# Droplet Digital RT-ddPCR™: Ultra-high Sensitivity Validation Technology for RNA-Seq

Shenglong Wang, Shawn Hodges, Dimitri Skvortsov, Svilen Tzonev, Serge Saxonov, and George, Karlin-Neumann

QuantaLife/Bio-Rad, 7068 Koll Center Parkway, Suite 401, Pleasanton, California 94566

## Abstract

As RNA-Seq increasingly assumes the discovery role once played by DNA expression microarrays, a highly precise and ultra-sensitive validation technology is needed to confirm its findings. Droplet digital PCR™ (ddPCR) is also a digital technology which counts individual molecules with high precision and linearity over a 5 log range. With its extremely low false-positive rate, it is possible to detect as little as a few molecules in a sample where precision is only limited by inherent sampling error. Furthermore, the minimal sample processing necessary in either 1-step or 2-step RT-ddPCR allows for maximal fidelity of determined transcript concentrations. In addition, where sample amount is less limited but high sensitivity is desired as for detecting a few percent of cells expressing a marker in a tumor or in plasma, relatively large amounts of RNA (>1ug of either total or polyA RNA) can be readily and accurately assayed, giving multiple logs greater sensitivity than achievable w/ 200M RNA-Seq reads. The greater simplicity and directness of the ddPCR process eliminates distortion of the sample composition and loss of sensitivity due to sampling error in RNA-Seq sample preparation. Comparisons between the two technologies and their inherent complementarity will be illustrated.

## Materials and Methods

- RNA-Seq library preparation was performed with Illumina Tru™Seq RNA Sample Prep Kit v2 following manufacture's protocol.
- Sequencing was performed on MiSeq.
- Human brain reference RNA and ERCC control Mix 1 and Mix 2 were purchased from Ambion.
- Four RNA-Seq libraries were generated with ERCC Mix spiked into human brain reference RNA as follows:
  1. 100ng human brain total RNA + 2ul 1:1000 diluted ERCC Mix 1
  2. 100ng human brain total RNA + 2ul 1:1000 diluted ERCC Mix 2
  3. 1000ng human brain total RNA + 2ul 1:100 diluted ERCC Mix 1
  4. 1000ng human brain total RNA + 2ul 1:100 diluted ERCC Mix 2
- Same spiked materials were also used in RT-ddPCR. The cDNA was made with Applied Biosystems MultiScribe, up to 10% RT reactions were loaded to each well of ddPCR™ assay.
- ddPCR™ was performed using the Bio-Rad QX100 platform including standard mastermix and reagents for droplet generation and reading.
- Assays were purchased from Applied Biosystems at 20x concentration
- For panel 1 only, cDNA was generated with Bio-Rad iScript.

TaqMan® Gene Expression Assays used in this study:

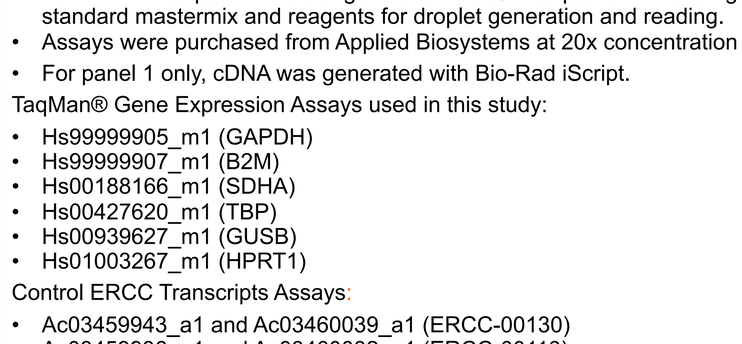
- Hs99999905_m1 (GAPDH)
- Hs99999907_m1 (B2M)
- Hs00188166_m1 (SDHA)
- Hs00427620_m1 (TBP)
- Hs00939627_m1 (GUSB)
- Hs01003267_m1 (HPRT1)

Control ERCC Transcripts Assays:

- Ac03459943_a1 and Ac03460039_a1 (ERCC-00130)
- Ac03459936_a1 and Ac03460032_a1 (ERCC-00113)
- Ac03459884_a1 and Ac03459980_a1 (ERCC-00022)
- Ac03459902_a1 and Ac03459998_a1 (ERCC-00053)
- Ac03459922_a1 and Ac03460018_a1 (ERCC-00084)
- Ac03459911_a1 and Ac03460007_a1 (ERCC-00069)
- Ac03459931_a1 and Ac03460027_a1 (ERCC-00104)
- Ac03459921_a1 and Ac03460017_a1 (ERCC-00083)

