Manatee RNA-Seq: Wet-lab protocol

Plasma samples from Florida manatees were collected during the U.S Geological Service (USGS) Sirenia Project 2018-2019 and Florida Fish and Wildlife Conservation Commission (USGS Research permit MA-791721-5 and FWC permit MA067115-1 and IACUC, University of Florida permit #201609674) at Crystal River (Citrus County, Florida, USA) and Brevard County (Florida, USA).

Random Florida manatees were captured with a net by a land set capture or a boat-based capture as in Takeuchi et al. 2016. A veterinarian collected whole blood from minimally restrained Florida manatees from the medial interosseous space between the ulna and radio from the pectoral flippers. Blood samples were drawn using Sodium Heparin 7-mL tubes. Samples were spun on-site and the plasma was aliquoted, stored in liquid nitrogen or ice, and transferred to -80 ºC once in the lab. The buffy coat was flash-frozen in liquid nitrogen on-site.

RNA extraction.

Total RNA was extracted from the buffy coat (white blood cells) using STAT 60 (Tel-test Friendswood, TX) reagent. The protocol followed was this:

1. Approximately 350 µL of the frozen buffy coat was added to 1 ml of STAT 60 and vortexed for 30 seconds.
2. 0,2 mL chloroform was added and the tube was centrifuged 20,800 x g for 15 minutes at 4 ºC, to extract the RNA. Collect the supernatant and repeat this step with the remaining phenol-chloroform phase.
3. RNA was precipitated from the supernatants overnight at -20ºC by the addition of 700 µL isopropanol with 1.5 µL of GlycoBlue™ (15 mg/mL) (Ambion, Invitrogen, Austin, TX) as a coprecipitant.
4. Centrifugation at 20,800 x g for 45 minutes.
5. The pellet obtained was washed with ethanol 70%, air-dried, and resuspended in 20 µL of RNA secure (Ambion, Austin, TX).
6. DNAse treatment was performed to remove possible contamination with genomic DNA by using Turbo DNA-free™ kit (Ambion, Austin, TX).

For LRGASP purposes, a total number 9 different individuals were sequenced. In the case of RNA-Seq, library preparation was performed individually for each of the samples. For Long-Read Sequencing (PacBio or ONT), RNA extracted from these 9 samples were pooled together by equal mass before entering the library preparation process. Information about the health status of the animals and the RNA quality are described in Table 1.

Table 1. Florida manatee data collected during health assessments showing the laboratory identification, sex, veterinary general evaluation category based on body condition (Fair, Good or Excellent), 260/280 ratio of RNA extraction and RIN.

Table 1: List of manatee RNA samples used to create sample pool						
Animal #	**Sample ID**	**Sex**	**Life Stage**	**Evaluation**	**260/280**	**RIN**
5	ManaT-061719-5_CBC-19002	Female	Adult	Excellent	2.05	8.4
10	ManaT-061719-10_CCR-1830	Female	Adult	Good	1.94	8.2
13	ManaT-061719-13_CCR-1827	Female	Adult	Good	2.03	8.2
22	ManaT-061719-22_CCR-1818	Male	Adult	Good	2.03	8.8
25	ManaT-061719-25_CCR-1815	Female	Adult	Excellent	2.04	8.5
26	ManaT-061719-26_CCR-1814	Female	Calf	Good	2.01	8
27	ManaT-061719-27_CCR-1813	Female	Lactating	Good	2.03	8.2
32	ManaT-061719-32_CCR-1808	Female	Adult	Good	2.05	8.4
39	ManaT-061719-39_CCR-1801	Female	Adult	Good	1.98	8.4

Library preparation for RNA-sequencing with HiSeq 4000

RNA samples from the same 9 manatees were sequenced with Illumina technology. DNAse treated RNA was processed using the Ovation SoLo RNA-Seq System (NuGEN Technologies, Redwood City, CA) integrating the AnyDeplete technology to deplete unwanted transcripts of rRNA and hemoglobin RNA. The Ovation SoLo RNA-Seq System allows for strand-specific RNA-Seq library construction using low sample input 10 pg – 10 ng of total RNA. It also integrates custom-designed NuGEN's AnyDeplete technology for targeted depletion of Manatee rRNA and hemoglobin sequences.

Ovation SoLo RNA-Seq System protocol incorporates a DNase treatment to remove any gDNA from the samples for total RNA input, which will be our starting point. In this step, primer annealing is also performed. Following the instructions of Ovation SoLo RNA-Seq System manual, first strand cDNA synthesis was done followed by cDNA processing and second-strand synthesis.

The newly generated cDNA fragments go through an end repair process followed by enzymatic adaptor ligation. These library constructs are purified with Agencourt AMPure XP Beads (Beckman Coulter, Cat. #A63881). After determining empirically which was the optimal number of PCR cycles required for each library, they were amplified and purified with AMPure XP Beads (Beckman Coulter, IN, USA).

