The sequencing and assembly of entire genomes

The genome holds all information that is necessary for development and function of an individual organism, and for its reproduction. In addition, genome sequences provide invaluable information about the evolutionary history of species and their positions in the tree of life (Mardis 2008). Ultimately, comparative genomics approaches aim at integrating each species into a network of organismal evolution. It is thus that sequencing the genomes of one to several individuals as representatives of an entire species has become a routine task (Fig. 1 & Fig. 2), and entire consortia are being formed to investigate global biodiversity on the genomic level (e.g. LOEWE-TBG, or the Earth BioGenome Project (EBP)). The first step in comparative genomics studies is, however, to decipher the genome sequence of the species of interest.

Figure 1: In 2017, the first large scale projects were proposed to sequence all life on earth. By that time, very few completed genomes (red lines in inner rim) for named eukaryotic species existed (green), but many more among bacteria (blue) and archaea (purple). Among the millions of eukaryotic species, there are even relatively few lower resolution genome sequences (blue, light and dark gray).

Figure 2: In the course of 4 years only, the number of genome assemblies finished or ongoing increased to more than 1.5 Million. Among these, there are more than 8,000 species with an assembly at least on scaffold level (Values taken from GoaT; Status: October 2021)

- DNA sequencing
- Shotgun assembly
- Shortcut to task list
- Simulation of shotgun reads (Supplementary information)

From:
https://fsbioinf.biologie.uni-frankfurt.de/teaching/wiki/ - **Teaching**

Permanent link:
https://fsbioinf.biologie.uni-frankfurt.de/teaching/wiki/doku.php?id=general:bioseqanalysis:genome_sequencing

Last update: **2021/10/20 12:10**