USER GUIDE

ambion®
by life technologies™

ERCC RNA Spike-In Control Mixes

ERCC RNA Spike-In Mix
ERCC ExFold RNA Spike-In Mixes

Catalog Number 4456740, 4456739

Publication Number 4455352
Revision D

life technologies™

For Research Use Only. Not intended for any animal or human therapeutic or diagnostic use.

Information in this document is subject to change without notice.

DISCLAIMERS

LIFE TECHNOLOGIES CORPORATION AND/OR ITS AFFILIATE(S) DISCLAIM ALL WARRANTIES WITH RESPECT TO THIS DOCUMENT, EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT. TO THE EXTENT ALLOWED BY LAW, IN NO EVENT SHALL LIFE TECHNOLOGIES AND/OR ITS AFFILIATE(S) BE LIABLE, WHETHER IN CONTRACT, TORT, WARRANTY, OR UNDER ANY STATUTE OR ON ANY OTHER BASIS FOR SPECIAL, INCIDENTAL, INDIRECT, PUNITIVE, MULTIPLE OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH OR ARISING FROM THIS DOCUMENT, INCLUDING BUT NOT LIMITED TO THE USE THEREOF.

The use of ERCC in the product name does not constitute an affiliation or sponsorship by the External RNA Controls Consortium.

NOTICE TO PURCHASER

Limited Use Label License No. 358: Research Use Only

The purchase of this product conveys to the purchaser the limited, non-transferable right to use the purchased amount of the product only to perform internal research for the sole benefit of the purchaser. No right to resell this product or any of its components is conveyed expressly, by implication, or by estoppel. This product is for internal research purposes only and is not for use in commercial applications of any kind, including, without limitation, quality control and commercial services such as reporting the results of purchaser's activities for a fee or other form of consideration. For information on obtaining additional rights, please contact **outlicensing@lifetech.com** or Out Licensing, Life Technologies, 5791 Van Allen Way, Carlsbad, California 92008.

TRADEMARKS

The trademarks mentioned herein are the property of Life Technologies Corporation or their respective owners. TaqMan is a registered trademark of Roche Molecular Systems, Inc. Illumina is a registered trademark of Illumina, Inc.

© 2012 Life Technologies Corporation. All rights reserved.

Contents

ERCC RNA Spike-In Control Mixes

Purpose

The Ambion® ERCC RNA Spike-In Control Mixes provide a set of external RNA controls that enable performance assessment of a variety of technology platforms used for gene expression experiments. Add one Spike-In Mix to each RNA sample, and run the Spike-In Mix-containing samples on your platform. Then compare the Spike-In Mix data to known Spike-In Mix concentrations and ratios to assess the dynamic range, lower limit of detection, and fold-change response of your platform.

A number of gene expression technologies can benefit from the use of the ERCC RNA Spike-In Control Mixes, including next-generation sequencing (NGS), microarrays and PCR-based assays.

Note: Microarrays must contain probes designed to interrogate the ERCC transcripts to benefit from the use of the ERCC RNA Spike-In Control Mixes. Consult with the array manufacturer for details.

Background on the ERCC

The External RNA Controls Consortium (ERCC) is an ad-hoc group of academic, private, and public organizations. The National Institute of Standards and Technology (NIST)-hosted ERCC has been working to develop a common set of RNA controls that can be used in multiple gene expression platforms such as quantitative RT-PCR, microarrays, and NGS technologies.

The outcome of the ERCC effort is a reference library of NIST-certified DNA plasmids that are designed to produce a set of transcripts 250–2000 nt in length that mimic natural eukaryotic mRNAs. See "Additional background on the ERCC plasmid library" on page 17 for further information.

ERCC RNA Spike-In Control Mixes

The ERCC RNA Spike-In Control Mixes are pre-formulated sets of 92 polyadenylated transcripts from the ERCC plasmid reference library. The transcripts are traceable through the manufacturing process to the NIST plasmid reference material. (The use of ERCC in the product name does not constitute an affiliation or sponsorship by the External RNA Controls Consortium.)

The transcripts are manufactured in two formulations, Spike-In Mix 1 and Spike-In Mix 2, each containing the full complement of 92 transcripts. The concentrations of the transcripts in each Spike-In Mix span an approximately 10^6-fold concentration range that is suitable for use with NGS platforms, such as the SOLiD® System, and microarray platforms, such as the Illumina® BeadArray System.

The transcripts in Spike-In Mix 1 and Spike-In Mix 2 are present at defined Mix 1:Mix 2 molar concentration ratios, described by 4 subgroups (Table 1 on page 6). Each subgroup contains 23 transcripts spanning a 10^6-fold concentration range, with approximately the same transcript size distribution and GC content.

Table 1 Transcript molar ratios in ERCC Spike-In Mixes

Subgroup	Mix 1:Mix 2†
A	4.00
B	1.00
C	0.67
D	0.50

† Applies only to Spike-In Mix 1 and Mix 2 with same manufacturing lot number.

