Protocol for Tissue RNA-Seq: Gingeras Lab

1. Description
The goals of these experiments is to generate RNA-Seq data for several mouse tissues (mainly from C57BL/6) and cell lines from many developmental timepoints. These data will be used to generate a breadth and depth of expression profiles as well as to improve the mouse gene and transcript annotation set.

2. Harvesting Tissue
All mice were euthanized by CO2 and embryos and pups less then 10 days old were decapitated. Tissues were be taken from embryos, pups and adult animals of all ages spanning the developmental lifecycle. The animals were washed with ETOH and body cavities exposed to allow removal of organs. The following organs were removed using forceps and scissors: Gonadal Adipose Tissue, Subcutaneous Adipose Tissue, Adrenal glands, bladder, CNS, spinal cord and brain, duodenum, large intestine, small intestine, stomach, limbs, mammary gland and ovaries and placed in RNA later.

3. Library Preparation
RNAs > 200 nucleotides were isolated from the tissue samples using commercial kits. Polyadenylated RNA was isolated using oligo-dT columns and DNAse treated. The A+ RNA was also treated using commercial kits to remove ribosomal RNAs. Illumina-compatible directional RNA-Seq libraries were prepared using the protocol outlined by *Parkhomchuk et al* [1]. To each RNA we add exogenous spike-in prior to library construction [2]. This is the same protocol that was used to generate the bolus of the ENCODE human RNA-Seq data.

4. Sequencing and Analysis
Samples were sequenced on both the Illumina GAIIx (as pair-end 76 reads) and the Illumina Hi-Seq (as pair-end 101 reads). Each sample we sequenced to an average depth of 100 million mate-pairs. The data were mapping with STAR (Gingeras lab: http://gingeraslab.cshl.edu/STAR/). We make contigs, Cufflinks gene and transcript models and splice junction files as elements that contain IDR (irreproducibility scores) as well as compare them against the current annotations to derive expression values for genes, transcripts and exons. The data are sent to the DCC.

[1] Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A. Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res*. 2009 Oct; 37(18):e123.

[2] Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, Gingeras TR, Oliver B. Synthetic spike-in standards for RNA-seq experiments. *Genome Res.* 2011 Aug 4.