for HTS is Next-Generation Sequencing (NGS). [4]

Metagenome

Metagenome is all genetic material of different organisms present in an environmental sample. [18]

Metagenomics

Metagenomics is the analysis of genetic material gathered from the environment, as opposed to analysing single specimens. [6]

Molecular species delimitation (-> GMYC, ABGD, PTP, BINs)

Molecular species delimitation is DNA-based species identification using different methods for defining OTUs that should resemble biological species. Tree-based methods such as GMYC (generalized mixed Yule-coalescent) and PTP (Poisson tree processes) are often used for analysis of single locus, while BPP (Bayesian Phylogenetics Phylogeography) focuses on multiple loci. Distance-based methods include ABGD (Automatic Barcode Gap Discovery) and BINs (Barcode Index Numbers).

MOTU (Molecular Operational Taxonomic Unit)

MOTU is a group of similar organisms based on sequence data comparisons. See also “OTU (Operational Taxonomic Unit)”. [19]

Multiplex-PCR

Multiplex polymerase chain reaction (PCR) is used for amplification of multiple targets in a single PCR. In a multiplexing assay, more than one target sequence can be amplified by using multiple primer pairs in one reaction mixture. As an extension to the practical use of PCR, this technique has the potential to produce considerable savings in time and effort within the laboratory without compromising on the utility of the experiment. There are two types of multiplex PCR:

1. Single Template PCR Reaction

This technique uses a single template which can be genomic DNA along with several pairs of forward and reverse primers to amplify specific regions within a template.

2. Multiple Template PCR Reaction

This technique uses multiple templates and several primer sets in the same reaction tube. Presence of multiple primers may lead to cross-hybridization with each other and the possibility of mis-priming with other templates.

Multiplex PCR has applications in a range of studies, including pathogen identification, high throughput SNP genotyping, mutation analysis, gene deletion analysis, template quantification, linkage analysis, RNA detection and forensic studies. [20]

NGS (Next-Generation Sequencing); also see HTS

Next-generation sequencing (NGS) is a DNA sequencing technology that was introduced in 2005 and replaced previously used Sanger (chain termination) method. NGS revolutionized genomic research by enabling the production of large volumes of sequence data in parallel. The term NGS is now often replaced by High-Throughput Sequencing. [21–23]

PCR (Polymerase Chain Reaction)

PCR was developed by Kary Mullis in 1983. During PCR, a selected DNA fragment (DNA sequence, DNA barcode) is amplified so that an appropriate quantity of this DNA sequence is available for further analysis (e.g. gel-electrophoresis, sequencing). The whole process is based on thermal cycling, which enables the different steps of the whole reaction to take place in order. [24, 25]

PCR requires:

- target DNA fragment
- a pair of short, single stranded oligonucleotides, called primers (forward and reverse primers) that are complementary to the two ends of the target DNA sequence
- DNA-polymerase that is an enzyme carrying out the DNA replication by using the primers as the template to start
- free nucleotides (dNTPs) as the building blocks for the new copies of target DNA fragment
- reaction medium (buffer or water) including Magnesium ions

Steps of PCR:

- denaturation: at ~96°C the two strands of DNA separate from each other
- annealing: at lower temperature (50–65°C, depending on primer GC content and length) the primers bind to their complementary sequence on the single stranded DNA
- extension: usually at 67–72°C (depending on polymerase) the DNA-polymerase enzyme starts the synthesis of the new DNA strand starting from the primer using available free nucleotides

Phylogenetics (vs. phylogenomics)

Phylogenetics is the study of evolutionary relatedness among groups of organisms. Classic phylogenetics deals mainly with physical or morphological features. More recent advances have led to the development of modern molecular phylogenetics, which uses DNA or amino acid sequence data to infer these relationships by comparing homologous sequences.

Phylogenomics (vs. phylogenetics)

Phylogenomics can be considered as a part of phylogenetics, referring to the analysis of genome data – large sets of sequence data – to infer evolutionary relationships between the sequences and the organisms they are derived from based on the degree of genomic similarity.

