USER BULLETIN

ion torrent

Create a Planned Run

Using the RNA Seq Analysis Plugin

Publication Number MAN0014286

Revision A

ThermoFisher SCIENTIFIC

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

DISCLAIMER

TO THE EXTENT ALLOWED BY LAW, LIFE TECHNOLOGIES AND/OR ITS AFFILIATE(S) WILL NOT BE LIABLE FOR SPECIAL, INCIDENTAL, INDIRECT, PUNITIVE, MULTIPLE, OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH OR ARISING FROM THIS DOCUMENT, INCLUDING YOUR USE OF IT.

Placeholder for other notices

Important Licensing Information

These products may be covered by one or more Limited Use Label Licenses. By use of these products, you accept the terms and conditions of all applicable Limited Use Label Licenses.

TRADEMARKS

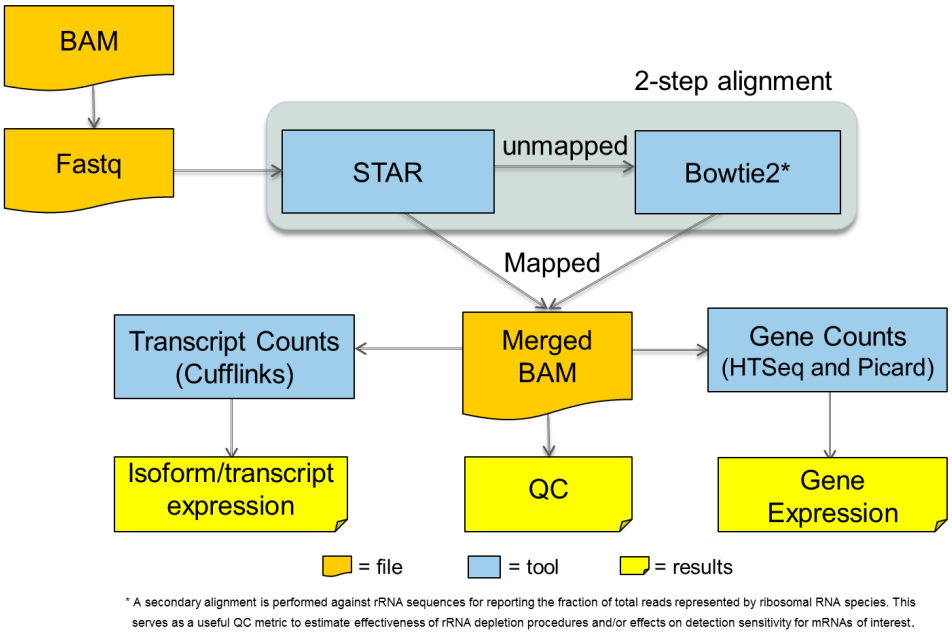
All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified.

Contents

Introduction

The following instructions are intended to assist RNA Seq plugin customers set up Torrent Suite planned runs incorporating the RNA Seq plugin.

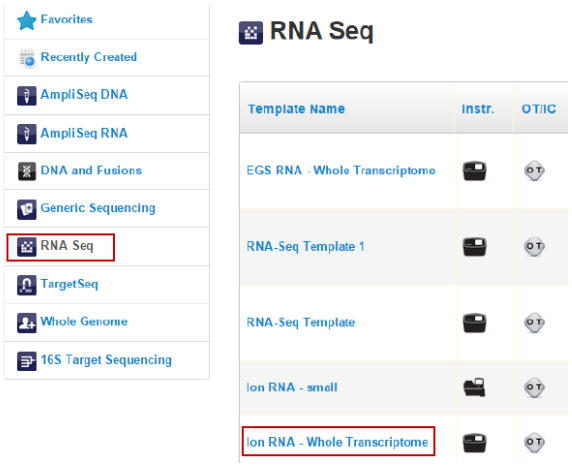
The Torrent RNASeqAnalysis Plugin is an RNA Transcript Alignment and Analysis tool for use with reference genomes hg19 and mm10. Use this plugin to analyze cDNA reads, as produced by RNA-Seq. Reads are aligned to the reference genome using STAR and bowtie2 aligners to find full and partial mappings. The alignments are analyzed by HTSeq and Picard tools to collect assigned read counts and cufflinks to extract gene isoform representation. For barcoded data, comparative representation plots across barcodes are created in addition to individual reports for each barcode. All alignment, detail and summary report files are available for download.



