**RNAseq Methods Text**

*Use appropriate sentences/paragraphs and text in red appropriate to your project.*

*<comments in angle brackets add clarification but should be removed in final methods text>*

*Adjust tense if needed (proposal vs publication).*

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

***<Use only if TREx isolates your RNA, or if you use similar methods>***

**RNA isolation**

Total RNA will be purified using Trizol (Thermo Fisher) according to the commercial protocol with the following additions: after the first phase separation, additional chloroform extraction step of the aqueous layer in Phase-lock Gel heavy tubes (Quanta Biosciences); addition of 1ul Glyco-blue (Thermo Fisher) immediately prior to isopropanol precipitation; two washes of the RNA pellet with 75% ethanol. If the RNA integrity results indicate co-purified genomic DNA, it will be removed with the RapidOUT DNA Removal kit (Thermo Fisher).

**RNA Quality Control**

RNA sample quality will be confirmed by spectrophotometry (Nanodrop) to determine concentration and chemical purity (A260/230 and A260/280 ratios) and with a Fragment Analyzer (Advanced Analytical) to determine RNA integrity.

*<OR, for low yield samples>*

RNA sample quality will be confirmed using a Qubit3 (RNA HS kit; Thermo Fisher) to determine concentration and with a Fragment Analyzer (Advanced Analytical) to determine RNA integrity.

*<OR, for ultra low yield samples>*

RNA sample quality will be confirmed using a Qubit3 (RNA HS kit; Thermo Fisher) to determine concentration and with a Fragment Analyzer (Advanced Analytical) to determine RNA integrity.

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

***<Enrichment:*** *most projects use one of the following methods>*

**Ribosomal RNA Subtraction**

Ribosomal RNA will be subtracted by hybridization from total RNA samples using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat v1/v2 *<OR>* Bacteria, New England Biolabs).

*<OR>*

**polyA+ RNA Isolation**

PolyA+ RNA will be isolated with the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs).

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**Illumina Library Preparation**

TruSeq-barcoded RNAseq libraries will be generated with the NEBNext Ultra II [Directional] RNA Library Prep Kit (New England Biolabs). Each library will be quantified with a Qubit 2.0 (dsDNA HS kit; Thermo Fisher) and the size distribution will be determined with a Fragment Analyzer (Advanced Analytical) prior to pooling.

*<Low input libraries (<20nt total RNA) will be generated with the Ultra II RNA Library Prep Kit (non-directional).>*

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**Illumina Sequencing**

Libraries will be sequenced on an Illumina instrument. At least 20M <OR> 10M reads will be generated per library.

*<Instrument and read length: we are transitioning to mostly 2x150nt PE reads on a HiSeq or NovaSeq, but some projects may be sequenced on a NextSeq500 [SE 75nt kit] or NextSeq-2k [SE 100nt kit].*

*Read depth: default 20M for RNAseq. Some projects with bacterial samples may target 10M raw reads, and some projects have custom deeper sequencing depth.>.*

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**Analysis**

*preprocessing***:** reads will be trimmed for low quality and adaptor sequences with TrimGalore v0.6.0 (ref 1), a wrapper for cutadapt (ref 2) and fastQC (ref 3).

Parameters: -j 1 -e 0.1 --nextseq-trim=20 -O 1 -a AGATCGGAAGAGC --length 50 --fastqc

*<optional: custom projects only, such as removal of host sequences for pathogen identification projects>*

*subtract unwanted reads*: unwanted reads will be removed with STAR v 2.7.0e (ref 4).

Parameters: --outReadsUnmapped Fastx

*mapping*: reads will be mapped to the reference genome/transcriptome (name/version) using STAR v2.7.0e (ref 4).

Parameters: --outSAMstrandField intronMotif , --outFilterIntronMotifs RemoveNoncanonical ,

--outSAMtype BAM SortedByCoordinate, --quantMode GeneCounts

*gene expression analysis*: SARTools and DESeq2 v1.26.0 will be used to generate normalized counts and statistical analysis of differential gene expression (ref 5,6).

Parameters: fitType parametric, cooksCutoff TRUE, independentFiltering TRUE, alpha 0.05,

pAdjustMethod BH, typeTrans VST, locfunc median

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

***Appendix: Current reference genomes***

*TREx uses Ensembl transcriptome annotations by default*

*Human: Ensembl GRCh38*

*Mouse: Ensembl GRCm38*

*Dog: Ensembl CanFam3*

*Cat: Ensembl Felis\_catus9.0*

*Chicken: Ensembl Galgal5*

*Other: available on request*

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

***Software References:***

1) TrimGalore: Felix Krueger

<http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/>

2) cutadapt: Marcel Martin

<https://cutadapt.readthedocs.io/en/stable/>

<http://journal.embnet.org/index.php/embnetjournal/article/view/200>

3) fastQC: Simon Andrews

<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

4) STAR: Alexander Dobin

<https://doi.org/10.1093/bioinformatics/bts635>

5) SARTools: Hugo Varet

<http://dx.doi.org/10.1371/journal.pone.0157022>

6) DEseq2: Michael Love

<http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html>

<https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8>