Using large data sets overcomes the problem of stochastic (or sampling) errors encountered when reconstructing phylogenies from single or few nucleotide or amino acid sequences. [26]

PlutoF

PlutoF is a web-based platform for storing and managing biodiversity data, which provides database and computing services for taxonomic, ecological and genetic research. [27, 28]

Primer

A primer is a short single strand of RNA or DNA of approximately 20 nucleotide bases. A primer defines the region of the starting point for DNA synthesis after separation of double-stranded DNA. These bases are needed for the DNA polymerase to start the replication reaction based on the opposite DNA strand.

Quantitative PCR (qPCR)

Quantitative polymerase chain reaction (qPCR) or real-time PCR is a molecular technique that follows the PCR workflow incorporating the monitoring of the DNA amplified at each cycle of PCR. qPCR is able to determine the absolute or the relative quantities of a specific DNA sequence. [29–33]

Sequence Read Archive (SRA)

Sequence Read Archive (SRA) makes biological sequence data available for all researchers. It is one of the 35 NCBI (National Center for Biotechnology Information) databases. It stores raw sequencing HTS (High-Throughput Sequencing) data and alignment information. It does not contain taxonomic information. SRA accepts data from all kinds of sequencing projects including clinically important studies that involve human subjects or their metagenomes. [34]

Shotgun-sequencing

Shotgun sequencing is a method for determining the sequence of long DNA strands, based on randomly breaking up DNA into numerous small fragments that can be sequenced separately. Sequenced fragments are then reassembled into the complete sequence by finding overlapping regions. This was one of the early technologies that enabled full genome sequencing. [35–37]

Species distribution models (SDMs)

SDMs use modeling software to estimate the likely occurrence (or likely absence) of taxa over geographic space and time using environmental data, i.e. the similarity of conditions at an unknown site compared to the conditions at a known site and presence/absence data of taxa. Species distribution modeling is also known under other names including climate envelope-modeling, habitat modeling and (environmental or ecological) niche-modeling. [38, 39]

Tag-switching

Tags are short nucleotide stretches used to individually label amplicon libraries. Typically, one tag-combination is used for one sample. Tag-switching, also known as tag-jumping, refers to the phenomenon when tags occur in combinations not used causing potential errors that you need to consider when interpreting your NGS results.

Tag-switching can occur if tags are added in the final PCR during the building of your library and the samples are then pooled prior sequencing. After the tagging PCR, individual samples might contain single-stranded amplicons, and when the samples are pooled these may form heteroduplexes thereby mixing different tags, making tag-switching likely to occur. Tag-switching can also occur from simple cross-contamination or during the sequencing process and therefore it is good to have long tags.

The likelihood of tag-switching is typically about 1 per 1000 reads. You can identify the places of tag-switching by adding tagging PCR replicates and then account for it in your data analysis. [40]

Other terms / ecological terms

Alpha-diversity (vs. beta- and gamma-diversity)

These terms refer to the ecological description of species diversity on different scales in ecological communities. Alpha deals with the smallest scale, the local species diversity, and the other terms refer to larger ecological units (beta = regional, gamma = total in the landscape). Alpha diversity measures the mean species diversity within local habitats.

Assemblage (vs. community)

An assemblage is a group of taxonomically related organisms from different species that share similar habitat requirements. It is also considered as a subset of the total species that coexist in a community and occur together in space and time. Assemblage is different from communities; the latter is used to describe the total set of organisms that share the same ecological habitat without any distinction about their taxonomic relationship. [41]

Benthic

The term benthic is used when relating to the bottom of a body of water, for example a lake bottom, from the shore to the deep parts. It is also used when relating to, or occurring in the depths of the ocean. [42, 43]

Benthos

Benthos is a community of organisms that inhabit the benthic zone of freshwater and marine systems. [44, 45]

Beta-diversity (vs. alpha- and gamma-diversity)

In ecology, beta diversity (β -diversity or true beta diversity) is the ratio between regional and local species diversity. It suggests how diverse given communities are when compared to each other. It is used to estimate gamma diversity (γ -diversity) which is the total species diversity in a landscape together with alpha diversity (α -diversity), which is the mean species diversity at the habitat level.