Create an RNA Seq analysis run from factory template

1. Log into Torrent Suite™ Software.
2. Click on the **Plan** Tab.
3. Click on the **Templates** link to view Templates.

4. Click the **RNA Seq-Whole Genome** option under Favorites in the left side navigation list.

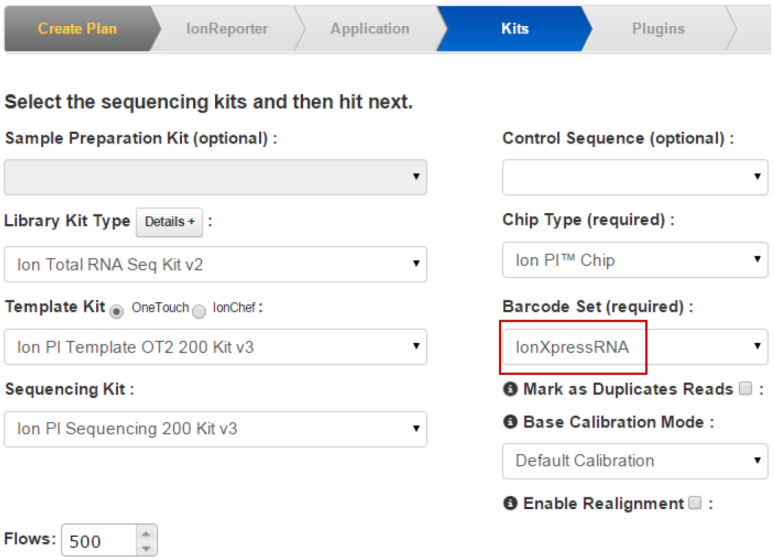


The Plan tab appears.

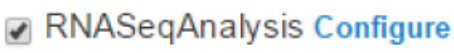
5. Click on the Ion Reporter tab, and click **None** and then click **Next**.

6. Click on the Application tab, select **RNA** and then click **Next**.

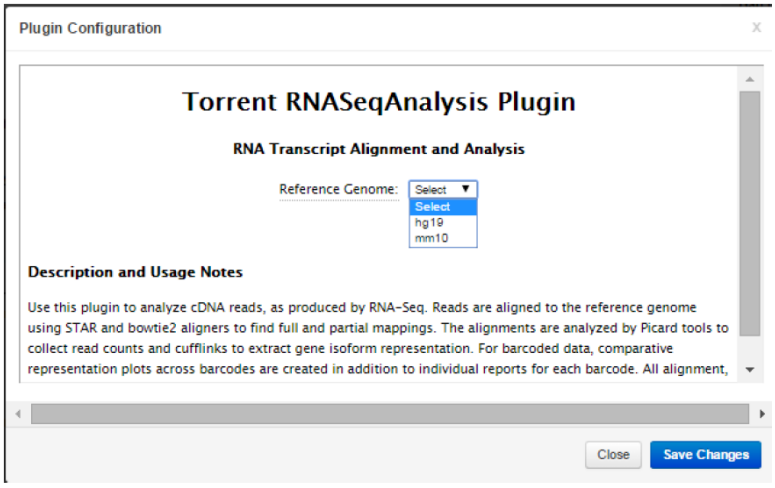
7. On the Kits tab, change Barcode Set to **IonXpressRNA**. Click **Next**.



8. On the Plugins tab, check **RNASeqAnalysis** and click **Configure**.



9. Select human or mouse reference and click **Save Changes**.

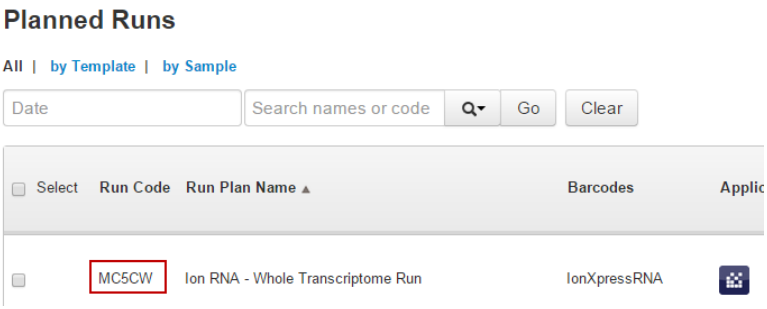


10. Click **Next**.

11. (Optional) On the Projects tab, select a project and click **Next**.

12. On the Plan tab, select **None** as the reference, enter a name for the run and add samples. Click **Plan Run**.

13. Run plan on your sequencing system.
Each planned run contains complete instructions for its sample, from sequencing on instrument to export of the results files to Ion Reporter software. Here is what a planned run from this example looks like on the **Plan ▸ Planned Run** page:



A planned run is ready to execute on the sequencing instrument and is executed by entering the 5-digit run code on the instrument. From the run code, all the plan run's settings are available on the instrument and to the Torrent Suite software. All of your selections, from original template and the planned run that you saved, are known to the Torrent system and software. The system carries out your instructions from sequencing to data export.

Modify RNASeq Template

If you wish to customize the template, create a copy of the template.

1. Click the **Gear** button in the **Ion RNA - Whole Transcriptome** row and select **Copy**.

2. Enter a name for the template and modify settings as required.
3. Click **Copy Template**.

Review run results

After your sequencing run completes, review results on the Run Summary page.