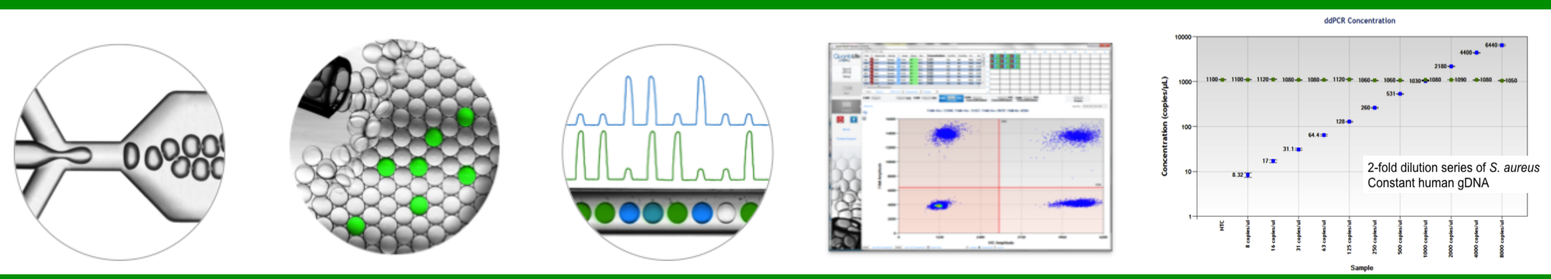
## ① ddPCR is an analytical tool with great precision and sensitivity

**ddPCR measurement of cDNA concentration: Dynamic Range, Sensitivity and Precision**

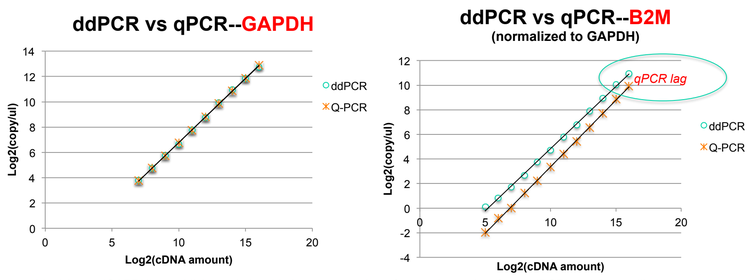



**Example of cDNA concentration measurement by ddPCR.** cDNA was generated with Bio-Rad iScript Advanced kit and 2-fold serially diluted. Two independent measurement sets were made, one at high concentration range, one at a lower range, with 4 points overlapping.

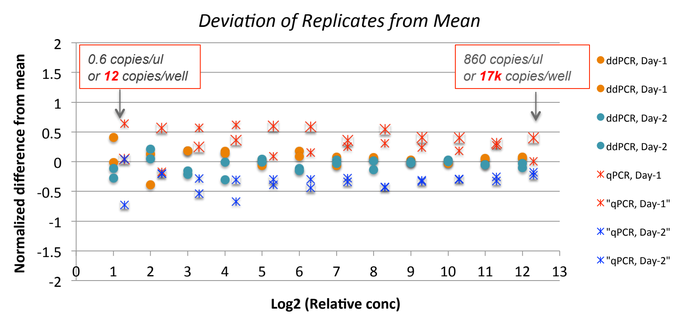
Generate | Cycle | Read | Analyze | Evaluate




## ② ddPCR is a quantification tool that does not require a standard curve



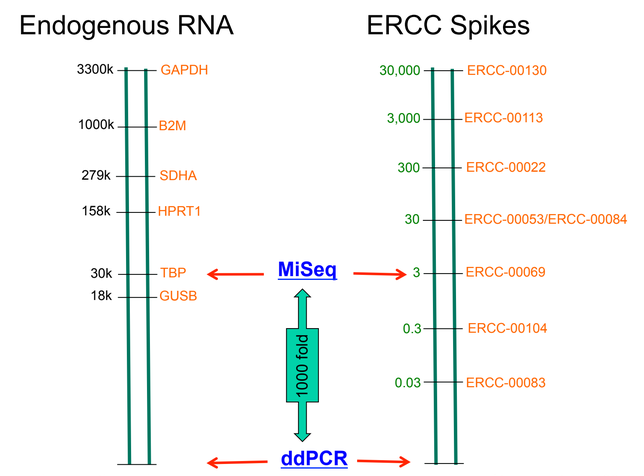

- Ideal Assay (GAPDH), establishes relationship between number of transcript copies (from ddPCR) and Ct per dilution.
- B2M deviates from this ideal behavior and lags behind expected Ct.