Bioassessment

Bioassessment refers to a process of evaluating the ecological integrity of both terrestrial and aquatic environments by measuring characteristics of organisms or assemblages of organisms that inhabit those environments.

For aquatic environments, bioassessment refers to the assessment of ecological integrity of a waterbody by measuring attributes of the assemblage of organisms inhabiting the waterbody. In conjunction with biological measurements, bioassessment of aquatic environments usually includes measurements of instream and riparian zone physical habitat constituents. Common assemblages of aquatic organisms used for bioassessment include fish, macroinvertebrates and algae. [46]

Bioindicator

Bioindicator is a term for an organism restricted to one or a few habitat types that potentially represent a better ecological indicator of environmental change than a habitat generalist. Organisms of that kind are used to monitor environmental changes and assess the impacts of disturbance on an ecosystem. [47, 48]

Biomonitoring

Biomonitoring is the practice of recording species diversity and abundance across different locations and times. The comparison of data with reference datasets is used to observe biological responses to influences and to assess the ecological condition and changes in the environment. The aim of biomonitoring is to assess the ecological status and to characterize the quality of biological environment. [49–51]

Brackish water

Brackish water is a broad term used to describe water which salinity is between that of fresh and marine water, and these are often transitional areas where such waters mix. An estuary, which is the part where river meets the sea, is the best-known example of brackish water. [52]

Calcareous

Calcareous refers to the principal composition of calcium salts (calcium carbonate), containing lime or its derived soil. The term is used in different scientific disciplines. [44, 53]

Conductivity

Conductivity is a measure of the ability of water to conduct electricity. It increases as the amount of dissolved minerals (ions) increases. These conductive ions originate from dissolved salts and inorganic materials such as alkalis, chlorides, sulfides and carbonate compounds.

Current

Current is a movement in a body of water caused by major ocean circulation or tides, by waves along shorelines or by gravity-induced flow in rivers. [44]

Community (vs. assemblage)

Community refers to all organisms, i.e. populations of two or more different species that interact in a common area, while assemblage refers to taxonomically related groups of species populations. [40]

Cryptic species

Cryptic species is a species in which the diagnostic morphological characters are not easily perceived (or absent) and that do not hybridize under normal conditions.

Detritivore

Detritivore is an animal that eats organic material (detritus) found in sediments, originating from dead and decaying plant and animal matter. [44, 54]

Ecosystem

An ecosystem is a dynamic complex in the environment between biotic and abiotic factors, which are linked together through nutrient cycles and energy flows. [44, 55]

Environment (vs. habitat)

Environment is defined as the physical, chemical and biological components that surround an organism. The difference between environment and habitat is that habitat is an environment that is specific to the certain organism or group of organisms, and contains all the components required by the organisms. [6]

EPT

EPT is an acronym for the collective reference to the insect orders Ephemeroptera, Plecoptera and Trichoptera. [4]

Eutrophication

Eutrophication is a consequence of nutrient overload in aquatic ecosystems usually noticed by the bloom of opportunistic green algae or cyanobacteria.

False-positive

False-positive is a result indicating that a given condition is present when it is not. It is analogous to a type I error in statistics. [56]

False-negative

A false negative error is a type II error occurring when the absence or non-existence of something is falsely inferred when it is real or does exist (i.e. a conclusion based on false information). [57]

Gamma-diversity (vs. alpha- and beta-diversity)

Gamma diversity (γ -diversity) represents the overall species diversity of a range of habitats or communities within a region. It is determined by the number of species occupying individual habitats (alpha diversity, α) and by the rate of change in species composition across habitats or among communities (beta diversity, β). The term was introduced by R. H. Whittaker in 1960. [58–60]

Habitat (vs. environment)

Habitat refers to all the environmental conditions, both biotic and abiotic, in which a species lives and reproduces. Habitat is a part of the environment that encompasses all the biotic and abiotic components in which several habitats exist. [61, 62]

Intercalibration

Intercalibration is a procedure or state achieved by a group of researchers or laboratories engaged in collecting and analyzing a certain type of data (such as in a monitoring program) in which they produce and maintain compatible data outputs.

