Optimization of RNA extraction from FFPE tissues for expression profiling in the DASL assay

Mark Abramovitz[1,2], **Maja Ordanic-Kodani**[1,3], **Yuefang Wang**[3], **Zhenhong Li**[2], **Charles Catzavelos**[4], **Mark Bouzyk**[5], **George W. Sledge Jr**[6], **Carlos S. Moreno**[3,7], and **Brian Leyland-Jones**[3]

2 *VM Institute of Research, Montreal, Quebec*

3 *Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia*

4 *Department of Pathology, St. Mary's Hospital, McGill University, Montreal, Quebec*

5 *Center for Medical Genomics, Emory University School of Medicine, Atlanta, Georgia*

6 *Indiana University Cancer Center, Indiana Cancer Pavilion, RT-473, 535 Barnhill Dr., Indianapolis, IN 46202*

7 *Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, Georgia 30322*

Abstract

Formalin-fixed paraffin-embedded (FFPE) breast tumor tissues are readily available and represent a largely untapped, vast resource for molecular profiling of clinical samples with long-term follow-up data. We have optimized the conditions and parameters that result in the preparation of total RNA that is of the necessary quality for use in the DASL (cDNA-mediated annealing, selection, extension and ligation) assay in which expression of 502 genes are analyzed simultaneously using as little as 100 ng of input RNA.

Keywords

FFPE; RNA isolation; DASL; genomics; extraction; expression profiling; microarray

Formalin-fixed paraffin-embedded (FFPE) tissue samples make up a vast archive of pathologically well-characterized clinical samples from randomized trials and are an immense virtually untapped resource that can be used for conducting biomarker investigations. Even though the degradation of RNA that occurs due to the formalin fixation process results in RNA species with an average size of ~200 nt (1), it is feasible to extract and purify RNA from such FFPE tissue and to perform real-time reverse transcription-polymerase chain reaction (RT-PCR)-based gene expression profiling. Some studies have addressed improvements to the process of isolating high quality FFPE RNA suitable for RT-PCR or high-throughput genome-wide gene expression profiling (2-4).

The DASL (cDNA-mediated Annealing, Selection, extension and Ligation) assay, that is based upon massively multiplex RT-PCR applied in a microarray format that allows for the determination of expression of up to 512 genes (502 genes in the Cancer panel used in this study) using RNA isolated from 96 FFPE tumor tissue samples in a high throughput format (5,6). The DASL assay has been recently used to identify a 16 gene set that correlates with prostate cancer relapse (7).

Here, we have compared four FFPE RNA preparation methodologies resulting in an optimized protocol for performance in the DASL assay. Although Illumina Inc., San Diego, CA recommends using the High Pure RNA Kit (Roche, Mannheim, Germany) based on earlier studies (5,6), several new kits (e.g. Ambion, Qiagen, and SuperArray) have not been previously tested for utility in downstream DASL assays. In our experiments, FFPE blocks from six different estrogen receptor positive (ER+), epidermal growth factor receptor 2-positive (HER2 +) (as determined by IHC) breast cancer patients were sectioned to enable direct comparison of the same tissues with six individual samples. All samples were then run in the DASL assay as two technical replicates. Here we report our findings on performance of these FFPE RNA extraction methods and finalized optimal protocol in the DASL assay.

Total cellular RNA was extracted with overall yields ranging from 0 to 9 μg with concentrations from 0 to over 300 ng/μL. For each kit, longer digestion with Proteinase K yielded more RNA (Figure 1A), consistent with previous studies (2). Often limited information may be available on how the tissue was processed prior to fixation and embedding, but it likely strongly influences the quantity and quality of nucleic acids recovered from the sample.

Most kits tested in this study produced RNA with an A260/A280 ratio close to two (Figure 1B). The most variability in the A260/A280 ratio was seen when using the SuperArray FFPE RNA Isolation Kit. When the A260/A280 ratio decreases, there is a tendency for the RPL13a TaqMan Ct values to increase and the DASL replicate reproducibility to decrease. Agilent 2100 Bioanalysis determined that the median RNA size was approximately 100-200 bp (Figure 1E), with profiles typical for FFPE RNA (8), but RIN values were not predictive of the utility of RNA samples.

