
Supplementary information

Tissue-specific and interferon-inducible expression of nonfunctional ACE2 through endogenous retroelement co-option

In the format provided by the
authors and unedited

Supplementary Code 1. Custom code used for TPM calculations.

This code was used to estimate expression of all Trinity assembled transcripts using quasi-mapping of Salmon in samples from TCGA or GTEx. Executable lines are marked with '>'. The bash script assumes the Salmon transcriptome index was generated prior to executing and is provided as first Argument. The bam or fastq file name is provided as second Argument.

```
> salmon index -t flattened_transcriptome.fasta -i ./transcriptome_index
```

Salmon version 0.8.2

SAMtools version 1.3.1

Python version 3.5.2

TCGA paired-end RNA-seq was downloaded through GDC portal as bam files. GTEx paired-end RNA-seq was downloaded with the SRA toolkit as fastq files. Hence, TCGA files were first converted to fastq files. Here, the following programs were used from github.com/A-N-Other/pedestal: deinterleave (commit 169dd81) and interleavei (commit 800e78f). They are available under the permissive MIT license and we encourage code re-use and comment.

#####

Step 1

file input & directory and test if file exists

```
> indexfile=$1
> bamname=$2 #collects string of input file from 1st argument (*.bam)

> fullfilename="${bamname##*/}"
> indirectory="${bamname%/*}"
> subdirectory="${indirectory##*/}"
> filename="${fullfilename%_gdc_realn_rehead.bam}"
  #> filename="${fullfilename%_1.fastq.gz}" # for GTEx fastq files

> if [ -s ./countFiles/"$subdirectory"/"$filename".counts.gz ]; then
    exit 3; fi
```

#####

#####

Step 2

separate first and second read of every pair with deinterleave script

```
> samtools fastq -n "$bamname" | interleave -u -
  | deinterleave ~/Scratch/tcgarecycle/fastq.files/"$filename"_1.fq
  ~/Scratch/tcgarecycle/fastq.files/"$filename"_2.fq
```

#####

#####

Step 3

Expression estimate with Salmon quasi-mapping

```
> ~/bin/salmon-0.8.2/bin/salmon quant \
    -p 8 -i "$indexfile" \
    -l IU --seqBias --gcBias \
    -1 ~/Scratch/tcgarecycle/fastq.files/"$filename"_1.fq \
    -2 ~/Scratch/tcgarecycle/fastq.files/"$filename"_2.fq \
    -o ./countFiles/"$subdirectory"/"$filename".tmpdirectory

> mv ./countFiles/"$subdirectory"/"$filename".tmpdirectory/quant.sf \
    ./countFiles/"$subdirectory"/"$filename".counts
> mv ./countFiles/"$subdirectory"/"$filename".tmpdirectory/aux_info/ambig_info.tsv
\  ./countFiles/"$subdirectory"/"$filename".ambig_info.tsv
```

Supplementary Table 1. Primers used in this study

5' RACE PCR primers

Primer name	Sequence (5' – 3')	Target
GSP1	TTTGTGCACATAAGGATCCTCC	Human <i>ACE2</i>
GSP2	TTTGTCCCAGCCTGCTTC	Human <i>ACE2</i>

RT-qPCR primers

Primer name	Sequence (5' – 3')	Target
<i>ACE2</i> Forward	GTAACGGACCCAGGAAAT	Human <i>ACE2</i>
<i>MIRb-ACE2</i> Forward	TACAAGTGCTTCATTGAGG	Human <i>MIRb-ACE2</i>
Human <i>ACE2</i> Common Reverse	CCATTGTCACCTTTGTGC	Human <i>ACE2</i> and <i>MIRb-ACE2</i>
Vervet <i>Ace2</i> Forward	CTAACTGATCCAGGAAAT	Vervet <i>Ace2</i>
Vervet <i>MIRb-Ace2</i> Forward	CACAAATGCTTCATTGAGG	Vervet <i>MIRb-Ace2</i>
Vervet <i>Ace2</i> Common Reverse	CCATTGTCACCTTTGTGC	Vervet <i>Ace2</i> and <i>MIRb-Ace2</i>
Rabbit <i>Ace2</i> Forward	CTAACTGAGCCAGGAGAT	Rabbit <i>Ace2</i>
Rabbit <i>MIRb-Ace2</i> Forward	TACAAGTGTTTCATTTAGA	Rabbit <i>MIRb-Ace2</i>
Rabbit <i>Ace2</i> Common Reverse	CCATTGTTACTTTTGTGC	Rabbit <i>Ace2</i> and <i>MIRb-Ace2</i>
Dog <i>Ace2</i> Forward	CTAACTGAGCCAAGTGAC	Dog <i>Ace2</i>
Dog <i>MIRb-Ace2</i> Forward	TACAAATGTTTCATTGAGG	Dog <i>MIRb-Ace2</i>
Dog <i>Ace2</i> Common Reverse	CCATCGTTACCTTTGTGC	Dog <i>Ace2</i> and <i>MIRb-Ace2</i>
Mouse <i>Ace2</i> Forward	ACTTCTCTAACATGTAAG	Mouse <i>Ace2</i>
Mouse <i>MIRb-Ace2</i> Forward	TACAAGTTGTGTATGTTCT	Mouse <i>MIRb-Ace2</i>
Mouse <i>Ace2</i> Common Reverse	GTCTCTTGTCACCTTACAC	Mouse <i>Ace2</i> and <i>MIRb-Ace2</i>
<i>CXCL10</i> Forward	TGCCATTCTGATTTGCTGCC	Human <i>CXCL10</i>
<i>CXCL10</i> Reverse	TGCAGGTACAGCGTACAGTT	Human <i>CXCL10</i>
<i>CD274</i> Forward	TACAGCTGAATTGGTCATCCCA	Human <i>CD274</i>
<i>CD274</i> Reverse	TCAGTGCTACACCAAGGCAT	Human <i>CD274</i>
<i>HPRT</i> Forward	TGACACTGGCAAAACAATGCA	<i>HPRT</i> (all species)
<i>HPRT</i> Reverse	GGTCCTTTTCACCAGCAAGCT	<i>HPRT</i> (all species)