Intercalibration overcomes the problems resulting from methodological differences (use of different field and laboratory protocols) and allows direct comparison of the results. If the sampling and analytical methods used by different researchers are not identical, results obtained from different protocols must be standardized (i.e. intercalibrated) to allow for meaningful comparison of the data.

Invasive (vs. non-native / non-indigenous)

Invasive alien species (IAS) are defined as non-indigenous species that, once introduced into an environment, successfully overcome the acclimatization period, establish a self-sustaining population and show good capacity of expansion in the new environment. They cause ecological, economic or social damages and represent a big problem for the local biodiversity. [63]

Kick-sampling

Kick-sampling is a technique used to sample benthic invertebrates in a stream or a river, in which a net is held underwater and the substrate upstream of the net is disturbed by kicking for a predetermined period of time. In order to sample different habitats in the stream, it is important to move around the site during sampling. [64, 65]

Lentic (vs. lotic)

Lentic is a term used to describe a body of water composed by standing water such as lakes, ponds, seasonal pools, ditches and seeps. Lentic systems are characterized by slow or non-existent movement of water. This makes the ecosystem stratified due to different light and oxygen availability according to the depth of the water, so the biotic elements found in each layer are often different. [66]

Lotic (vs. lentic)

Lotic is a term used to describe continuously flowing water bodies such as rivers and streams, from rapid torrent to slow-moving waters. These environments are usually shallower than lentic ones (such as lakes and ponds), with water flow and temperature as key abiotic factors for the ecology. These environments are also greatly affected by seasonal precipitation and snow melting.

Macrohabitat

Macrohabitat is a habitat big enough to contain multiple environments within it and support multiple types of organisms.

Macroinvertebrates

Macroinvertebrates are macroscopic invertebrates that are large enough to be seen with the naked eye (>0.5 mm). For aquatic macroinvertebrates, at least one life stage is bound to water (e.g. streams and rivers, lakes, groundwater or marine environments). [4]

Marine Strategy Framework Directive (MSFD)

The Marine Strategy Framework Directive adopted in 2008 aimed to protect the marine environment across Europe more effectively. The Commission produced a set of detailed criteria and standard methodology to help achieve Good Environmental Status (GES) in the EU's marine waters by 2020 and to protect the resource base upon which marine-related economic and social activities depend. The directive requires the use of several indicators including species diversity, seafloor integrity, food web structure as well as non-indigenous and commercial species. [67]

Meiofauna

Meiofauna are organisms with intermediate size between the microbes and macrofauna. The term *meio* is a Greek term for “smaller”. Meiofauna are generally classified as protists and invertebrates between 50

and 1000 µm, although some researchers use a 500 µm upper size limit. Meiofauna occur in all aquatic ecosystems (marine to freshwater) and include representatives from two-thirds of the known animal phyla. [68, 69]

Microhabitat

Microhabitat is a small or limited habitat differentiated from extensive adjacent habitat by distinct flora and fauna and environmental characteristics, such as substrate type, water velocity, light, temperature or depth. [44, 70, 71]

Mock-community

Mock-communities are constructed communities of mixes of individuals of known identity or taxonomic composition, used as positive controls or to perform pilot studies.

Multi-habitat sampling

Multi-habitat-sampling is a sampling technique that aims at sampling all habitats encountered at an aquatic sampling site related to its proportion at the site. A typical example in the context of the WFD: initially, 20 sampling units are taken, each of which represents 5 % substrate coverage of the bottom of the waterbody and pooled into one multi-habitat sample. [44]

Nagoya protocol

The Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity, also known as the Nagoya Protocol on Access and Benefit Sharing (ABS), is a 2010 supplementary agreement to the 1992 Convention on Biological Diversity (CBD). Its aim is the implementation of one of the three objectives of the CBD: the fair and equitable sharing of benefits arising out of the utilization of genetic resources, thereby contributing to the conservation and sustainable use of biodiversity.