Of greater predictive power for assessment of RNA performance in the DASL assay is TaqMan real-time PCR analysis of RPL13a (6). TaqMan analysis suggested that most samples with a Ct value of 29 or lower had sufficient quality RNA to give reproducible results in the DASL assay (Figure 1C). As the Ct value increased, the replicate reproducibility value decreased. The lowest Ct values, and thus the most usable RNA, were obtained for RNA prepared with the Ambion and Roche kits paired with overnight Proteinase K digestion and an RNA concentration of ≥ 20 ng/μL. The RNA prepared with the Qiagen kit achieved higher Ct values in this assay at both the 15 min and 3 hr Proteinase K digestion timepoints compared with Ambion or Roche at the 3 hr or overnight Proteinase K digestion timepoints.

In order to determine which RNA isolation method yielded the highest quality RNA for the DASL assay (Figure 1D), we analyzed A260/A280 ratios, RPL13a Ct values, and DASL assay replicate data. The results for all analyses are summarized in Table 2. We found that by using at least 100 ng (preferably 200 ng) of input RNA (at 20 ng/μL), an A260/A280 ratio ≥ 1.5, and RPL13a TaqMan assay Ct values ≤ 29, we could achieve a Log R^2 value of > 0.9 for 92% of the replicate samples (n = 26) in the DASL assay. Therefore, these were the minimum RNA quality parameters required for obtaining an acceptable replicate reproducibility score in the DASL assay (see Table 1 for optimized protocol).

Although many manufacturers recommend a shorter Proteinase K digestion in order to save time, the additional incubation time significantly increases RNA yield and overall quality. Upon examination of replicate reproducibility between two samples from the same patient

(Figure 2), longer Proteinase K digestion also produced tighter scatter plots with higher Log R^2 values in the DASL assay for all kits tested. The best replicate reproducibility was achieved with the Roche kit employing an overnight Proteinase K digestion, but the Ambion kit also gave comparable reproducibility with an overnight digestion (Figures 1D and 2).

We were not able to compare the performance of RNA from fresh frozen with RNA from FFPE samples because frozen specimens were not available from this cohort of patients. However, previous studies have found that the DASL assay generates data that is highly correlated between frozen and fixed tissues from autopsies (8). Our technical replicate reproducibility was generally higher than those studies using similar RNA preparation methods, which may be due to the fact that autopsy specimens may have slightly greater RNA degradation than surgical specimens.

In the DASL assay, the number of detected genes ($p < 0.01$) and reproducibility of technical replicates (Log R^2 values) is largely determined by the signal to noise ratio. We found that while the expression patterns of RNA prepared using different kits clustered together for each patient, regardless of the RNA extraction method (Figure 3A), certain kits performed better. In fact, the DASL replicate reproducibility of samples prepared from the same patient using different RNA isolation methods was dependent upon the kits used, indicating that RNA preparation methods had important effects on the overall gene expression pattern (Figure 3B). There were significant differences in gene expression between patients, as evidenced by both our clustering and scatterplot comparisons (Figures 3A and 3B, respectively). A heat map generated for all 502 genes for the 6 ER and, HER2 positive samples (Figure 3C) showed that there were a substantial number of differentially expressed genes, pointing to significant and potentially important differences in the gene expression patterns between patient samples from the same major breast cancer subtype. These data suggest that the DASL assay is robust, and while quite tolerant to differences in RNA extraction methods, optimization and standardization of the RNA extraction protocol will enable the undertaking of future large studies.

Acknowledgements

References

1. Masuda N, Ohnishi T, Kawamoto S, Monden M, Okubo K. Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. Nucleic Acids Res 1999;27:4436–4443. [PubMed: 10536153]

2. Chung JY, Braunschweig T, Hewitt SM. Optimization of recovery of RNA from formalin-fixed, paraffin-embedded tissue. Diagn Mol Pathol 2006;15:229–236. [PubMed: 17122651]

3. Rupp GM, Locker J. Purification and analysis of RNA from paraffin-embedded tissues. Biotechniques 1988;6:56–60. [PubMed: 2483655]

4. Penland SK, Keku TO, Torrice C, He X, Krishnamurthy J, Hoadley KA, Woosley JT, Thomas NE, et al. RNA expression analysis of formalin-fixed paraffin-embedded tumors. Lab Invest 2007;87:383–391. [PubMed: 17297435]