Each Spike-In Mix is ready to be diluted and added to the RNA sample before processing for gene expression measurements. (The ERCC RNA Spike-In Control Mixes are not recommended for small RNA expression analysis, because of the size range of the ERCC transcripts.)

ERCC_Analysis Plugin

The ERCC_Analysis plugin is intended to help with ERCC RNA Spike-in Controls. It enables you to quickly determine whether or not the ERCC results indicate a problem with library preparation or the PGM run.

For more information about the ERCC_Analysis Plugin, refer to the *ERCC_Analysis Plugin User Bulletin* (Pub no. 4479068).

Choosing a kit

Choose the kit configuration that meets your experimental needs (Table 2).

- The **ERCC RNA Spike-In Mix** (Part no 4456740) includes Spike-In Mix 1 alone; it can be used to assess platform dynamic range and lower limit of detection.
- The **ERCC ExFold RNA Spike-In Mixes** (Part no 4456739) include both Spike-In Mix 1 and Spike-In Mix 2. In addition to the performance measurements described above for the ERCC RNA Spike-In Mix, this kit can be used to assess the accuracy of measurements of differential gene expression on your platform.

Table 2 Performance measurements and experimental strategies with ERCC RNA Spike-In Control Mixes

Performance measurement	ERCC RNA Spike-In Mix	ERCC ExFold RNA Spike-In Mixes	Experimental strategy
Dynamic range and lower limit of detection (dose response)	+	+	Add *either* Spike-In Mix 1 or Spike-In Mix 2 to one or more RNA samples (page 13).
Fold-change response: assess the platform response to known transcript ratios	–	+	Add Spike-In Mix 1 and Spike-In Mix 2 to *separate* samples (page 14).† For example: • Add Mix 1 to the control RNA sample, and add Mix 2 to the treated samples. - *or* - • Randomize Mix 1 and Mix 2 among samples.

† Use only the same lot number of Spike-In Mix 1 and Mix 2 when assessing fold-change ratios.

Kit contents and storage

Each tube of Spike-In Mix is sufficient for approximately 100 next-generation sequencing or 200 microarray sample preparation reactions, depending on the amount of RNA used.

Table 3 ERCC RNA Spike-In Control Mixes: kit configurations

Component	ERCC RNA Spike-In Mix (Part no. 4456740)	ERCC ExFold RNA Spike-In Mixes (Part no. 4456739)	Storage Temperature‡
ERCC RNA Spike-In Mix 1†	10 µL	–	–20°C§
ExFold Spike-In Mix 1†	–	10 µL	–20°C§
ExFold Spike-In Mix 2	–	10 µL	–20°C§
Nuclease-free Water	1.75 ml	1.75 ml	–20°C, 4°C, or room temperature

† Although ERCC RNA Spike-In Mix 1 and ExFold Spike-In Mix 1 contain the same formulation of ERCC transcripts, do ***not*** substitute ERCC RNA Spike-In Mix 1 for ExFold Spike-In Mix 1 for fold-change assessment. Use only ExFold Spike-In Mix 1 and Mix 2 with the same manufacturing lot number.

‡ Do not freeze and thaw the Spike-In Mixes more than 8 cycles.

§ Do not store in a frost-free freezer.

Materials and equipment required

For optional materials and equipment, see "Related products" on page 18.

Item	Source
Non-stick RNase-free Microfuge Tubes	AM12350
Microcentrifuge	MLS†
Pipettor, positive-displacement or air-displacement recommended	MLS
Pipette tips, RNase-free	MLS
Vortex mixer	MLS

† Major laboratory supplier.

Workflow

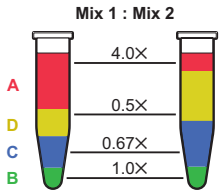
Prepare sample RNA with ERCC RNA Spike-In Control Mixes

Start with purified total RNA, poly(A), or rRNA-depleted RNA

↓

Add Spike-In Mix 1 ***or*** Mix 2 to the RNA sample(s)

↓



Perform gene expression measurements

Process the ERCC Control Mix-containing RNA sample(s)