However, there are concerns that the added bureaucracy and legislation will, overall, be damaging to the monitoring and collection of biodiversity, to conservation, to the international response to infectious diseases and to research. [72]

Non-indigenous (vs. invasive)

Non-indigenous species (NIS), also called non-native, alien or exotic species, are species introduced intentionally or accidentally by anthropogenic disturbance outside their original, past or present, area of distribution, that might survive and subsequently reproduce. NIS might eventually become invasive alien species if they become sufficiently abundant, causing adverse effects on native plants and animals. [73, 74]

Non-native (vs. invasive)

See “Non-indigenous”.

OTU (Operational Taxonomic Unit)

Operational taxonomic unit (OTU; also molecular operational taxonomic unit, MOTU) is a concept used to classify and group organisms based on similarities of their DNA sequence within the selected marker region of the genome. [75]

Pelagic

Pelagic organisms are micro- and macroorganisms of flora and fauna that live in the water column of lakes, streams and oceans.

Phytobenthos

Phytobenthos refers to photosynthetic organisms (plants, algae and some prokaryotes) living at the bottom of aquatic environments. [44, 76]

Rare species

Rare species are those that are represented by only a few individuals or restricted to particular habitats to a level that is demonstrably less than the majority of other organisms of comparable taxonomic entities. The actual rarity cutpoint is a subjective decision. [77]

Saprobic Index

Saprobic index is one of the biological indicators for water quality assessment. It was introduced by Pantle and Buck in 1955 and it is used to measure the impact of organic pollution. Saprobic indices are based on a selection of aquatic macroinvertebrate taxa or aquatic microorganisms that are proven to be either sensitive or tolerant to organic pollution. Several saprobic indices are currently used in Europe, e.g. the Austrian, Czech and German indices. [44, 78–80]

Shannon Diversity Index

An ecological diversity index is a quantitative measure that reflects how many different species there are in a community, and how evenly the individuals are distributed. The Shannon index has been a popular diversity index in the ecological literature, where it is also known as Shannon's diversity index, the Shannon–Wiener index, the Shannon–Weaver index or the Shannon entropy. The measure was originally proposed by Claude Shannon. [81]

Species (concept)

Species concept deals with the definition of the term “species” and the delimitation criteria for species that have been continuously changing through time, but it also depends on the field of study where it is used. Species concept is one of the most debated topics in biology. Today there are at least 26 different species concepts. Maybe the simplest definition is that species is the basic unit of classification of living organisms. It is often referred as the largest group of organisms in which two individuals of the appropriate sexes or mating types can engage in sexual reproduction and produce offspring. This definition is exclusive for organisms that reproduce sexually, but does not fit, for instance, to bacteria. [82–84]

Taxonomic Assignment

Taxonomic assignment is one of the principal tasks in analyzing sequences obtained from samples containing DNA from unknown organisms (e.g. environmental DNA). Taxonomic assignment is the identification of the taxonomic affiliation of DNA sequences in the sample, preferably to the species or genus level.

Sequence reads from a sample are compared (aligned) to a set of sequences in reference databases whose species/taxon assignment is truly known. A sequence read can be equally similar to more than one reference sequence, and these ambiguities are then solved by assigning the read to a consensus sequence, such as the lowest common ancestor of all the candidate sequences in a reference taxonomy or some more specific candidate reference sequence.

Reference databases and taxonomies are essential to determine the phylogenetic affiliation of sequence reads. The precision of identification depends on the number of identified sequences in the reference databases.

In the absence of a reference database, or in the presence of environmental sequences from unknown organisms, taxonomic assignment is not possible and sequence reads are grouped into clusters of related species, usually called operational taxonomic units (OTUs).