1. Go to **Data** ▸ **Completed Runs & Results** and search for your run. The Run Summary page appears.
2. Scroll down to the **RNASeqAnalysis** section and you'll see the Barcode Summary table. This is an overview table that includes columns for Barcode Name, Sample, Total Reads, Aligned Reads, Percent Aligned Mean Read Length, Genes Detected, and Isoforms Detected.

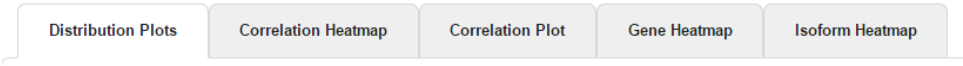
Barcode Summary

Reference: hg19

Barcode Name	Sample	Total Reads	Aligned Reads	Percent Aligned	Mean Read Length	Genes Detected	Isoforms Detected
IonXpress_001	Sample 1	222	222	100%	101	1	131
IonXpress_004	None	2,727,892	2,713,002	99.45%	107.9	9,451	17,723
IonXpress_005	None	2,839,827	2,828,560	99.6%	108.2	9,455	17,346
IonXpress_006	None	2,842,945	2,824,513	99.35%	108	9,336	17,355
IonXpress_007	None	2,642,234	2,623,932	99.31%	107.8	8,860	16,831
IonXpress_008	None	2,757,938	2,737,906	99.27%	108	9,232	17,578
IonXpress_010	None	11,283,208	10,997,469	97.47%	102.4	16,626	58,457
IonXpress_011	None	10,408,573	10,127,505	97.3%	96.7	16,095	57,737
IonXpress_012	None	12,817,747	12,521,818	97.69%	102.2	16,843	58,273
IonXpress_013	None	13,774,036	13,398,953	97.28%	92.4	16,261	58,927
IonXpress_014	None	11,895,869	11,614,056	97.63%	102.9	15,842	57,181

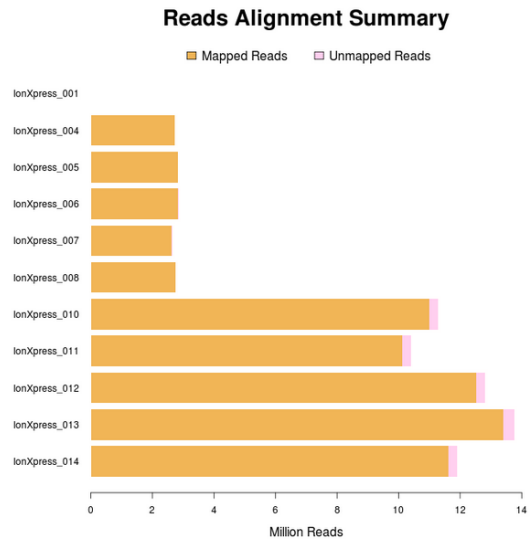
1 20 items per page 1 - 11 of 11 items

3. Click on the **RNASeqAnalysis.html** link to view other components of the report. Below the Barcode Summary are tabs to view the results graphically.

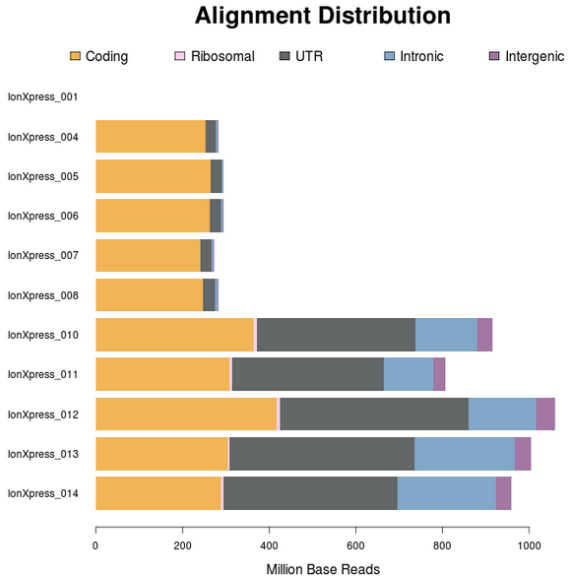


Distribution Plots

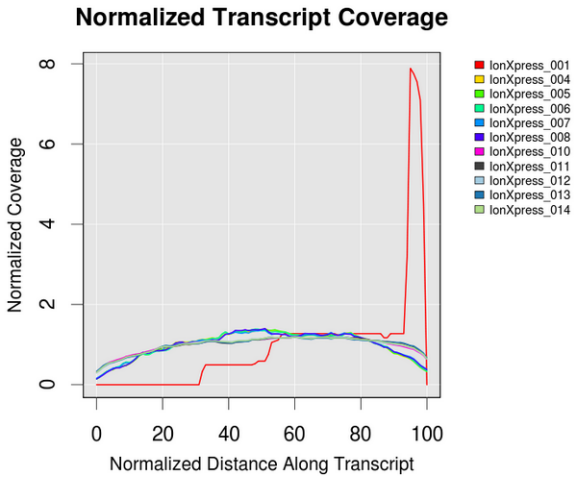
Reads Alignment Summary- A graphical summary of the number of mapped and unmapped reads across barcodes, as reported in the barcode summary table.