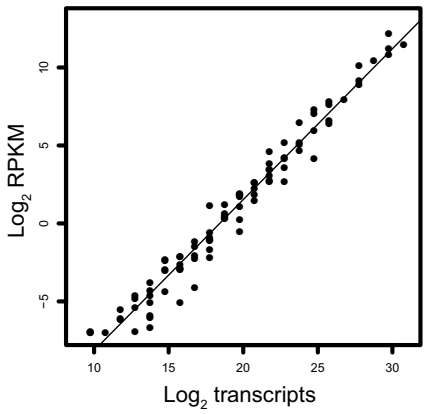
↓

Perform the gene expression measurements on your platform

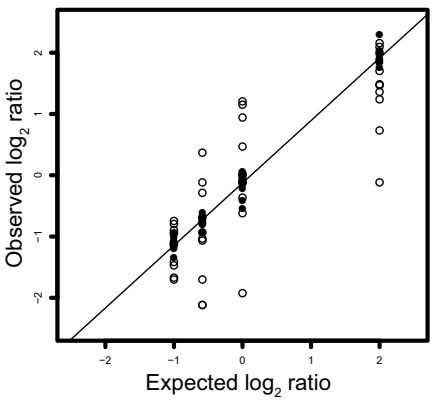
↓

Analyze the ERCC RNA Spike-In Control Mixes data

- Assess the lower limit of detection and dynamic range (dose response)



- Assess the fold-change response



Prepare sample RNA with ERCC RNA Spike-In Control Mixes

Guidelines

- Use the ERCC RNA Spike-In Control Mixes with:
  - Purified total RNA
  - Poly(A) RNA
  - rRNA-depleted RNA

    (Recommended) Add the Spike-In Mix to total RNA samples *before* poly(A) selection or rRNA depletion, to control for variation through the entire sample processing workflow. See "Determine the amount of Spike-In Mix to add", below.

    If required by your experimental needs, you can add the Spike-In Mix to RNA samples after the selection or depletion procedure instead of before. (This strategy controls only for sample processing variation that occurs downstream of the selection or depletion procedure, however.)
- Add Spike-In Mix to RNA samples before processing for gene expression measurements. For example:
  - For the SOLiD® Total RNA-Seq Kit, add Spike-In Mix to each whole transcriptome RNA sample before RNA fragmentation.

    Note: The ERCC RNA Spike-In Control Mixes are not recommended for use with the small RNA analysis procedure of the SOLiD® Total RNA-Seq Kit.
  - For the Illumina® TotalPrep™ RNA Amplification Kit, add Spike-In Mix to total RNA before the reverse transcription step.
- Add only one Spike-In Mix (Mix 1 or Mix 2) to each RNA sample.
- If RNA sample concentration is required by your downstream RNA processing procedure, add Spike-In Mix before the concentration step. Unless otherwise specified by your RNA processing procedure, a centrifugal vacuum concentrator (for example, SpeedVac) is recommended.

Determine the amount of Spike-In Mix to add

Determine the amount of Spike-In Mix 1 or Mix 2 to add, using the guidelines in Table 4 on page 10, and the sample RNA input requirements for your specific downstream application.

- ***For total RNA that will undergo poly(A) selection or rRNA depletion***, follow the guidelines in Table 4 for the starting amount of total RNA that will be used in the selection or depletion procedure.

  For example, if the starting amount of total RNA for the rRNA-depletion procedure is 5 µg, add 1 µL of a 1:10 dilution of Spike-In Mix 1 or Mix 2 to 5 µg of total RNA, then use the Spike-In Mix-containing sample in your rRNA-depletion procedure.
- ***For RNA that has already undergone poly(A) selection***, follow the guidelines in Table 4.
- ***For rRNA-depleted RNA,*** follow the guidelines in Table 4, based on the starting amount of total RNA that was used in the depletion procedure.

  For example, if the rRNA-depletion procedure used 5 µg of total RNA, add 1 µL of 1:10 dilution of Spike-In Mix 1 or Mix 2 to the *total recovered amount* of rRNA-depleted RNA.

Table 4 Guidelines for adding Spike-In Mixes to sample RNA

Amount of sample RNA	Volume of Spike-In Mix 1 or Mix 2 (dilution)†	
	Total RNA	Poly(A) RNA
20 ng	4 µL (1:10000)	2 µL (1:100)
50 ng	1 µL (1:1000)	5 µL (1:100)
100 ng	2 µL (1:1000)	1 µL (1:10)
500 ng	1 µL (1:100)	5 µL (1:10)
1000 ng	2 µL (1:100)	–
5000 ng	1 µL (1:10)	–

† ERCC RNA Spike-In Mix 1, ExFold Spike-In Mix 1, or ExFold Spike-In Mix 2.

Add Spike-In Mix 1 or Mix 2

IMPORTANT! For fold-change response assessment, use Spike-In Mix 1 and Mix 2 from the same manufacturing lot.

1. Add Spike-In Mix 1 or Mix 2 to each RNA sample.
   a. Prepare the appropriate dilution of each Spike-In Mix needed. Prepare a fresh dilution of the Spike-In Mix for each procedure. Discard unused diluted Spike-In Mix.

      Scale the volumes accordingly if more than 10 µL of the dilution are needed.