Water Framework Directive (WFD)

The Water Framework Directive 2000/60/EC is an EU directive that commits European Union member states to achieve good qualitative and quantitative status of all water bodies (including marine waters up to one nautical mile from shore) by 2015. [85]

References

1. Wikipedia contributors. Bioinformatics. In Wikipedia, The Free Encyclopedia. Retrieved from <https://en.wikipedia.org/wiki/Bioinformatics>
2. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403–410. doi:10.1016/S0022-2836(05)80360-2.
3. www.boldsystems.org
4. <http://dnaqua.net/glossary>
5. Wikipedia contributors. Barcode of Life Data System. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Barcode_of_Life_Data_System
6. <https://bioportal.bioontology.org/ontologies/>
7. Horswell, J., Cordiner, S. J., Maas, E. W., Martin, T. M., Sutherland, K. B., Speir, T. W., Nogales, B., Osborn, A. M. (2002). Forensic comparison of soils by bacterial community DNA profiling. *Journal of Forensic Sciences*, 47(2), 350–353.
8. Doi, H., Uchii, K., Takahara, T., Matsushashi, S., Yamanaka, H., Minamoto, T. (2015). Use of Droplet Digital PCR for Estimation of Fish Abundance and Biomass in Environmental DNA Surveys. *PLoS ONE*, 10(3).
9. Hellberg, R. S., Pollack, S. J., Hanner, R. H., in *Seafood Authenticity and Traceability*, 2016.
10. <https://www.allgenetics.eu/index.php/services/dna-metabarcoding.html>
11. Staats, M., Arulandhu, A. J., Gravendeel, B., Holst-Jensen, A., Scholtens, I., Peelen, T., Prins, T. W., Kok, E. (2016). Advances in DNA metabarcoding for food and wildlife forensic species identification. *Analytical and Bioanalytical Chemistry*, 408(17), 4615–4630.
12. Taberlet, P., Coissac, E., Hajibabaei, M., Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, 21(8), 1789–1793.
13. Barnes, M. A., and Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, 17, 1–17. doi:10.1007/s10592-015-0775-4.
14. Turner, C. R., Uy, K. L., Everhart, R. C. (2015). Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, 183, 93–102.
15. Torti, A., Lever, M. A., Jørgensen, B. B. (2015). Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Marine Genomics*, 24, 185–196.
16. Deiner, K., Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS ONE*, 9(2).
17. Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéoz-Perret-Gentil, L., Beja, P., Boggero, A., ... & Feio, M. J. (2018). The future of biotic indices in the ecogenomic era: Integrating (e) DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment*, 637, 1295–1310.
18. <https://en.wiktionary.org/wiki/metagenome>
19. http://www.nematodes.org/research/barcoding/motu_defined.html
20. http://www.premierbiosoft.com/tech_notes/multiplex-pcr.html
21. Metzker M.L. (2010) Sequencing technologies — the next generation, *Nature Reviews Genetics*, 11, 31–46.

