USER BULLETIN

applied biosystems®
by *life* technologies™

ERCC_Analysis Plugin

For use with: ERCC RNA Spike-In Control Mixes

Publication Part Number: 4479068   **Revision Date:** 04 June 2012 (Rev. A)

SUBJECT: ERCC Analysis

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 the PGM run. This bulletin provides the following information:

What are the ERCC Spike-in Controls?

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 next-generation sequencing (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.

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.

For more information on ERCC RNA Spike-In Control Mixes, please refer to the *ERCC RNA Spike-In Control Mixes User Guide* (Pub no. 4455352).

life technologies™

Requirements for the ERCC_Analysis plugin

Tool	Requirement
Software	Torrent Suite v2.2 or higher
	IMPORTANT! To ensure proper trimming of the sequences prior to alignment with a reference, for all libraries prepared with the Ion Total RNA-seq Kit v2, you must select an option from the barcode pull-down menu.
Web browser	• Chrome™ Opera®
	• Firefox® Safari®
	• Internet Explorer® (Version 7+)
	Note: In order to upload the reference file described on this page, a browser with either Adobe Flash® or Microsoft Silverlight® software is required. Reference file installation will need to be done only once, and is not a requirement for day-to-day use of the ERCC_Analysis plugin.

Upload and create the reference file

To properly run, the ERCC_Analysis plugin requires a FASTA format reference sequence file. This procedure describes how to:

- Download the FASTA file *ERCC92.fa* to your machine.
- Upload *ERCC92.fa* to a Torrent server and create the reference file.

Download ERCC92.fa

To download the *ERCC92.fa* file:

- Go to the Life Technologies website ERCC RNA Spike-In Mix:
  http://products.invitrogen.com/ivgn/product/4456740
  Scroll to Additional Information and select **Analysis File: ERCC92.fa**.
  – OR –
- Go directly to the *ERCC92.fa* file:
  http://tools.invitrogen.com/downloads/ERCC92.fa
  Then RIGHT-click on the link to save the file to your computer

  Note: If your browser is Internet Explorer, click **Save Target As**. In Firefox or Chrome, click **Save Link As**. In Safari, click **Save Linked File** or **Save Linked File As**.
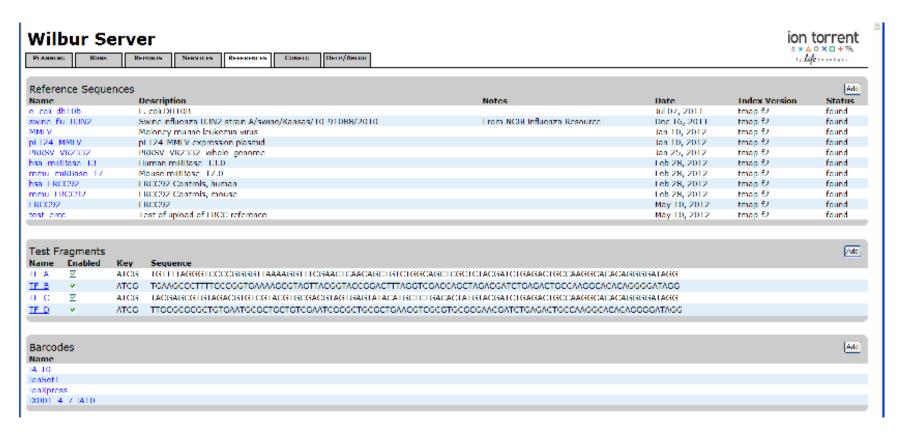
Upload ERCC92.fa and create the reference file

Note: This procedure requires either Microsoft Silverlight or Adobe Flash to upload the *ERCC92.fa* file. We apologize for any inconvenience.

To upload the FASTA file *ERCC92.fa* to the Torrent Suite and create the reference file:

1. Return to the Torrent Browser and click the **REFERENCES** tab, shown on page 3.

2. In the *Reference Sequences* section, click the **Add** button.

3. Enter **ERCC92** in the *Description*, *Short form of genome name*, and *Genome version* fields, and leave the other fields blank.