Dilution	Spike-In Mix†	Nuclease-free Water
1:10	1 µL undiluted	9 µL
1:100	1 µL of 1:10	9 µL
1:1000	1 µL of 1:100	9 µL
1:10,000	1 µL of 1:1000	9 µL

      † ERCC RNA Spike-In Mix 1, ExFold Spike-In Mix 1, or ExFold Spike-In Mix 2.

   b. Add the volume of the diluted Spike-In Mix (determined from Table 4) to each RNA sample. (Add only one Spike-In Mix to each sample.)
2. (Optional) If required by your downstream RNA processing workflow, concentrate the sample RNA containing Spike-In Mix (centrifugal vacuum concentration recommended).

Perform gene expression measurements

After adding a Spike-In Mix to RNA samples, process the Spike-In Mix-containing samples for the platform of choice and perform the gene expression measurements.

For example, for whole transcriptome analysis using SOLiD® System next-generation sequencing, prepare SOLiD® cDNA libraries using Spike-In Mix-containing RNA samples and the SOLiD® Total RNA-Seq Kit, then continue processing the libraries in the SOLiD® System sequencing workflow.

Prepare for data analysis

Overview

Although there are a number of potential uses for the ERCC RNA Spike-In Control Mixes, in general, analysis with the ERCC Spike-In Mixes can be separated into two broad categories:

- Dose response (platform dynamic range and lower limit of detection), discussed on page 13.
- Fold-change analytics, discussed on page 14.

Prepare for data analysis of Spike-In Mix-containing samples

Data analysis with the ERCC Spike-In Mixes requires ERCC_Controls_Analysis.txt, a tab-delimited file containing the transcript ERCC_IDs, the molar concentration of each transcript in Spike-In Mix 1 and Mix 2, and the expected fold-change ratio for each transcript when comparing Mix 1 to Mix 2. This file provides reference values for comparison with the expression profiling data generated from Spike-In Mix-containing samples.

Additionally, ERCC_Controls_Annotation.txt contains the sequence of each individual control along with associated TaqMan® Gene Expression Assay IDs for each control sequence.

Download ERCC_Controls_Analysis.txt and ERCC_Controls_Annotation.txt through the ERCC RNA Spike-In Control Mixes product web page. Figure 1 shows a screenshot of ERCC_Controls_Analysis.txt opened in a spreadsheet program.

Note: Before data analysis, it is often convenient to transform the Mix 1 and Mix 2 concentration values to reflect the dilution scheme used from Table 4 on page 10. For example, if 2 µL of a 1:100 dilution of Spike-In Mix was added to 1 µg of total RNA, multiply columns D and E in Figure 1 by 0.02 to give new concentration values, expressed as attomoles of ERCC transcript/1 µg total RNA. The concentration values can be expressed in terms of absolute number by conversion of moles to molecules with Avogadro's number (N_A; $6.02214179 \times 10^{23}$ mol^{-1}).

Figure 1 Spreadsheet view of ERCC_Controls_Analysis.txt

B	C	D	E	F	G
ERCC ID	subgroup	concentration in Mix 1 (attomoles/ul)	concentration in Mix 2 (attomoles/ul)	expected fold-change	log2(Mix 1/Mix 2
ERCC-00130	A	30000	7500	4	2
ERCC-00004	A	7500	1875	4	2
ERCC-00136	A	1875	468.75	4	2
ERCC-00108	A	937.5	234.375	4	2
ERCC-00116	A	468.75	117.1875	4	2
ERCC-00092	A	234.375	58.59375	4	2
ERCC-00095	A	117.1875	29.296875	4	2
ERCC-00131	A	117.1875	29.296875	4	2
ERCC-00062	A	58.59375	14.6484375	4	2
ERCC-00019	A	29.296875	7.32421875	4	2
ERCC-00144	A	29.296875	7.32421875	4	2
ERCC-00170	A	14.6484375	3.66210938	4	2
ERCC-00154	A	7.32421875	1.83105469	4	2
ERCC-00085	A	7.32421875	1.83105469	4	2
ERCC-00028	A	3.66210938	0.91552734	4	2
ERCC-00033	A	1.83105469	0.45776367	4	2
ERCC-00134	A	1.83105469	0.45776367	4	2
ERCC-00147	A	0.91552734	0.22888184	4	2
ERCC-00097	A	0.45776367	0.11444092	4	2
ERCC-00156	A	0.45776367	0.11444092	4	2
ERCC-00123	A	0.22888184	0.05722046	4	2
ERCC-00017	A	0.11444092	0.02861023	4	2
ERCC-00083	A	0.02861023	0.00715256	4	2
ERCC-00096	B	15000	15000	1	0
ERCC-00171	B	3750	3750	1	0
ERCC-00009	B	937.5	937.5	1	0
ERCC-00042	B	468.75	468.75	1	0
ERCC-00060	B	234.375	234.375	1	0
ERCC-00035	B	117.1875	117.1875	1	0
ERCC-00025	B	58.59375	58.59375	1	0
ERCC-00051	B	58.59375	58.59375	1	0
ERCC-00053	B	29.296875	29.296875	1	0
ERCC-00148	B	14.6484375	14.6484375	1	0
ERCC-00126	B	14.6484375	14.6484375	1	0
ERCC-00034	B	7.32421875	7.32421875	1	0
ERCC-00150	B	3.66210938	3.66210938	1	0
ERCC-00067	B	3.66210938	3.66210938	1	0
ERCC-00031	B	1.83105469	1.83105469	1	0
ERCC-00109	B	0.91552734	0.91552734	1	0
ERCC-00073	B	0.91552734	0.91552734	1	0
ERCC-00158	B	0.45776367	0.45776367	1	0
ERCC-00104	B	0.22888184	0.22888184	1	0
ERCC-00142	B	0.22888184	0.22888184	1	0
ERCC-00138	B	0.11444092	0.11444092	1	0
ERCC-00117	B	0.05722046	0.05722046	1	0
ERCC-00075	B	0.01430512	0.01430512	1	0
ERCC-00074	C	15000	22500	0.67	-0.58
ERCC-00113	C	3750	5625	0.67	-0.58
ERCC-00145	C	937.5	1406.25	0.67	-0.58
ERCC-00111	C	468.75	703.125	0.67	-0.58
ERCC-00076	C	234.375	351.5625	0.67	-0.58
ERCC-00044	C	117.1875	175.78125	0.67	-0.58
ERCC-00162	C	58.59375	87.890625	0.67	-0.58