22. Behjati S., Tarpey P.S. (2013) What is next generation sequencing? *Archives of Disease in Childhood: Education & Practice*, 98(6): 236–238.
23. Van Dijk E.L., Auger H., Jaszczyszyn Y., Thermes C. (2014) Ten years of next-generation sequencing technology, *Trends in Genetics*, 30(9), 418–426
24. Mullis, K. B., Erlich, H.A., Arnheim, N., Horn, G.T., Saiki, R.K., Scharf, S.J. (1989) Process for amplifying, detecting, and/or-cloning nucleic acid sequences. U.S. Patent 4800159.
25. Wikipedia contributors. (2019, March 21). Polymerase chain reaction. In Wikipedia, The Free Encyclopedia. Retrieved 19:51, March 23, 2019, from https://en.wikipedia.org/w/index.php?title=Polymerase_chain_reaction&oldid=888806498
26. Chan, C. X., & Ragan, M. A. (2013). Next-generation phylogenomics. *Biology Direct*, 8, 3.
27. Abarenkov, K., Tedersoo, L., Nilsson, R. H., Vellak, K., Saar, I., Veldre, V., Parmasto, E., Proust, M., Aan, A., Ots, M., Kurina, O., Ostonen, I., Jõgeva, J., Halapuu, S., Põldmaa, K., Toots, M., Truu, J., Larsson, K. H., Kõljalg, U. (2010). PlutoF – a Web Based Workbench for Ecological and Taxonomic Research, with an Online Implementation for Fungal ITS Sequences. *Evolutionary Bioinformatics*, 6, 189 – 196
28. PlutoF user manual [PDF file]. PlutoF. Retrieved on March 21, 2019 from https://plutof.ut.ee/assets/varia/manuals/docs/plutof_manual_3_en.pdf
29. <https://www.sigmaaldrich.com/life-science/molecular-biology/molecular-biology-products.html?TablePage=9620611>
30. <http://www.bio-rad.com/es-es/applications-technologies/what-real-time-pcr-qpcr?ID=LUSO4W8UU>
31. <http://www.enzolifesciences.com/science-center/technotes/2017/march/what-are-the-differences-between-pcr-rt-pcr-qpcr-and-rt-qpcr?/>
32. <https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/real-time-polymerase-chain-reaction>
33. Wikipedia contributors. Real-time polymerase chain reaction. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Real-time_polymerase_chain_reaction
34. <https://www.ncbi.nlm.nih.gov/sra>
35. Wikipedia contributors (2019, February 17). Shotgun sequencing. In Wikipedia, The Free Encyclopedia. Retrieved on March 21, 2019, from https://en.wikipedia.org/w/index.php?title=Shotgun_sequencing&oldid=883830254
36. What is shotgun sequencing? Yourgenome. Retrieved on March 21, 2019, from <https://www.yourgenome.org/facts/what-is-shotgun-sequencing>
37. Medical Definition of Shotgun sequencing. MedicineNet. Retrieved on March 21, 2019, from <https://www.medicinenet.com/script/main/art.asp?articlekey=20667>
38. https://rspace.org/sdm/1_sdm_introduction.html
39. https://en.wikipedia.org/wiki/Species_distribution_modelling
40. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/1755-0998.12402>
41. Stroud, J. T., Bush, M. R., Ladd, M. C., Nowicki, R. J., Shantz, A. A., Sweatman, J. (2015). Is a community still a community? Reviewing definitions of key terms in community ecology. *Ecology and Evolution*, 5(21), 4757–4765.

42. <https://www.merriam-webster.com/dictionary/benthic>
43. Wikipedia contributors. Benthic zone. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Benthic_zone
44. <http://www.freshwaterplatform.eu/index.php/glossary.html>.
45. Wikipedia contributors. Benthos. In Wikipedia, The Free Encyclopedia. Retrieved from <https://en.wikipedia.org/wiki/Benthos>
46. BioAssessment Services (BAS):
http://www.jtkbioservices.com/PDFs/Bioassessment_overview_6_20_11_mm_tk.pdf
47. De Cáceres, M., Legendre, P., Moretti, M. (2010). Improving indicator species analysis by combining groups of sites. *Oikos*, 119: 1674–1684
48. Carignan, V., and Villard, M. A. (2002). Selecting Indicator Species to Monitor Ecological Integrity: A Review. *Environmental Monitoring and Assessment*, 78(1), 45–61.
49. https://www.norman-network.net/?q=node/9#pos_M
50. <https://www.sciencedirect.com/topics/earth-and-planetary-sciences/biomonitoring>
51. <http://www.kvvm.hu/szakmai/karmentes/kiadvanyok/karmutmutato6/202-02.htm>
52. Rich, V. I., Maier, R. M. in *Environmental Microbiology* (Third Edition), 2015.
53. Wikipedia contributors. Calcareous. In Wikipedia, The Free Encyclopedia. Retrieved from <https://en.wikipedia.org/wiki/Calcareous>
54. Dodson, S. (2005). *Introduction to Limnology*. California: McGraw-Hill Education.
55. Worster D. (1994). *Nature’s Economy: A history of Ecological Ideas*. 2nd edition. Cambridge; New York: Cambridge University Press.
56. Wikipedia contributors. False positives and false negatives. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/False_positives_and_false_negatives
57. Wikipedia contributors. Type I and type II errors. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Type_I_and_type_II_errors
58. Whittaker, R. H. (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, 30, 279–338.
59. Wikipedia contributors. Gamma diversity. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Gamma_diversity
60. <http://www.oxfordreference.com/view/10.1093/oi/authority.20110803095842382>
61. Wikipedia contributors. (2019, March 21). Natural environment. In Wikipedia, The Free Encyclopedia. Retrieved 20:37, March 24, 2019, from https://en.wikipedia.org/w/index.php?title=Natural_environment&oldid=888777080
62. Wikipedia contributors. (2019, February 25). Habitat. In Wikipedia, The Free Encyclopedia. Retrieved 20:38, March 24, 2019, from <https://en.wikipedia.org/w/index.php?title=Habitat&oldid=884957304>
63. <https://www.iucn.org/theme/species/our-work/invasive-species>
64. Wikipedia contributors (2018, September 28). Kick sampling. In Wiktionary, The Free Dictionary. Retrieved on March 21, 2019 from https://en.wiktionary.org/w/index.php?title=kick_sampling&oldid=50399636

