

5.1 RNA Hydrolysis Protocol (M.musculus)

Input

Download

[Reference Annotation](#)

[Parameter File](#)

[Reference Genome](#)

Parameter

Expression		
NB_MOLECULES	5,000,000	Number of RNA molecules initially in the experiment
TSS_MEAN	25	Average deviation from the annotated transcription start site (TSS)
POLYA_SCALE	300	Scale of the Weibull distribution, shifts the average length of poly-A tail sizes
POLYA_SHAPE	2	Shape of the Weibull distribution describing poly-A tail sizes
Fragmentation		
FRAG_SUBSTRATE	RNA	Specifies RNA as the substrate of fragmentation
FRAG_METHOD	UR	Uniform random fragmentation
FRAG_UR_ETA	170	Average expected framgent size after fragmentations, i.e., number of breaks per unit length (exhaustiveness of fragmentation)
FRAG_UR_D0	1	Minimum length of fragments produced by UR fragmentation
FRAG_UR_DELTA	NaN	Geometry of molecules in the UR process depends logarithmically on molecule length
Reverse Transcription		
RTRANSCRIPTION	YES	Switch on the reverse transcription
RT_PRIMER	RH	Use random hexamer primers used for first strand synthesis
RT_MOTIF	default	A default PWM of the current Illumina protocol is used
RT_LOSSLESS	YES	Flag to force every molecule to be reversely transcribed
RT_MIN	500	Minimum length observed after reverse transcription of full-length transcripts
RT_MAX	5,500	Maximum length observed after reverse transcription of full-length transcripts
Amplification and Size Segregation		
PCR_DISTRIBUTION	default	Default PCR distribution with 15 rounds and 20 bins
GC_MEAN	0.5	Mean value of a gaussian distribution that reflects GC bias amplification probability
GC_SD	0.1	Standard deviation of a gaussian distribution that reflects GC bias amplification probability
FILTERING	YES	Enables size filtering of fragments
SIZE_DISTRIBUTION	null	Employ an empirical Illumina fragment size distribution
SIZE_SAMPLING	MH	The Metropolis-Hastings algorithm is used for filtering
Sequencing		
READ_NUMBER	15,000,000	Produce 15 million reads
READ_LENGTH	75	Each read sequence is 75nt long
PAIRED_END	NO	Single reads are simulated (one per fragment)

Output

--

```

[INFO] I am collecting information on the run.
      initializing profiler *****
[INFO] Checking GTF file
***** OK (00:00:03)
[PROFILING] I am assigning the expression profile
***** OK (00:00:05)
      Reading reference annotation ***** OK (00:00:06)
      found 28045 transcripts
[PROFILING] Parameters
      NB_MOLECULES      5000000
      EXPRESSION_K      -0.6
      EXPRESSION_X0     5.0E7
      EXPRESSION_X1     9500.0
      PRO_FILE_NAME     /Users/micha/Desktop/mm9_hydrolysis.pro
      profiling ***** OK (00:00:00)
      Updating .pro file ***** OK (00:00:00)
      molecules      4999480
[LIBRARY] creating the cDNA library
      Initializing Fragmentation File ***** OK (00:00:06)
      4999480 mol initialized
[LIBRARY] Fragmentation UR
[LIBRARY] Configuration
      DO: 1.0
      Delta: Not specified, depends on sequence length
      Eta: 170.0
      Processing Fragments ***** OK (00:03:20)
      99433550 mol: in 4999480, new 94434070, out 99433550
      avg Len 154.13617, maxLen 499
      preparing transcript sequences ***** OK (00:02:04)
[INFO] Initializing PWM cache
[INFO] Done
[LIBRARY] Reverse Transcription
[LIBRARY] Configuration
      Mode: RH
      PWM: motif_1mer_0-5.pwm
      RT MIN: 500
      RT MAX: 5500
      Processing Fragments ***** OK (00:18:53)
      99436361 mol: in 99433550, new 2811, out 99436361
      avg Len 226.49129, maxLen 718
      initializing Selected Size distribution
[LIBRARY] Segregating cDNA (Acceptance)
      Processing Fragments ***** OK (00:02:32)
      99436361 mol: in 99436361, new 0, out 3935454
      avg Len 183.18074, maxLen 299
      start amplification
[INFO] Loading default PCR distribution
[INFO] Initializing PWM cache
[INFO] Done
[LIBRARY] Amplification
[LIBRARY] Configuration
      Rounds: 15
      Mean: 0.5
      Standard Deviation: 0.1
      Processing Fragments ***** OK (00:00:19)
      Amplification done.
      In: 3935454 Out: 106111525
      3935454 mol: in 3935454, new 0, out 106111525
      avg Len 183.16734, maxLen 299
      Copied results to /Users/micha/Desktop/mm9_hydrolysis.lib
      Updating .pro file ***** OK (00:00:00)
[SEQUENCING] getting the reads
      Initializing Fragment Index
      Indexing ***** OK (00:00:03)
      2112053 lines indexed (106111525 fragments, 16421 entries)
      sequencing ***** OK (00:04:11)
      106111525 fragments found (2112053 without PCR duplicates)
      15000653 reads sequenced
      2333333 reads fall in poly-A tail
      511978 truncated reads
      Moving temporary BED file

```

```
Updating .pro file ***** OK (00:00:00)
Updating .pro file ***** OK (00:00:00)
Updating .pro file ***** OK (00:00:00)
Updating .pro file ***** OK (00:00:00)
[END] I finished, took me 2291 sec.
```