Assess platform dynamic range and lower limit of detection (dose response)

Overview

To assess platform dynamic range and lower limit of detection, use either the ERCC RNA Spike-In Mix kit or the ERCC ExFold RNA Spike-In Mixes kit. Add either Spike-In Mix 1 or Mix 2 to an RNA sample, then process and run the sample on your gene expression platform.

After applying appropriate normalization and filtering to the expression data (RPKM, C_T, RFU, etc.), plot the signal for each ERCC transcript against its known molar concentration or amount, and use linear regression to determine the best-fit line. This plot is often described as a *dose-response curve*.

- The dynamic range can be measured as the difference between the highest and lowest concentration of ERCC transcript detected in each sample.
  - Some platforms, such as microarrays, have a fairly restricted linear range. In such cases the dynamic range can be defined by the lower limit of detection (LLD; described below) and the region of signal saturation. Both linear and non-linear regression can be used to model the dose response and determine these regions empirically. For other methods, refer to documentation for your specific gene expression platform.
  - NGS platforms do not exhibit a region of saturation, so the dynamic range can be determined by observing the concentration difference between the highest-concentration ERCC transcript detected and the LLD.
- The LLD is a measure of sensitivity, and it is defined as the lowest molar amount of ERCC transcript detected in each sample, with user-defined threshold values for determining detection.

Example

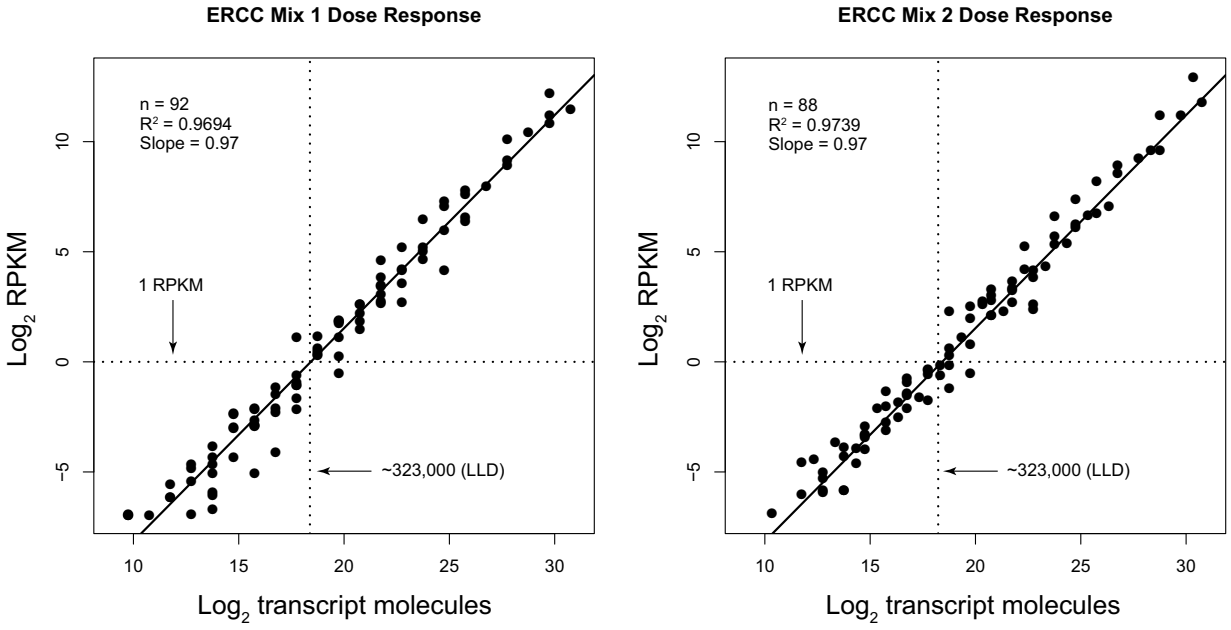
In the example shown in Figure 2 on page 14, cDNA libraries were prepared from HeLa poly(A) RNA containing Spike-In Mix 1 or Spike-In Mix 2 using the SOLiD® Total RNA-Seq Kit, then processed and run on the SOLiD® System. The data were normalized to reads per kilobase of exon model per million mapped reads (RPKM), and filtered using a sensitivity threshold set arbitrarily at 1 RPKM (shown by the horizontal dotted line in Figure 2 at $\log_2$ RPKM = 0; Mortazavi et al., 2008). Using these stringent filtering criteria, the dynamic range of the sequencing runs shown in Figure 2 was determined to be ~12.5 $\log_2$ units of concentration. If a more relaxed criteria is used for the threshold, such as detection with at least 1 mapped read, the dynamic range increases to more than 20 $\log_2$ units.