5. Bibikova M, Talantov D, Chudin E, Yeakley JM, Chen J, Doucet D, Wickham E, Atkins D, et al. Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. Am J Pathol 2004;165:1799–1807. [PubMed: 15509548]

6. Fan JB, Yeakley JM, Bibikova M, Chudin E, Wickham E, Chen J, Doucet D, Rigault P, et al. A versatile assay for high-throughput gene expression profiling on universal array matrices. Genome Res 2004;14:878–885. [PubMed: 15123585]

7. Bibikova M, Chudin E, Arsanjani A, Zhou L, Garcia EW, Modder J, Kostelec M, Barker D, et al. Expression signatures that correlated with Gleason score and relapse in prostate cancer. Genomics 2007;89:666–672. [PubMed: 17459658]

8. Haller AC, Kanakapalli D, Walter R, Alhasan S, Eliason JF, Everson RB. Transcriptional profiling of degraded RNA in cryopreserved and fixed tissue samples obtained at autopsy. BMC Clin Pathol 2006;6:9. [PubMed: 17144909]
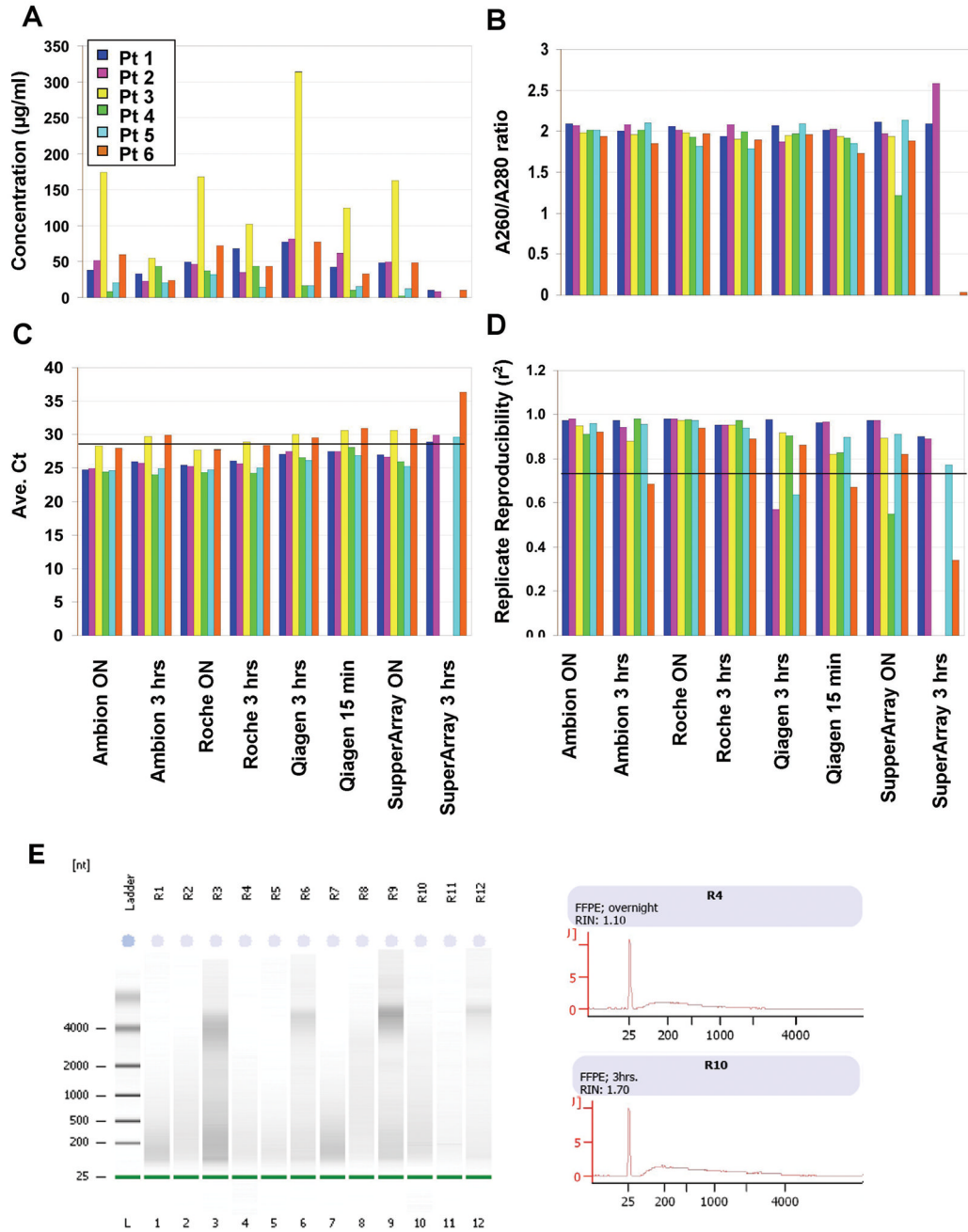
Figure 1. Quality control analysis of total RNA samples prepared using four commercially available kits