4. Click **Select file**.

5. Click the **Upload file and create reference** button.

You are returned to the REFERENCES tab, which now shows "ERCC92" in the list of Reference Sequences. Notice that the Status is "found."
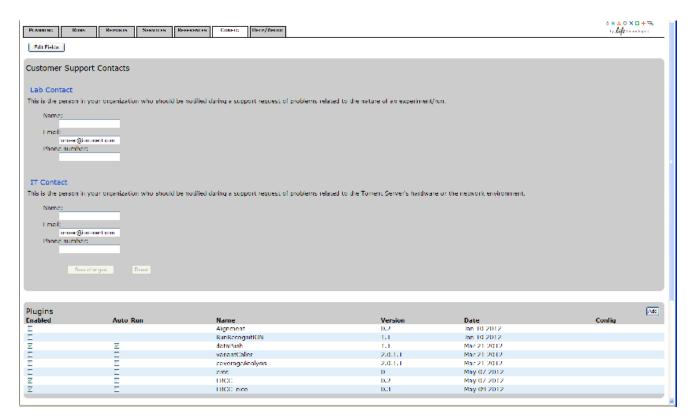
More information   For more detailed instructions about uploading a reference on your Torrent Server:

1. Click the **HELP/ABOUT** tab.

2. Click **Local documentation**, type **upload reference** in the search data entry field, then click the **Search** button.

3. Select the article **Torrent Suite User Documentation: Working With Reference Sequences**.

Enable the ERCC_Analysis Plugin

To enable the ERCC_Analysis plugin:

1. Click the **CONFIG** tab, scroll to the **Plugins** section and the **ERCC** analysis plugin.

2. Select **Enabled** to manually run the plugin. If you want every run to automatically call the ERCC_Analysis plugin, select **Auto-Run**. You can select both Auto-Run and Enabled.
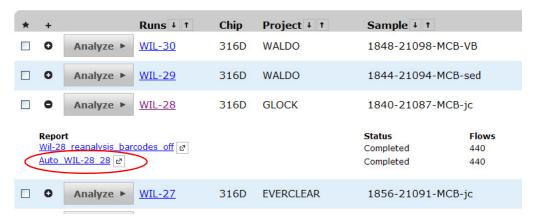
Manually run the ERCC_Analysis plugin

To manually run the ERCC_Analysis plugin:

1. Click the RUNS tab.

2. In the Report column, scroll to the run results you want to view and select it to open the default report.

   Note: To open the report in a new browser window, click the reports icon ⬀ .

3. Scroll to the bottom of the report to **Plugin Summary**, then click the button **Select Plugins To Run** to open the Plugin List.

   If you do not see the ERCC_Analysis plugin listed, refer to the previous section "Enable the ERCC_Analysis Plugin" on page 4.

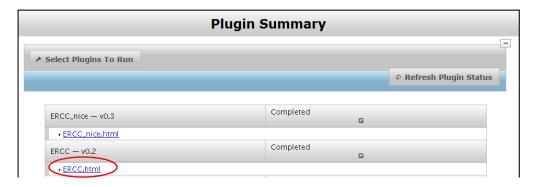4. In the Plugin List window, select **ERCC**.

   The Status of the ERCC_Analysis plugin changes to "Started." You can check the progress of the run by clicking the button **Refresh Plugin Status**.

   For analysis runs with total reads under 1,000,000, the plugin will normally take 2-3 minutes to run on the hardware described on the **Torrent Server** website. For larger runs, the plugin will approximately take an additional 1-2 minutes per million total reads (for example, a run with 5 million reads may take 10-15 minutes). These run times are offered only as a guideline. If your Torrent Server is busy with other processing, run times will be longer.
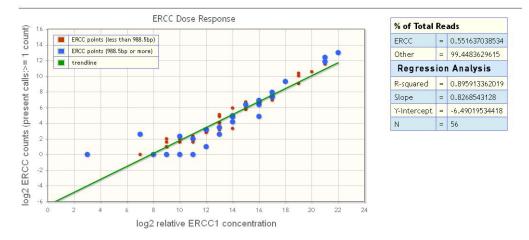
   After the ERCC plugin status shows "Completed," you can view the analysis results.

View analysis results

After the status of the ERCC plugin has changed to "Completed," click **ERCC.html** in the Plugin Summary section to open the ERCC Report and view analysis results:
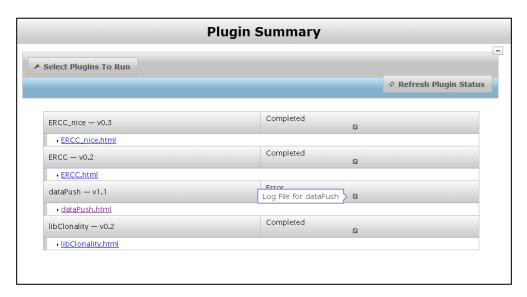