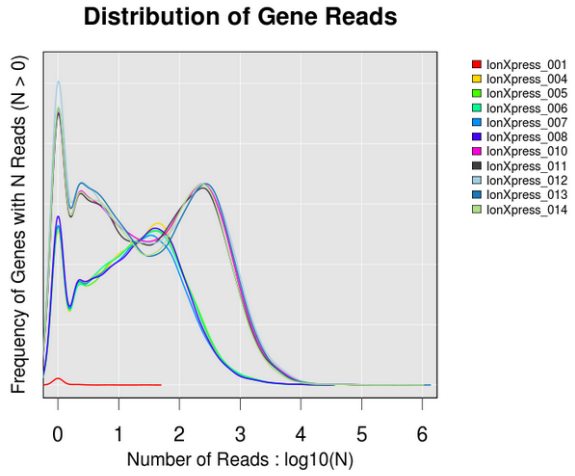
Alignment Distribution - A graphical summary of the distribution of reads to genomic features.



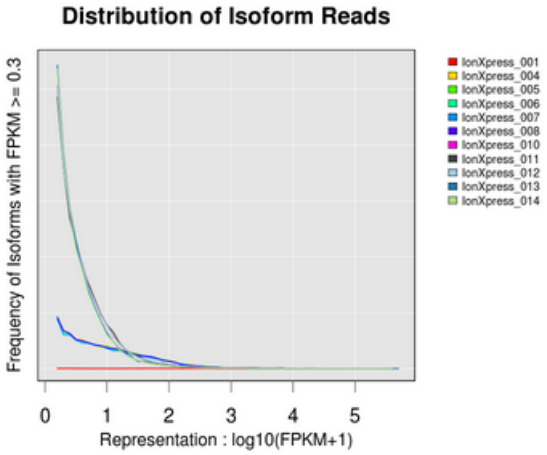
Normalized Transcript Coverage - An overlay of individual normalized transcript coverage plots for each barcode.



Distribution of Gene Reads - Distribution of genes across barcodes showing the frequency of numbers of genes having similar log10 read counts. All curves are plotted on the same axis scale. The counts data is fitted to a Gaussian kernel using the default R 'density' function.

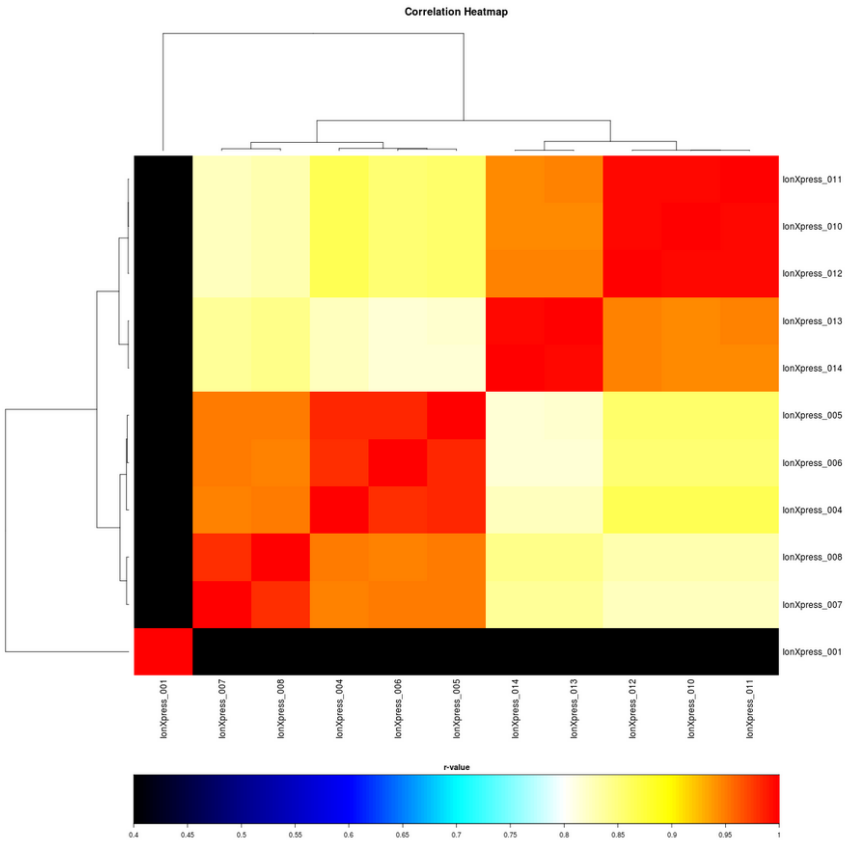


Distribution of Isoform Reads - Distribution of transcript isoforms across barcodes showing the counts of isoforms having similar FPKM values. All curves are plotted on the same y-axis, normalized to the highest count and scaled for FPKM values ≥0.3.