(A) Concentration of samples varied from sample to sample, and from kit to kit, but it was clearly tissue-dependent. The overnight incubation with Proteinase K increased the yield of total RNA across the methods tested. (B) The A260/A280 ratio was close to the ideal value of two, with the exception of the SuperArray kit. (C) The lowest Ct values were obtained using the Ambion and Roche kits, while the SuperArray kit provided the least amount of usable RNA. (D) Replicate reproducibility was highest in samples that were incubated overnight, with the best results obtained using the Ambion and Roche kits. (E) Representative Agilent 2100

Bioanalysis of the same patient (sample 4) using the Roche kit with overnight and 3hr Proteinase K digestion is shown. Median RNA size is approximately 100-200 nt.
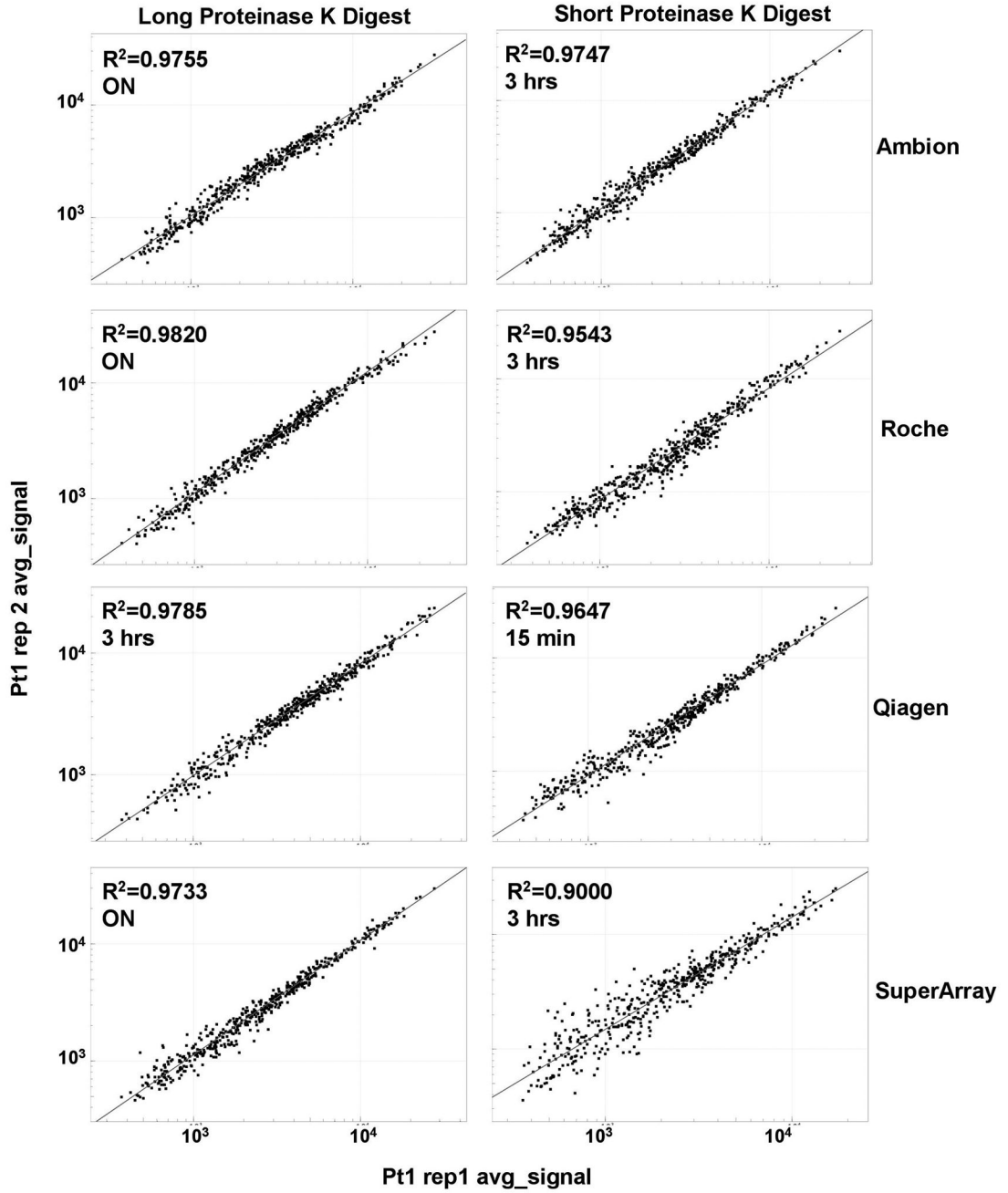
Figure 2. Longer Proteinase K digestion increases the replicate reproducibility