ERCC Report
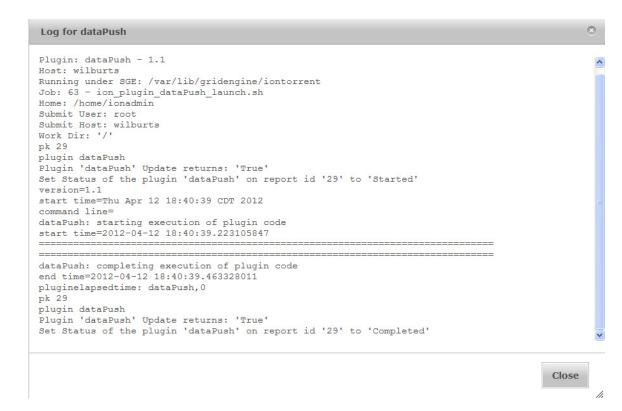
View error information

If the ERCC_Analysis plugin status changes to "Error" after you click ERCC.html, then something went wrong during the running of the plugin. In this case, look at the error log:

1. Click the log file icon in the Plugin Summary section to see the error log:

2. Scroll to the bottom of the system log file to view information about what went wrong with the analysis.

   An Error status for the ERCC_Analysis plugin should be a rare event, An error does not indicate that the ERCC_Analysis results are bad, but rather that the ERCC_Analysis plugin itself failed to run.

Interpret the data

The ERCC Report screen (shown on page 6) displays the ERCC Dose Response plot. The points are color-coded, based on whether they represent transcripts that are larger or smaller than the median value of points found. There is also a trendline, based on the parameters shown in tabular form to the right of the graph.

The y-axis of the plot is the log (base 2) of the raw counts found for the transcript in question. The x-axis is also logarithmic, but represents the known relative concentration of the ERCC transcripts. Ideally, the points all fall on a straight line.

More realistically, in the good case, the raw counts and relative concentration should at least correlate with a high R-squared (for example, 0.9 or higher). The table to the right of the plot (shown on page 6) shows the R-squared value found for this plot, as well as the Slope, Y-intercept, and N (number of transcripts found) values. Although there are 92 transcripts in the ERCC mix, it is not expected that all 92 will be detected. The number of transcripts detected depends on the sequencing depth. For more information on ERCC analysis, refer to the following resources:

- Figure 2, *Analysis of ERCC read counts*, in ***Sensitivity of RNA-Seq using Ion semiconductor sequencing a comparison to microarrays and qPCR***
- The Ion Torrent white paper, ***Methods, tools, and pipelines for analysis of Ion PGM™ Sequencer miRNA and gene expression data***
- The information on the **ERCC ExFold RNA Spike-In Mix** product page.

View transcript details

If you want to look at the details regarding a particular transcript, there are two methods you can use:

- Hover your mouse-cursor over a point on the ERCC Dose Response plot, to display a popup window that shows details about that transcript (the name, reads, and coverage plots).

  If several points are very close together on the plot and it is difficult to hover over the point you are interested in, you can zoom in on the plot to more easily distinguish points:

  a. Use your mouse to draw a box around the point of interest and magnify it.

  b. To zoom out to the full view of the ERCC Dose Response plot, either double-click the plot, or click the button labeled **Reset Zoom**.

  – OR –

- Scroll to the particular transcript, and click the **[+]** next to the transcript name. This method shows the same information, plus a few additional pieces. See Appendix A, "Definitions" if the meaning of any of these pieces is unclear.

Appendix A Definitions

Coverage Depth   The minimum and maximum number of reads covering bases in the transcript. If coverage is 100%, the minimum value will be > 0.

Coverage   The number of base positions covered by at least one read.

Start Sites   The number of base positions that are the start site for a read.

Unique Start Sites   The number of start sites that have only one read starting at the site.

Coverage CV   Coefficient of Variation for coverage = average coverage / stddev coverage for the entire transcript.

Appendix B Parameters used for tmap

For information about how the ERCC_Analysis plugin is working, use the `tmap` call:

```
tmap mapall -f $REFERENCE -r ${TSP_FILEPATH_FASTQ} -a 1 -g 0 -n
8 stage1 map1 --seed-length 18 stage2 map2 map3 --seed-length 18
```

$REFERENCE is a constant that refers to the reference file location.

${TSP_FILEPATH_FASTQ} similarly refers to the *.fastq file location.

At the time of publication of this bulletin, it is not possible to change these parameters when you are using the ERCC_Analysis plugin.

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

The information in this bulletin is subject to change without notice.

NOTICE TO PURCHASER: LIMITED USE LABEL LICENSE: 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 and/or its affiliate(s) or their respective owners. Silverlight, Microsoft, and Internet Explorer are registered trademarks of Microsoft Corporation. Safari is a registered trademark of Apple, Inc. Firefox is a registered trademark of Mozilla Foundation. Chrome is a trademark of Google, Inc. Flash and Adobe are registered trademarks of Adobe Systems Incorporated. Opera is a registered trademark of Opera Software ASA.