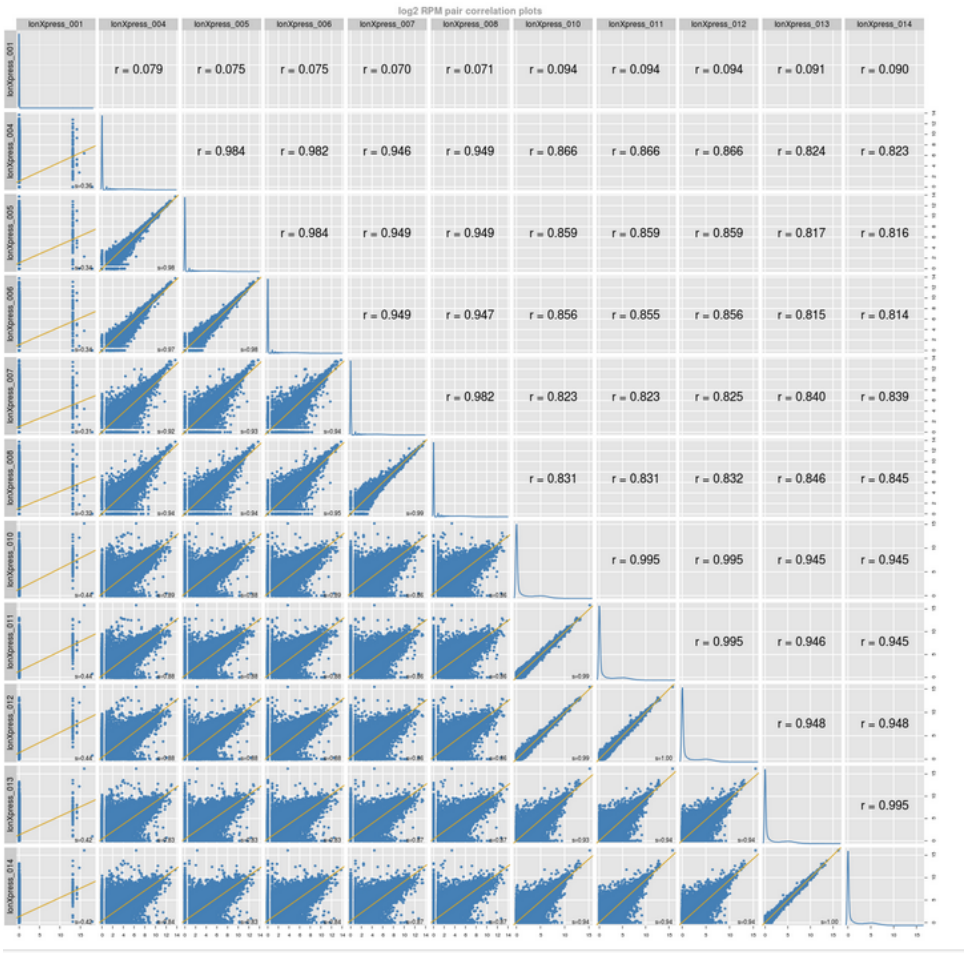
Correlation heatmap

A heatmap of Spearman correlation r-values for comparing log2 RPM reads pair correlation barcodes, with dendrogram reflecting ordering of barcodes as being most similar by these values.