- Duplicates of the same 2-fold cDNA dilution series run on both platforms on each of 2 days

- In this experiment, the cDNA was generated from human brain total RNA with the Applied Biosystem High Capacity cDNA Reverse Transcription Kit. A serial dilution of cDNA was made for both ddPCR and qPCR. Components in ddPCR and in qPCR were identical for direct comparison.
- The ddPCR absolute concentration measurements were used to calibrate $C_t$ to concentration for GAPDH. This relationship was then applied to B2M.
- The top right figure shows that qPCR for B2M was lagging, meaning that the same number of B2M molecules detected by ddPCR required more cycles to reach the threshold than that of GAPDH molecules.
- The bottom figure shows that two identical qPCR runs done on two different days could differ by as much as two cycles in $C_t$ number, but ddPCR yielded identical concentrations on both days.

## ③ ddPCR has ~1000X lower detection limit than RNA-Seq




The numbers on the endogenous gene scale above were copy numbers of transcripts in 100ng human brain total RNA measured by RT-ddPCR.

**Table 1a. Detection of House-Keeping Genes: ddPCR sensitivity is enhanced proportional to the input amount of RNA; RNA-Seq is not.**

| Gene ID | ddPCR (copy/well) | | MiSeq (RPKM) | | | |
|---|---|---|---|---|---|---|
| | 100ng RNA 4 replicates | 1000ng 4 replicates | 100ng Two replicates | | 1000ng Two replicates | |
| GAPDH | 1671±115* | 16275±479* | 974 | 953 | 1077 | 1061 |
| B2M | 504±46* | 3450±155* | 233 | 251 | 229 | 237 |
| SDHA | 139±14* | 1131±81* | 51 | 50 | 59 | 65 |
| HPRT 1 | 15781±2310 | 140705±11059 | 34 | 22 | 25 | 25 |
| TBP | 3650±178 | 31625±1010 | 3 | 3 | 6 | 3 |
| GUSB | 1794±53 | 15731±1134 | 9 | 12 | 13 | 12 |

Numbers marked with * were obtained with cDNA diluted 200 fold, due to their high abundance.

- In this experiment, four combinations of human brain total RNA and ERCC Mixes were subjected to RNA-Seq library prep or RT-ddPCR (see M&M).
- Table 1a above shows that ddPCR detected thousand of copies per well on the lower abundance transcripts (TBP and GUSB genes) in 100ng total RNA input, but MiSeq only detected a single digit RPKM on the same transcripts.
- When the RNA input was increased 10-fold to 1000ng, ddPCR detected 10-fold more copies on each transcript, but RNA-Seq detected the same RPKM.

This is because detectability in RNA-Seq was limited by the total reads each run can produce; ddPCR on the other hand, can handle a much larger amount of material, therefore achieving higher detection sensitivity.

**Table 1b. Detection of ERCC Spike-Ins: ddPCR sensitivity is ~1000 fold greater than MiSeq**

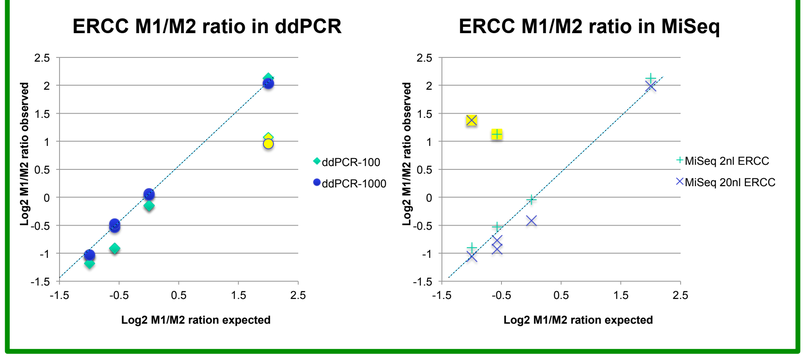
| | ERCC Stock | | ddPCR (copy/well) | | | | MiSeq (RPKM) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mix 1 | Mix 2 | 0.002ul ERCC | | 0.02ul ERCC | | 0.002ul ERCC | | 0.02ul ERCC | |
| | Stock (amol/ul) | | Mix 1 | Mix 2 | Mix 1 | Mix 2 | Mix 1 | Mix 2 | Mix 1 | Mix 2 |
| ERCC-00130 | 30000.00 | 7500.00 | 1456* | 335* | 7775* | 1925* | 21117 | 4835 | 24223 | 6119 |
| ERCC-00113 | 3750.00 | 5625.00 | 195* | 368* | 1523* | 2199* | 2833 | 4081 | 3186 | 5430 |
| ERCC-00022 | 234.38 | 468.75 | 2156 | 4475 | 15325 | 31375 | 137 | 256 | 167 | 349 |
| ERCC-00084 | 29.30 | 43.95 | 414 | 775 | 2863 | 4000 | 17 | 8 | 22 | 41 |
| ERCC-00053 | 29.30 | 29.30 | 346 | 388 | 2588 | 2476 | 22 | 23 | 19 | 25 |
| ERCC-00069 | 1.83 | 3.66 | 26 | 59 | 179 | 364 | 0 | 0 | 4 | 2 |
| ERCC-00104 | 0.23 | 0.23 | 16 | 18 | 49 | 48 | 0 | 0 | 0 | 0 |
| ERCC-00083 | 0.03 | 0.01 | 4 | 2 | 10 | 5 | 0 | 0 | 0 | 0 |

