**Arthropod CO1 Sequencing Protocol**

**Updated August 2016**

**Jonah Ventures**

**gDNA extraction**

0.25g of sample is used for gDNA extraction using MoBio PowerSoil Kit (cat#12955) according to the manufacturer’s protocol

**PCR**

**Primers used for Arthropod CO1 sequencing**

Forward – 5’ AGATATTGGAACWTTATATTTTATTTTTGG 3’

Reverse – 5’WACTAATCAATTWCCAAATCCTCC 3’

**Target-specific PCR reaction (per rxn) – Arthropod CO1**

12.5ul Mastermix (Promega cat # M5133)

1ul Forward primer (0.4uM final)

1ul Reverse primer (0.4uM final)

3ul gDNA

7.5 ul DNase/RNase-free H2O

**Arthropod CO1 Cycling**

94C – 5 minutes

94C – 30 seconds

45C- 45 seconds x 45 cycles

72C – 45 seconds

72C – 10 minutes

10C - ∞

**Agarose gel**

Visualize 5ul PCR sample on a 2% agarose gel

**PCR Clean Up**

20ul of 1st step PCR amplicon is used for PCR clean-up using ExoI/SAP reaction.

8.85ul water, 0.023ul ExoI, 0.2275ul SAP is added to each PCR reaction and incubated 37C for 30 minutes. ExoI/SAP are then inactivated by incubation at 95C for 5 minutes.

**Indexing PCR reaction (per rxn)**

20ul Mastermix

2ul Forward Barcoded primer (0.5uM final)

2ul Universal Reverse primer (0.5uM final)

4ul template

12 ul DNase/RNase-free H2O

**Indexing PCR Cycling**

95C – 3 minutes

95C – 30 seconds

55C- 30 seconds x 8 cycles

72C – 30 seconds

72C – 2 minutes

10C - ∞

**Agarose gel**

Visualize 5ul PCR sample on a 2% agarose gel. Amplicon should be ~80bp longer due to Indexing addition

**Normalization and Pooling**

25ul of PCR amplicon is purified and normalize using the Life Technologies SequalPrep Normalization kit (cat#A10510-01) according to the manufacturer’s protocol. Samples are then pooled together

**Sequencing**

Amplicons are sequenced using an Illumina MiSeq housed in the CU Boulder BioFronteirs Sequencing Center using the v2 500-cycle kit (cat#MS-102-2002)

**Sequence Processing**

The following summarizes how the CO1 amplicons were processed via a joint QIIME [1]and UPARSE[2] pipeline similar to that of Andrei et al. [3], with modification.

Sequences were demultiplexed by taking advantage of Golay barcodes [4] via QIIME v1.9.1 [1]. The following options were used to output raw unfiltered fastq files for both forward and reverse reads: split\_libraries\_fastq.py -q 0 --max\_bad\_run\_length 250 --min\_per\_read\_length\_fraction 0.0001 --sequence\_max\_n 250 --store\_demultiplexed\_fastq… . Primer sequences were trimmed using cutadapt v1.8.1 [5] in ‘paired-end mode’ to remove the primers ZBJ-ArtF1c (5'-AGATATTGGAACWTTATATTTTATTTTTGG-3') and ZBJ-ArtR2c (5'-WACTAATCAATTWCCAAATCCTCC-3') from Zeale *et al*. [6]. Trimmed paired-ends where then merged by the –fastq\_mergepairs option of usearch v8 [7]. From here, the general quality filtering and OTU construction was completed as per the UPARSE pipeline [2], with the following modifications: OTUs were generated by clustering the reads at 99% sequence similarity, and the OTU table was generated by mapping quality filtered reads back to the OTU seeds by setting the following parameters: -maxaccepts 128 -maxrejects 1024. These parameters help to ensure that individual reads are correctly mapped to their respective OTUs.

Taxonomy was assigned by recording the top BLAST [8, 9] hit for any sequence in which the query coverage and identity exceeded 95% and 80% respectively. Any OTUs with taxonomy assignments not meeting these criteria were removed from the OTU table.

References:

1. Caporaso JG, Kuczynski J, Stombaugh J, et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. doi: 10.1038/nmeth.f.303

2. Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10:996–998. doi: 10.1038/nmeth.2604

3. Andrei A-Ş, Robeson MS II, Baricz A, et al (2015) Contrasting taxonomic stratification of microbial communities in two hypersaline meromictic lakes. ISME J 9:2642–2656. doi: 10.1038/ismej.2015.60

4. Caporaso JG, Lauber CL, Walters WA, et al (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6:1621–1624. doi: 10.1038/ismej.2012.8

5. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17:pp. 10–12.

6. Zeale MRK, Butlin RK, Barker GLA, et al (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. Molecular Ecology Resources 11:236–244. doi: 10.1111/j.1755-0998.2010.02920.x

7. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. doi: 10.1093/bioinformatics/btq461

8. Camacho C, Coulouris G, Avagyan V, et al (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421. doi: 10.1186/1471-2105-10-421

9. Altschul SF (2001) BLAST Algorithm. doi: 10.1002/9780470015902.a0005253.pub2