The R^2 value between two replicates of the same sample was superior for samples that were treated with Proteinase K overnight than for the samples that were incubated with Proteinase K for only three hours or less. In addition, the replicate reproducibility of the samples incubated in Proteinase K for a shorter time was the best for the Ambion kit, then Qiagen, Roche, and finally SuperArray kit.
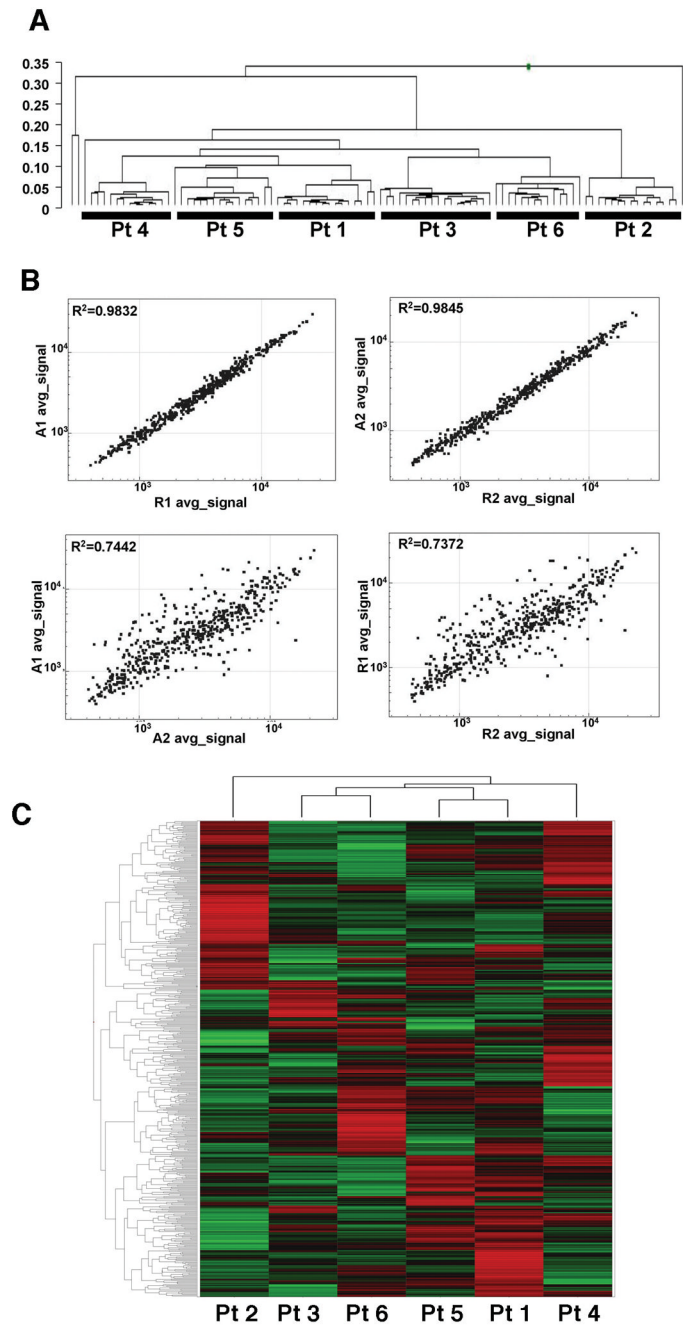
Figure 3. Samples cluster by patient regardless of the RNA isolation method used
(A) Clustering of all of the patients using all of the RNA samples, 8 per patient. In an initial analysis of data for the 6 patients obtained in the DASL assay, clustering of the samples grouped by patient regardless of the kit or method used. This suggests that the DASL assay can tolerate, to some extent, differences in RNA quality. (B) Scatterplots of gene expression patterns of samples from Patient 1 prepared with Ambion (A1) or Roche (R1) kits and of Patient 2 prepared with Ambion (A2) or Roche (R2) kits. The replicate reproducibility of these scatterplots indicates that there were significant differences in gene expression between patients 1 and 2 and high correlation of expression patterns for these patients' RNA prepared by the two different RNA extraction methods. (C) Hierarchical clustering of Z-score normalized data from