Numbers marked with * were obtained with cDNA diluted 200 fold, due to their high abundance.

- Droplet ddPCR can detect almost every transcript that is converted to cDNA. In Table 1b above, the lowest abundance spike-ins in the 0.002ul sample have a few tens of molecules in the whole RT reaction (only 10% were assayed/well).
- The precision and accuracy of ddPCR detection of such low abundance molecules are determined by sampling, not its detectability.
- In RNA-Seq, the samples have to go through a lengthy library prep procedure, where several steps in the process are known to be very inefficient, such as polyA selection and ligation of cDNA to adaptors, which lead to the permanent loss of low abundant transcripts.

## ④ Ratio of transcripts in ERCC Mix 1 vs Mix 2 are well maintained in ddPCR

| | ERCC Stock (amol per ul) | | | M1/M2 in ddPCR | | M1/M2 in MiSeq | |
|---|---|---|---|---|---|---|---|
| Transcript | Mix 1 | Mix 2 | M1/M2 | Spiked in100ng | Spiked in1000ng | Spiked in100ng | Spiked in1000ng |
| ERCC-00130 | 30000.00 | 7500.00 | 4 | 4.36 | 4.09 | 4.37 | 3.96 |
| ERCC-00113 | 3750.00 | 5625.00 | 0.67 | 0.53 | 0.69 | 0.69 | 0.59 |
| ERCC-00022 | 234.38 | 468.75 | 0.5 | 0.48 | 0.49 | 0.54 | 0.48 |
| ERCC-00084 | 29.30 | 43.95 | 0.67 | 0.53 | 0.72 | 2.18 | 0.53 |
| ERCC-00053 | 29.30 | 29.30 | 1 | 0.90 | 1.04 | 0.97 | 0.75 |
| ERCC-00069 | 1.83 | 3.66 | 0.5 | 0.44 | 0.49 | | 2.60 |
| ERCC-00104 | 0.23 | 0.23 | 1 | 0.90 | 1.03 | | |
| ERCC-00083 | 0.03 | 0.01 | 4 | 2.10 | 1.94 | | |




## Conclusions

**ddPCR is ideally suited for validation of RNA-Seq discoveries in both low and high throughput workflows for multiple reasons.**

**ddPCR is:**

- **Precise**, **accurate** and **reproducible** over ~5 logs, and **sensitive** enough to detect as little as a few molecules/sample.
- **~1000X more sensitive** than RNA-Seq (assuming 1 ddPCR well and 1 HiSeq lane).
- **Low cost**: Cost of running a few ddPCR wells is at least 100-fold less than a single run on a MiSeq or larger NGS sequencer.
- **High fidelity**: Requires minimal manipulation of the RNA sample (only cDNA synthesis) before ddPCR.
- **Versatile**: Works equally well with total or polyA-selected RNA, using all types of cDNA synthesis priming (gene-specific, N6, oligo dT).
- **Unbiased**: Allows unbiased interrogation of transcripts all along their length, ***whether intact or fragmented*** (eg. FFPE, plasma).
- **Simple and fast to implement and run**: Uses standard Taqman assay chemistry and thermocyclers.

## Contact

Shenglong Wang and
George Karlin-Neumann

Bio-Rad/QLBU
7068 Koll Center Pkwy, Suite 401
Pleasanton CA 94566

BIO-RAD