65. Collecting freshwater macroinvertebrate samples [PDF file]. Freshwater Biological Association. Retrieved on March 21, 2019 from <https://www.fba.org.uk/sites/default/files/CourseInvertSamplingProtocol.pdf>
66. Marsh G.A., Fairbridge R.W. (1999) Lentic and lotic ecosystems. In *Environmental Geology. Encyclopedia of Earth Science*. Springer, Dordrecht.
67. http://ec.europa.eu/environment/marine/eu-coast-and-marine-policy/marine-strategy-framework-directive/index_en.htm
68. Hakenkamp, C. C., Palmer, M. A., in *Streams and Ground Waters*, 2000.
69. Stevenson, R. J., Bothwell, M. L., Lowe, R. L., in *Algal Ecology*, 1996.
70. <https://www.dictionary.com/browse/microhabitat>
71. <https://www.bing.com/search?q=microhabitats+definition&FORM=QSRES>
72. <https://www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf>
73. http://ec.europa.eu/environment/marine/good-environmental-status/descriptor-2/index_en.htm
74. Lehtiniemi, M., Ojaveer, H., David, M., Galil, B., Gollasch, S., McKenzie, C., Minchin, D., Occhipinti-Ambrogi, A., Olenin, S., Pederson, J. (2015). Dose of truth—Monitoring marine non-indigenous species to serve legislative requirements. *Marine Policy*, 54, 26–35.
75. http://www.nematodes.org/research/barcoding/motu_defined.html
76. http://wiki.reformrivers.eu/index.php/Category:Macrophytes_and_phytobenthos
77. Mouillot, D., Bellwood, D. R., Baraloto, C., Chave, J., Galzin, R., Harmelin-Vivien, M., Kulbicki, M., Lavergne, S., Lavorel, S., Mouquet, N., Paine, C. E. T., Renaud, J., Thuiller, W. (2013) Rare Species Support Vulnerable Functions in High-Diversity Ecosystems. *PLoS Biology*, 11(5).
78. Pantle, R., Buck, H. (1955). Die biologische uberwachung der Gewasser und die Darstellung der Ergebnisse. *Gas- und Wasserfach*, 96, 604.
79. <https://www.eionet.europa.eu/gemet/en/concept/7449>
80. http://wgbis.ces.iisc.ernet.in/energy/lake2002/proceedings/5_1.html
81. Wikipedia contributors. Diversity index. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Diversity_index
82. Wikipedia contributors. (2019, March 17). Species. In Wikipedia, The Free Encyclopedia. Retrieved 20:59, March 24, 2019, from <https://en.wikipedia.org/w/index.php?title=Species&oldid=888187004>
83. Wikipedia contributors. (2019, March 11). Species concept. In Wikipedia, The Free Encyclopedia. Retrieved 20:59, March 24, 2019, from https://en.wikipedia.org/w/index.php?title=Species_concept&oldid=887280863
84. Wilkins, John S. (1 October 2006). "A List of 26 Species Concepts". Science Blogs.
85. Wikipedia contributors. Water Framework Directive. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Water_Framework_Directive