Correlation plot

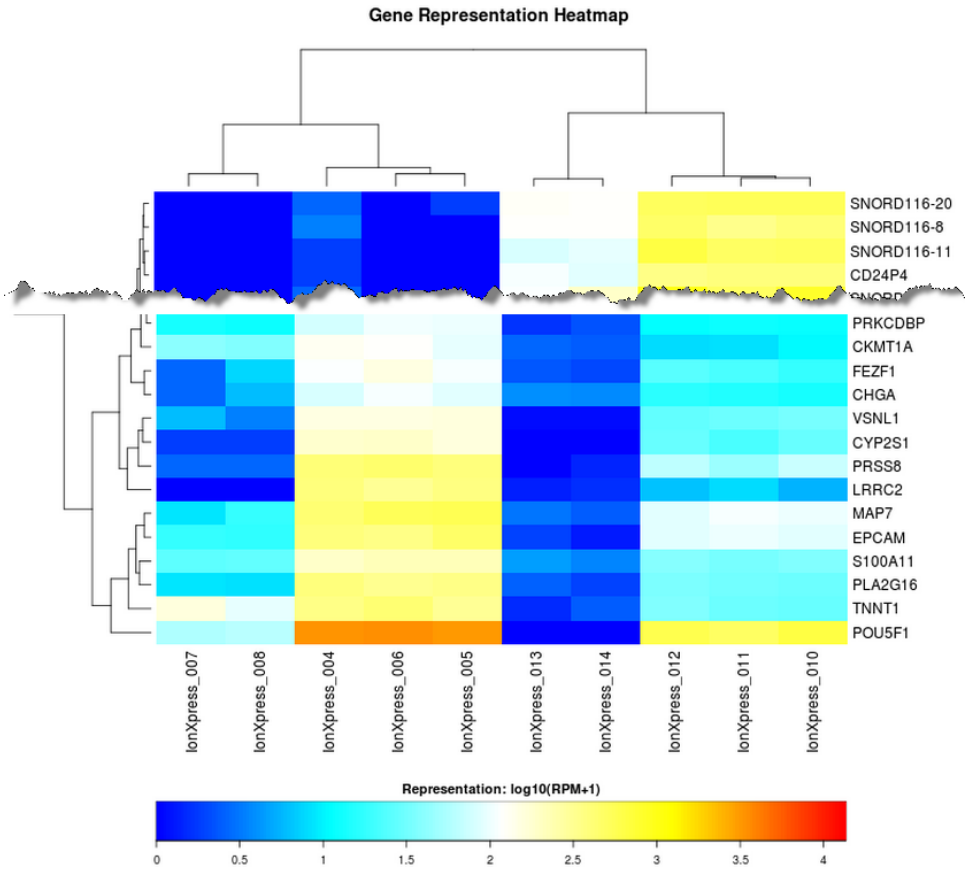
Barcode read pair correlation plot. Lower panels show log2(RPM+1) values plotted for each pair of barcodes, with linear least squares regression line overlaid and line slope reported. Upper panels show Pearson correlation r-values for the regression line. Diagonal panels show the frequency density plot for the individual log(RPM+1) values for each barcode. (If only one barcode has reads, a density plot is displayed.) Click the plot to open an expanded view in a new window.



Gene heatmap

A gene representation heatmap of 250 genes showing the most variation in representation across barcodes as measured by the coefficient of variant (CV) of normalized read counts for genes that have at least one barcode with at least 100 RPM

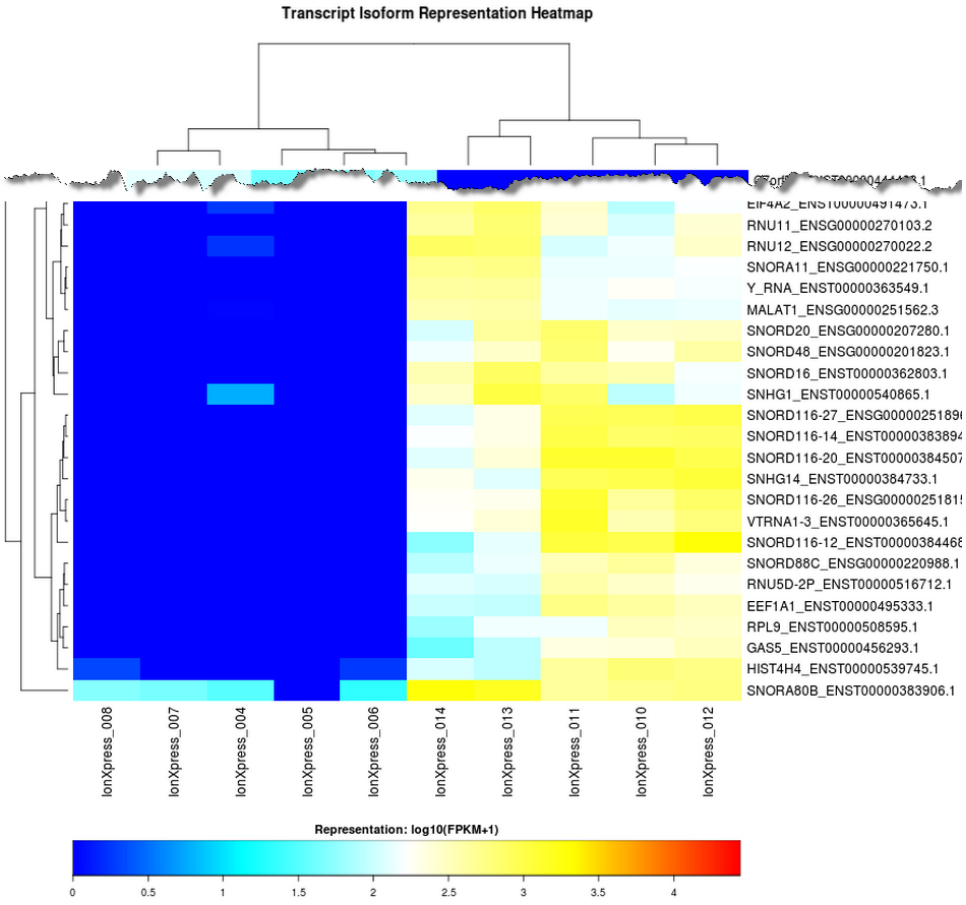
reads, plotted using log10 of those counts. For this plot, barcodes will be omitted if they have less than 100,000 total reads.



