HOKKAIDO UNIVERSITY

Title	MicroRNA expression profiling of cat and dog kidneys
Author(s)	Ichii, Osamu; Otsuka, Saori; Ohta, Hiroshi et al.
Citation	Research in Veterinary Science, 96(2), 299-303 https://doi.org/10.1016/j.rvsc.2014.01.003
Issue Date	2014-04
Doc URL	https://hdl.handle.net/2115/56453
Type	journal article
File Information	Supplemental methods.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Supplemental methods

Before aligning sequences to a reference genome, the sequencing adaptors (Illumina; San Diego, CA, USA), which were used in the sequencing library preparation, and low quality sequences evaluated using the Quality Value (QV) Score were eliminated using Trimmomatic-0.27 (Lohse et al., 2012). In this case, the average QV Score was calculated using a sliding window method (window size; 4 bp) from both ends of a sequence. When the average score was ≤15 (error rate approximately 3.2%), those four successive nucleotides were discarded from the read. In addition, for filtering the reads from RNAs not considered small and error-prone short reads, those longer than 30 bp or shorter than 18 bp were also discarded. We used COBWeB, the alignment module in Avadis NGS 1.4.7 (Agilent) to generate sequence alignments. Because comprehensive annotation information is not available for small RNAs of dogs and cats, we utilized a mouse reference genome, mm9, and the comprehensive annotation included. One mismatch and no gap between a read and reference sequence were allowed. In the alignment process, the reads matched to mouse rRNA sequences were screened by extracting mouse rRNA sequences from the ncRNA FASTA dataset available at Ensembl Web site (ftp://ftp.ensembl.org/pub/release-72/fasta/mus_musculus/) and constructed an rRNA

screening database.

The raw read counts allocated to the small RNA loci in mm9 were obtained using this alignment procedure, and those raw counts were normalized to the “reads per kilobase per million” (RPKM) value (Mortazavi et al., 2008) using Avadis NGS 1.4.7. RPKM was defined as follows: RPKM = R/(L*N) (R = number of reads mapping to the region of interest; N, total number of reads from the sample; L, length of the region of interest).

For counting the miRNAs, we considered reads that were precisely aligned to the 5’ end of active/mature region of an miRNA. Two-fold increases in miRNAs compared with CO vs. MD for the dog and the cat were obtained and subjected to target gene prediction using PITA (Kertesz et al., 2007), PicTar (Krek et al., 2005), TargetScan (Lewis et al., 2005), and microRNAorg (Betel et al., 2008). A significant GO term was identified using the Generic Gene Ontology Term Finder (http://go.princeton.edu/cgi-bin/GOTermFinder).

1. Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., Segal, E., 2007. The role of site accessibility in microRNA target recognition. Nat. Genet. 39, 1278–84.

2. Krek, A., Grün, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., Rajewsky, N., 2005. Combinatorial microRNA target predictions. Nat. Genet. 37, 495–500.

3. Lewis, B.P., Burge, C.B., Bartel, D.P., 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120, 15–20.

4. Betel, D., Wilson, M., Gabow, A., Marks, D.S., Sander, C., 2008. The microRNA.org resource: targets and expression. Nucleic Acids Res. 36, D149–53.