Metabarcoding analysis of Wild dog dung

The client wishes to identify species consumed by wild dogs by isolating DNA from wild dog scat and selectively PCR amplifying the mitochondrial 12S rRNA gene from these samples and sequencing it on an IonTorrent sequencer. The amplicon sequences are then compared to a database of known 12S sequences and the origin of the DNA deduced.

A crucial step in the bioinformatic analysis of these amplicon sequences is the identification of representative sequences. This can be achieved using a clustering approach or by denoising raw sequencing reads. DADA2 is a widely adopted algorithm, released as an R library, that denoises marker-specific amplicons from next-generation sequencing and produces a set of representative sequences referred to as 'Amplicon Sequence Variants' (ASV).

ASVs sequences are compared to a reference database to verify the taxonomic composition using the "Blastn" function of the program Blast+ (Camacho et al., 2009) for the 10 best hits and an e-value of < 0.001. ASVs blasted > 98% similarity will be considered the matched species.

We proposed to use DADA2 (Callahan et al., 2016) in this analysis.

Output

List of species identified by DADA2

Units required

40 hours

References

Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-Resolution Sample Inference from Illumina Amplicon Data. Nat. Methods **2016**, 13, 581–583.

C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T.L. Madden. BLAST+: architecture and applications. BMC Bioinformatics, 10 (2009), p. 421, 10.1186/1471-2105-10-421