The X-axis value where the regression line crosses the threshold is the LLD concentration. In the example shown in Figure 2, the X-axis value where the sensitivity threshold of 1 RPKM ($\log_2$ RPKM = 0) crosses the regression line is ~18.3 $\log_2$ molecules. This translates to ~323,000 control molecules detected per 100 ng poly(A) RNA.

The threshold used for defining the LLD may be determined empirically or arbitrarily, or with a combination of these strategies. In the example shown in Figure 2, the threshold was set somewhat arbitrarily based on recommendations from the published literature. The actual threshold used should be based on the sensitivity and accuracy required by the application. For example, in differential expression analysis, high

stringency thresholds (0.3–1.0 RPKM) have been suggested to increase the accuracy of the differential expression calls. Conversely, if the experimental purpose is discovery of alternative splicing or fusion transcripts, lower stringency thresholds to increase sensitivity, such as detection with at least 1 mapped read, may be appropriate.

Figure 2 SOLiD® System dynamic range and lower limit of detection with ERCC RNA Spike-In Control Mixes



Assess platform fold-change response

Overview

To assess platform fold-change response, use the ERCC ExFold RNA Spike-In Mixes kit.

Add ExFold Spike-In Mix 1 and Mix 2 individually to two (or more) RNA samples, then process and run the samples on your gene expression platform.

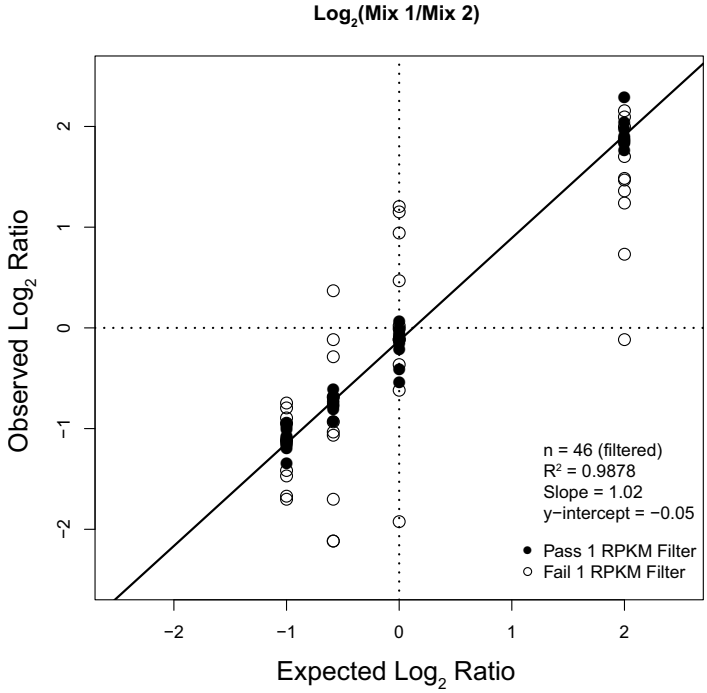
For each ERCC control, compare the observed fold-change ratio (Mix 1:Mix 2) to the expected ratio provided in ERCC_Controls_Analysis.txt. Determine the concordance of the ratios by linear regression to assess the platform performance.

Example

Figure 3 shows the $\log_2$ fold-change ratios formed from Spike-In Mix 1- and Mix 2-containing samples seen in Figure 2 above. The observed Mix 1:Mix 2 fold-change ratios were calculated for each ERCC control and compared with its expected fold-change value. After removing controls with an RPKM less than 1 in either sample, the linear fit shows highly accurate fold-change estimates (filled circles; $R^2 = 0.9878$).