QC analysis of the library was performed using Qubit dsDNA High Sensitivity Assay and Agilent Tape Station. Libraries were treated with custom-designed Manatee rRNA targeted depletion protocol followed by a second round of amplification and beads purification. Anydeplete probes

were designed based on hemoglobin partial sequences available at NCBI for the Florida manatee (*Trichechus manatus latirostris*, txid127582):

```
>hemoglobin_subunit_alpha_Trichechus_manatus_latirostris
ACCACCAAGACCTACTTTCCTCACTTCGACCTGAGCCACGGCTCCGGCCAGATCAAGGCCCACGGC
AAGAAGGTGGCGGACGCGCTGACCCGCGCTGTCGGCCACCTGGAGGACCTGCCTGGCACTCTGTCT
GAGCTGAGCGACCTGCATGCTCACAGGTTGAGGGTGGACCCCGTCAACTTCAAGGTGAGGGGAGAG
GTTGACCTGGGCCATCAGGGGCAGCGGGCCTTCCTGGCAGGGCCGAGATTTGCGGAGAGGAGGGTT
TCCTAAGGGCCCCCTCATCGCCCCCTCTCTCCGCAGCTCCTGAGCCACTGCCTGCTGGTGACTCTG
AGCAGCCACCTCCGTGAAGATTTCACCCCTTCCGTCCACGCTTCCCTGGACAAGTTCCTCAGCAGC
GTTAGCACCGTGCTGACCTCCAAATATCGTTAAGCTGGAGCCCAGGAGAGGCGCCACCGGTCTTCC
TCCCCTCCCTGCACGTGCCTCCGAGTCTTTGAATAAAGTCTGAGTGGGCGGCAACCTGGTGTGG
>hemoglobin_subunit_alpha-like_Trichechus_manatus_latirostris
CCGCGCGCTCCCTGCGGGGACACGCTTCTGATCTCCACACAGACTCAGAAACAACCCACCATGGTG
CTGTCTGATGAGGACAAGACCAACGTGAAGACCTTCTGGGGCAAGATTGGCACCCACACCGGAGAG
TATGGCGGCGAGGCCCTGGAGAGGTGAGGATCCCCCTCTTCCCTTCTCCCGCGCACCCAGGCTCAA
CACCACCCCACCGCCCCTCGTTGGGCTACGACACGGCCGGCGCCCAGGCCCCGAGGCCCCGGGGCC
AGACCGCCCCGCCAGCCTTCCTTCGGAGGCGGGGTGCGCTCTCCCCGCGCCCCTGCCCCACCCCTC
CACCGCCACTCACCTGCCGCTTCTCCACGCAGGATGTTCCTCTCCTTCCCCACCACCAAGACCTAC
TTTCCTCACTTCGACATGAGCCACGGCTCCGGCCAGATCAAGGCCCACGGCAAGAAGGTGGCGGAC
GCGCTGACCCACGCAGTCCACCACCTGGACGACCTGCCGACAAGCTGAGGGTGGACCCTGTCAACT
TCAGGATGAGGACCCAGCCGGGCAGCCCTGCCGTCGCTATCGCTAGGGAGGCGCGTTCCCCATCCG
GGCCAGGATCTAGGGGGTGACGAGACCCAGAGGGATGCGTGGAAGGGGTGCTACGGAGCCGTGGGG
CCGAAGGCAGCTCGACCCCGTCCTCTGCAGCTCCTTGGGCACTGCCTGCTGGTAACGCCGGCCCGG
CACTACCCCGGGTACTTCAGTCCCACCATGCAGGCCTCGCTGGACAAGTTTCTGAGCCACGTGATC
TCCGCGCCTGTCTCCAAGTACCGCTGA
>n5S_Trichechus_manatus_latirostris
GTCTACGGCCATACCACCCTGAACGCGCCCGATCTCATCTGATCTCGGAAGCTAAGCAGGGTCAGG
CCTGGTTAGTACGTGGATGGGAGACTGCCTGGGAATACCAGGTGCTGTAGGAT
```

At this step, beads to sample ratio was used to size select for libraries for optimal sequencing range. The final library product was QC on the Agilent TapeStation and Roche KAPA™ Library Quantification for Illumina sequencing. Libraries were normalized and clustered on Illumina cBot using a patented PE cluster kit.

For sequencing we used the Illumina HiSeq 4000 instrument with pair-end SBS chemistry 300 cycles kit (HiSeq 3000/4000 SBS Kit, 300 cycles (Illumina, cat. no. FC-410-1003)). During sequencing, we used the Ovation SoLo Custom R1 primer. The SoLo RNA-Seq System uses 8-base barcodes for sample multiplexing followed by 8 random bases used for duplicate read determination. The sequencer was set at 151 cycles for Read 1 and 2 and 16 cycles for the index sequencing. We used Illumina bcl2fastq2 version 2.17 software to de-multiplex the data for each sample and generate fastq format files with attention paid to the unique 16 cycle index set up with Ovation SoLo RNA-Seq System.

WARNING:
After an exploratory analysis, we realized that rRNA depletion didn't work as expected and we were able to detect a significant amount of reads that belong to rRNA genes. These samples only have "identification" purposes, we do not recommend to use them for quantitative analysis.