Isoform heatmap

A transcript isoform representation heatmap of up to 250 gene transcript isoforms showing the most variation in representation across barcodes as measured by the coefficient of variation (CV) of FPKM values for isoforms that have an FKPM value ≥

100 for at least one barcode, plotted using log10 of FKPM+1. Barcodes are excluded if they have less than 1,000 isoforms detected at FPKM values ≥ 0.3.



Downloadable reports

At the bottom of the screen are links for downloading raw analysis output files:

Download Barcode Summary Report
Download absolute reads table
Download absolute normalized reads table
Download aligned reads distribution table
Download isoform FPKM values table

Barcode Summary Report - This report produces a Microsoft® Excel® table listing each barcode's sample name, total reads, aligned reads and percent aligned.

Absolute Reads Table - This Microsoft® Excel® table lists absolute reads for the genes found on each barcode.

Absolute Normalized Reads Table - This Microsoft® Excel® table lists absolute normalized reads for the genes found on each barcode.

Aligned Reads Distribution Table - Distribution of genes across barcodes showing the frequency of numbers of genes having similar log10 read counts.

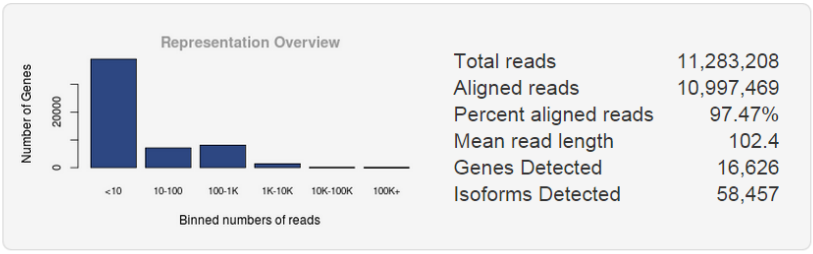
Isoform FPKM Values Table - Table format of the Isoform gene heatmap.

Individual barcode view

Click on any barcode of interest to see similar graphs of the barcode alone.

Reference table - Plot showing the number of genes with reads in log10 counting bins.

Reference: hg19



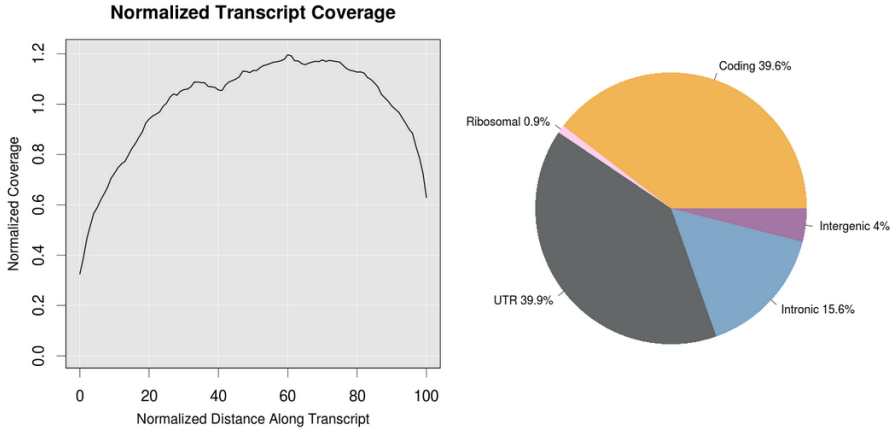
Gene Mapping Summary - Summary of reads mapped to genes of annotated reference.

Gene Mapping Summary	
Reference genes	55,765
Reads mapped to genes	7,390,706
Genes with 1+ reads	26,969
Genes with 10+ reads	16,626
Genes with 100+ reads	9,531
Genes with 10,00+ reads	1,429
Genes with 10,000+ reads	35
Isoforms Annotated	230,756
Isoforms Detected	58,457

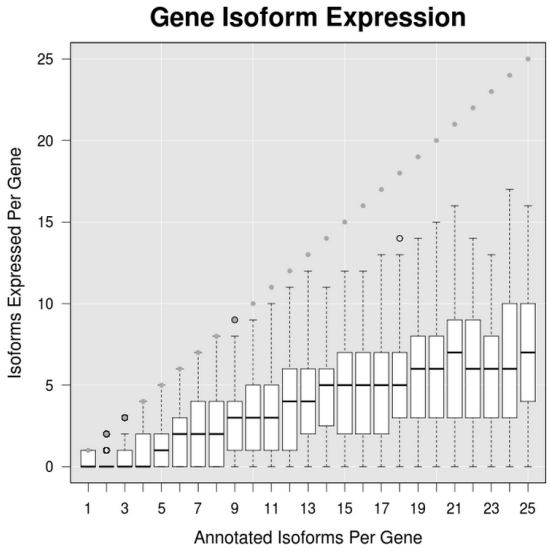
Base Mapping Summary - Summary of base reads aligned to genetic features of an annotated reference.