The open circles in Figure 3 show fold-change ratios for the ERCC controls that failed to pass the 1 RPKM threshold in one or more of the ERCC Spike-In Mixes. Adjusting the threshold settings allows you to attain the proper level of accuracy needed for your application.

Figure 3 SOLiD® System fold-change response with ERCC ExFold RNA Spike-In Mixes



Troubleshooting

Observation	Possible cause	Recommended action
Poor correlation of ERCC transcript signal with input amount.	Poor quality of sample RNA.	Confirm RNA integrity using Agilent® bioanalyzer or denaturing gel electrophoresis.
	Contamination of sample RNA with inhibitors of reverse transcriptases or ligases.	Re-purify sample RNA using ethanol precipitation, glass-fiber filter purification, or bead-based methods.
	Improper data analysis or normalization methods.	Review platform provider's instructions for data analysis and normalization where appropriate.
	Sample preparation bias due to gene expression kit.	Consult with the kit manufacturer.
	Up to 5% of the individual transcripts may fall outside the expected regression line.	Contact Technical Service (see page 23) for information about specific manufacturing lots of Spike-In Mix.
Poor correlation of observed fold-change ratios with expected ratios.	Improper data analysis or normalization methods.	Confirm that the same methods were applied to samples containing Mix 1 and those containing Mix 2.
	Mix 1 and Mix 2 from different manufacturing lots were used.	Repeat the sample preparation using Mix 1 and Mix 2 from the same lot.
	Sample preparation bias due to gene expression kit.	Consult with the kit manufacturer.

A Supplemental Information

Additional background on the ERCC plasmid library

The ERCC plasmid reference library produces well-characterized transcripts generated largely from random unique sequences. Sequence comparisons have been made to multiple databases available at the time of design including mouse, rat, human, *Drosophila*, bacteria, mosquito, and other nonhuman species. The control collection contains some sequences with homology to *Bacillus subtilis*.

Potential control sequences were contributed by members of the ERCC and tested across multiple platforms and at a variety of locations. The final set of ERCC control sequences was optimized based on sequence performance in studies designed to evaluate cross-hybridization, quantitative performance in titration experiments, and general performance of the control sequences in complex samples.

The resulting plasmid library incorporates a set of well-characterized sequences that have been cloned with a 20-mer poly(A) tail into the identical public-domain vector suitable for *in vitro* transcription. The sequences of these control sequence-containing plasmids have been certified by NIST as a reference material.

For further information, see External RNA Controls Consortium (2005, 2005a).

Map ERCC reads in an NGS run

The reads generated from an NGS run must be mapped against the ERCC reference sequences and counted against an ERCC feature file using publicly available or proprietary mapping software. We recommend incorporating the ERCC reference into the genome reference, to allow ERCC reads to be mapped and counted in the same mapping run that performs genome mapping and transcript counting. The output from the run includes the ERCC reads as additional transcripts with their associated counts.

ERCC92.fa is available to allow mapping of NGS reads to the ERCC reference sequences. ERCC_bioscope_readme.txt provides instructions for mapping and counting using SOLiD® BioScope™ v1.2 or v1.3 software. See Table 5 on page 18. All files are available through the ERCC RNA Spike-In Control Mixes product web page.

Using the amounts of Spike-In Mix recommended in Table 4 on page 10, <1–5% of the resulting SOLiD® sequencing reads should map to the ERCC reference sequences, depending on the type of starting RNA.

Files for analysis of ERCC data

These files are available through the ERCC RNA Spike-In Control Mixes product web page at **www.appliedbiosystems.com**.

Table 5 Files for analysis of ERCC data

File name	Description
ERCC_Controls_Analysis.txt	A tab-delimited file containing: • The ERCC_IDs • The molar concentration of each transcript in Spike-In Mix 1 and Mix 2 • The expected fold-change ratio for each transcript when comparing Mix 1 to Mix 2
ERCC_Controls_Annotation.txt	A tab-delimited file containing, for each ERCC transcript: • TaqMan® Gene Expression Assay IDs • Sequence information
ERCC_bioscope_readme.txt	Provides instructions for mapping and counting using SOLiD® BioScope™ v1.2 or v1.3 software
ERCC92.fa	FASTA-format file that provides sequences of the ERCC transcripts for mapping
ERCC92.gtf	GTF-format annotation file

Related products

Product	Part number (source)†
SOLiD® Total RNA-Seq Kit	4445374 (LT)
Ambion® RNA-Seq Library Construction Kit	4454073 (IVGN)
Illumina® TotalPrep™ RNA Amplification Kit	AMIL1791 (IVGN)
Illumina® TotalPrep™-96 RNA Amplification Kit	4393543 (IVGN)
MicroPoly(A)Purist™ Kit	AM1919 (IVGN)
RiboMinus™ Eukaryote Kit for RNA-Seq	A10837-08 (IVGN)
TaqMan® Gene Expression Assays for ERCC targets	See below
TaqMan® Gene Expression Master Mix	4369016, 4369510 (LT)

† LT: Life Technologies, **www.lifetechnologies.com**; IVGN: Invitrogen, **www.invitrogen.com**.