six patients. Each column is the average of four DASL assays, two technical replicates prepared with the Ambion kit, and two technical replicates prepared with the Roche Kit. All four were prepared with an overnight Proteinase K digestion.

Table 1
Optimized Protocol for RNA preparation for DASL assay

Optimized Protocol

Cut three 5 μm sections per block, and place into a 1.5 mL sterile microfuge tube

Deparaffinize with 100% xylene for 3 min at 50 °C, centrifuge, perform two ethanol washes, centrifuge, and air dry

Digest with Proteinase K at 50 °C overnight

Prepare RNA using Ambion RecoverAll Kit

Perform nanodrop quantitation and Rpl13a Taqman assay

Use samples with 200 ng RNA (20 ng/ul), A260/A280 > 1.5, and Rpl13a C_T < 29

Table 2

RNA QC values for each sample preparation. The average and range are given for the R^2 correlation for technical replicates, the RNA Integrity Number (RIN) from the Agilent Bioanalyzer, the Cycle Threshold (CT) for the TaqMan Rpl13a assay, the DNA concentration, and the A_{260}/A_{280} ratios.

	DASL R^2	RIN	Rpl13A CT	[DNA] ng/ul	A260/A280
Ambion O/N	0.95 (0.91 - 0.98)	1.2 (0 - 2.3)	25.8 (24.4 - 28.3)	58.8 (8.4 - 173.6)	2.0 (1.9 - 2.1)
Ambion 3hr	0.90 (0.69 - 0.98)	1.0 (0 - 1.7)	26.7 (24 - 30)	33.0 (21 - 54.2)	2.0 (1.9 - 2.1)
Roche O/N	0.97 (0.94 - 0.98)	1.5 (1.1 - 2.3)	25.9 (24.3 - 27.8)	67.3 (31.5 - 167.9)	2.0 (1.8 - 2.1)
Roche 3 hr	0.94 (0.89 - 0.97)	1.8 (1.0 - 3.0)	26.4 (24.2 - 28.9)	50.8 (14.5 - 101.6)	1.9 (1.8 - 2.1)
Qiagen 3hr	0.81 (0.57 - 0.98)	2.0 (1.1 - 2.4)	27.8 (26.2 - 30.1)	97.2 (16.2 - 314.3)	2.0 (1.9 - 2.1)
Qiagen 15 min	0.86 (0.67 - 0.97)	1.7 (0 - 2.4)	28.5 (26.8 - 30.9)	47.6 (10.3 - 124.5)	1.9 (1.7 - 2.0)
SuperArray O/N	0.85 (0.55 - 0.97)	0.9 (0 - 1.7)	27.7 (25.2 - 30.8)	53.9 (1.6 - 163)	1.9 (1.2 - 2.1)
SuperArray 3 hr	0.73 (0.34 - 0.9)	0.3 (0 - 1.0)	31.2 (28.9 - 36.4)	9.6 (8.3 - 10.3)	1.2 (0 - 2.6)