Base Mapping Summary	
Total base reads	1,155,834,791
Total aligned bases	914,778,477
Percent aligned bases	79.14%
Percent coding bases	39.68%
Percent UTR bases	39.93%
Percent ribosomal bases	0.94%
Percent intronic bases	15.65%
Percent intergenic bases	3.98%
Strand balance	0.4980

Normalized Transcript Coverage - A plot of normalized transcript coverage; the frequency of base reads with respect to the length of individual transcripts they are aligned to in the 3' to 5' orientation.



Gene Isoform Expression - Box plots showing variation of isoforms expressed at FPKM ≥ 0.3 for each set of genes grouped by the number of anticipated (annotated) isoforms. Whiskers are defined by points within Q1-1.5xIQR to Q3+1.5xIQR. Only genes with 25 or less isoforms are represented in this plot. The data and a plot for all genes are available for download using the download reports links at the bottom of the screen.



Downloadable reports

At the bottom of the screen are links for downloading raw analysis output files for the individual barcodes:

- Download the Statistics Summary
- Download the Gene Read Counts
- Download Output Files (page)
- Download Cufflinks Output Files (page)

Statistics Summary - Provides an overview of the individual barcodes RNA Seq Analysis results.

```
RNASeqAnalysis Summary Report

Sample Name: None
Reference Genome: hg19
Adapter Sequence: None
Reads Sampled: 100.0%
Alignments: IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes

Total Reads:      11283208
Aligned Reads:    10997469
Pct Aligned:      97.47%
Mean Read Length: 102.4
Strand Balance:   0.4980

Reference Genes:  55765
Reads Mapped to Genes: 7390706
Genes with 1+ reads: 26969
Genes with 10+ reads: 16626
Genes with 100+ reads: 9531
Genes with 1000+ reads: 1429
Genes with 10000+ reads: 35

Total Base Reads:     1155834791
  Pct Aligned Bases:  79.14%
  Pct Usable Bases:   63.01%
Total Aligned Bases:  914778477
Pct mRNA Bases:       79.61%
  Pct Coding Bases:   39.68%
  Pct UTR Bases:      39.93%
Pct Ribosomal Bases:  0.94%
Pct Intronic Bases:   15.65%
Pct Intergenic Bases: 3.98%

Isoforms Annotated: 230756
Isoforms Detected:  58457
```

Gene Read Counts - Lists the number of times a gene was counted for the individual barcode.

	A	B
1	Gene	Reads
2	5S_rRNA	3
3	7SK	547
4	A1BG	3
5	A1BG-AS1	34
6	A1CF	0
7	A2M	14
8	A2M-AS1	16
9	A2ML1	45
10	A2ML1-AS	0
11	A2ML1-AS	0
12	A2MP1	0
13	A3GALT2	0
14	A4GALT	45
15	A4GNT	0
16	AAAS	492

Output Files - Provides a directory for various output files for this barcode.

File Size	Date	File
871M	2015-06-02	alignedSTAR.bam
72M	2015-06-02	Chimeric.out.junction
495M	2015-06-02	Chimeric.out.sam
90	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.bam
27K	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.geneisoexp_all.png
19K	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.geneisoexp.png
1.3M	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.geneisoexp.xls
660K	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.genereads.xls
4.4K	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.generep.png
129	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.genes.fpkm_tracking
132	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.isoforms.fpkm_tracking
19K	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.mareads.png
107	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.mareads.xls
121	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.skipped.gtf
1.3G	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.STARBowtie2.bam
3.5M	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.STARBowtie2.bam.bai
660K	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.STARBowtie2.gene.count
20K	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.STARBowtie2.RNAmetrics.png
2.9K	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.STARBowtie2.RNAmetrics.txt
897	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.stats.txt
125	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.transcripts.gtf
1.7K	2015-06-02	Log.final.out
12K	2015-06-02	Log.out
32K	2015-06-02	output_cufflinks
19K	2015-06-02	maseq.log
5.0M	2015-06-02	SJ.out.tab
92	2015-06-02	xrRNA.bam
2	2015-06-02	xrRNA.basereads

Cufflinks Output Files - Provides a list of links to Cufflinks output files.

File Size	Date	File
5.5M	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.genes.fpkm_tracking
24M	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.isoforms.fpkm_tracking
0	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.skipped.gtf
305M	2015-06-02	IonXpress_010_R_2015_02_12_15_16_34_sc_P19-753-P2bead_on_p1--R79599_Update_for_less_barcodes.transcripts.gtf

For support visit **thermofisher.com/techresources** or email **techsupport@lifetech.com**
lifetechnologies.com

26 June 2015

ThermoFisher SCIENTIFIC