TaqMan® Gene Expression Assays for ERCC targets

TaqMan® Gene Expression Assays are available for all 96 sequences in the NIST plasmid reference library. These assays can be useful for monitoring RNA preparatory procedures. For example, you can compare observed to known ERCC transcript amounts and ratios in SOLiD® System whole transcriptome libraries before preparing templated beads from the libraries. In general, TaqMan® Gene Expression Assays for ERCC targets can serve as valuable quality control and troubleshooting tools for RNA-based workflows in your laboratory.

TaqMan® Gene Expression Assay IDs for each ERCC transcript are included in ERCC_Controls_Annotation.txt, which is available through the ERCC RNA Spike-In Control Mixes product web page.

Examples and more information on using ERCC TaqMan® Gene Expression Assays to monitor these controls in RNA-based workflows are available at the ERCC RNA Spike-In Control Mixes product web page.

B

Safety

Chemical safety

WARNING! GENERAL CHEMICAL HANDLING. To minimize hazards, ensure laboratory personnel read and practice the general safety guidelines for chemical usage, storage, and waste provided below, and consult the relevant SDS for specific precautions and instructions:

- Read and understand the Safety Data Sheets (SDSs) provided by the chemical manufacturer before you store, handle, or work with any chemicals or hazardous materials. To obtain SDSs, see the "Documentation and Support" section in this document.
- Minimize contact with chemicals. Wear appropriate personal protective equipment when handling chemicals (for example, safety glasses, gloves, or protective clothing).
- Minimize the inhalation of chemicals. Do not leave chemical containers open. Use only with adequate ventilation (for example, fume hood).
- Check regularly for chemical leaks or spills. If a leak or spill occurs, follow the manufacturer's cleanup procedures as recommended in the SDS.

Obtaining SDSs

Safety Data Sheets (SDSs) are available from **www.appliedbiosystems.com/sds**

Note: For the SDSs of chemicals not distributed by Life Technologies, contact the chemical manufacturer.

Obtaining Certificates of Analysis

The Certificate of Analysis provides detailed quality control and product qualification information for each product. Certificates of Analysis are available on our website. Go to **www.invitrogen.com/support** and search for the Certificate of Analysis by product lot number, which is printed on the box.

Documentation and Support

Related documentation

Document	Publication number (source)†
SOLiD® Total RNA-Seq Kit *Protocol*	4452437 (LT)
Ambion® RNA-Seq Library Construction Kit Protocol	4452440 (IVGN)
Illumina® TotalPrep™ RNA Amplification Kit *Protocol*	See the product web page‡ (IVGN)
MicroPoly(A)Purist™ Kit Protocol	See the product web page‡ (IVGN)
RiboMinus™ Eukaryote Kit for RNA-Seq Manual	See the product web page‡ (IVGN)
TaqMan® Gene Expression Assays Protocol	4333458 (LT)

† LT: Life Technologies, **www.lifetechnologies.com**; IVGN: Invitrogen, **www.invitrogen.com**.

‡ At **www.invitrogen.com**, search for the *product* part number, then find the protocol under **Manuals**.

Obtaining support

For the latest services and support information for all locations, go to:

www.lifetechnologies.com/support

At the website, you can:

- Access worldwide telephone and fax numbers to contact Technical Support and Sales facilities
- Search through frequently asked questions (FAQs)
- Submit a question directly to Technical Support
- Search for user documents, SDSs, vector maps and sequences, application notes, formulations, handbooks, certificates of analysis, citations, and other product support documents
- Obtain information about customer training
- Download software updates and patches

Limited Product Warranty

Life Technologies Corporation and/or its affiliate(s) warrant their products as set forth in the Life Technologies' General Terms and Conditions of Sale found on Life Technologies' website at **www.lifetechnologies.com/termsandconditions**. If you have any questions, please contact Life Technologies at **www.lifetechnologies.com/support**.

Bibliography

External RNA Controls Consortium (2005) Proposed methods for testing and selecting the ERCC external RNA controls. *BMC Genomics* 6:150. Available at **www.biomedcentral.com//1471-2164/6/150**.

External RNA Controls Consortium (2005a) The External RNA Controls Consortium: a progress report. *Nature Methods* 2:731–734.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5:621–628.

Headquarters
5791 Van Allen Way | Carlsbad, CA 92008 USA | Phone +1 760 603 7200 | Toll Free in USA 800 955 6288
For support visit **lifetechnologies.com/support** or email **techsupport@lifetech.com**

lifetechnologies.com

September 2012