#### Learning objectives:

Specific emphasis should be placed on the process used to find the answer. Be as comprehensive as possible e.g. provide URLs for web sources, literature citations, etc.  
*(Reminders for how to format links, etc in RMarkdown are in the RMarkdown Cheat Sheets)*

#### Specific Questions:

* How many prokaryotic divisions have been described and how many have no cultured representatives (microbial dark matter)?
  + At least 89 bacterial and 20 archaeal phyla are recognized via small subunit ribosomal RNA databases, although the true phyla count is certainly higher and could range up to 1,500
    - As there are prokaryotes that live in the “shadow biosphere” –> which is a hypothetical microbial biosphere unknown to life
  + 26 of the approximately 52 identifiable major phyla, within the domain Bacteria have cultivated representatives
    - Thus, 52-26 of the major phyla of Bacteria are uncultured
  + Point is most of the life is uncultured. Only information we have about life is from seqeuncing.
  + Specific references for the question above:
    - Solden, L, Lloyd, K, Wrighton, K. 2016. The bright side of microbial dark matter: lessons learned from the uncultivated majority. Curr. Opin. Microbiol. 31:217-226. doi: 10.1016/j.mib.2016.04.020.
    - Youssef, NH, Couger, MB, McCully, AL, Criado, AEG, Elshahed, MS. 2015. Assessing the global phylum level diversity within the bacterial domain: A review. Journal of Advanced Research. 6:269-282. doi: 10.1016/j.jare.2014.10.005.
    - Rappé, MS, Giovannoni, SJ. 2003. THE UNCULTURED MICROBIAL MAJORITY. Annual Reviews in Microbiology. 57:369-394. doi: 10.1146/annurev.micro.57.030502.090759.
* How many metagenome sequencing projects are currently available in the public domain and what types of environments are they sourced from?
* Shot-gun metagenomics:
  + Assembly: EULER, IMG -/M
  + Binning: S-GCOM, IMG-RAST
  + Annotation: KEGG, NCBI
  + Analysis: Megan 5
* Marker Gene Metagenomics:
  + Standalone software:OTU base
  + Analysis: SILVA
  + Denoising: Amplicon Noise
  + Datapases: Ribosomal Database Project (RDP) rences: from a paper, review artile
* What types of on-line resources are available for warehousing and/or analyzing environmental sequence information (provide names, URLS and applications)?
  + IMG/-M (<https://img.jgi.doe.gov/m/>)
  + MEGAN (<https://ab.inf.uni-tuebingen.de/software/megan/>)
  + MetaPathways (<https://hallam.microbiology.ubc.ca/MetaPathways/>)
* What is the difference between phylogenetic and functional gene anchors and how can they be used in metagenome analysis?
  + Phylogenetic:
    - veritcal gene transfer
    - carry phylogenetic info
    - taxonomic
    - ideally single copy
  + Functional:
    - more horizontal gene transfer
    - identify specific biogeochemical functions associated with measureable effects
    - not as useful for phylogentic construction
* What is metagenomic sequence binning? What types of algorithmic approaches are used to produce sequence bins? What are some risks and opportunities associated with using sequence bins for metabolic reconstruction of uncultivated microorganisms?
  + Binning: process of grouping sequences that comes from a single genome
  + Types of algorithms:
  1. Align sequences to database
  2. Group to each based on DNA characteristics: GC content, codon usage
  + Risks & Oppurtunities: Risks:
  + incomplete coverage of genome sequences
  + contamination of different phylogeny. Questions to ask what is considered a contamination?
    - Threshold should be be minimum 10% for contamition
  + What is a genome from a metagenome?
* Is there an alternative to metagenomic shotgun sequencing that can be used to access the metabolic potential of uncultivated microorganisms? What are some risks and opportunities associated with this alternative?
  + functional screens (biochemical, etc)
  + 3rd gene seqeuecing (nanopore)
  + Single sequencing
  + FISH probe