

High-Throughput and Quantitative Detection of Residual NS0 and CHO Host Cell Genomic DNA

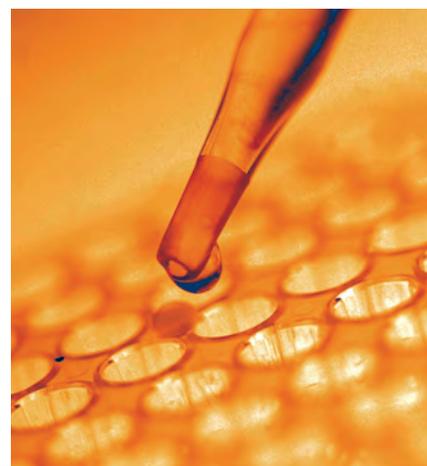
Daryl Venable, Guillermo Miro-Quesada, John Calley, Ellen Monson*, and Luhong He

Residual host cell DNA removal in the processing of biological pharmaceuticals is an important metric for consistency of the purification process. There are also safety concerns with respect to the oncogenic potential of residual DNA from continuous cell lines. The acceptable residual amount of DNA is 10 ng/dose for parentally administrated drugs produced from continuous cell lines, based on WHO (World Health Organization) guidelines (1). Traditional methods to evaluate residual DNA level during process development, such as the Threshold assay, are expensive, low throughput, and labor intensive (Molecular Devices Corporation Threshold total DNA assay system, www.moleculardevices.com).

To provide quick and high-throughput evaluation of the DNA

removal process, we developed TaqMan qPCR (quantitative polymerase chain reaction) assays coupled with 96-well DNA extraction methodology. To meet the sensitivity requirements, we required that the TaqMan target sequence for the hosts be at least 100 base pairs (bp) long, present in multiple copies, and interspersed throughout the host genome. Our expectation was that dispersed target sequences with high occurrence in the genome would yield assays of greater sensitivity and guard against the possibility of nondetection of genome portions lacking the targeted sequence.

Two TaqMan assays (B1 and R-repeat assays) were developed for detection of genomic DNA of mouse cells (NS0), and three (CHO-49, CHO-63 and CHO G-repeat assays) were designed for detection of genomic DNA of Chinese hamster ovary cells (CHO). An additional TaqMan assay was designed to target luciferase sequence carried by a pGL3-Basic vector, which served as a spike-control to evaluate host cell DNA extraction and TaqMan quantitation. The 96-well extraction conditions were optimized through statistical experimental design. TaqMan assays showed good correlation with the Threshold assay when evaluating samples from actual purification process streams. The R^2 is 0.9843 in log scale of DNA, with 95% confidence intervals obtained from the replicated assays.



WWW.PHOTOS.COM

EXPERIMENTAL: MATERIALS AND METHODS

TaqMan PCR: The primer-probe sequence was designed using Applied Biosystems' Primer Express software (www.appliedbiosystems.com). All probes were TAMRA quenched, 5'-VIC labeled (except CHO49-probe with 6-FAM at its 5') and acquired from Applied Biosystems. Two primer-probe sets, referred to as B1 and R-repeat assays, were designed for detection of mouse genomic DNA (2, 3). Three primer-probe sets were designed for detection of CHO DNA. Two sets, referred to as CHO49 and CHO63, were designed against CHO Alu-equivalent repeats, the sequence of clone 49C and Clone 63, respectively (4); and the third, CHO-G repeat, was designed against the novel hamster G-repeat sequence (5). The primer-probes sequence (5'-3') is listed in the "Sequences" box.

PRODUCT FOCUS: MONOCLONAL ANTIBODIES

PROCESS FOCUS: DOWNSTREAM PROCESSING

WHO SHOULD READ: PROCESS DEVELOPMENT AND PRODUCT QC SCIENTISTS

KEYWORDS: RESIDUAL HOST CELL DNA REMOVAL, TAQMAN qPCR AND THRESHOLD ASSAYS, NS0, CHO, 96-WELL DNA EXTRACTION

LEVEL: INTERMEDIATE

SEQUENCES (5'-3')

TaqMan PCR

B1 Assay: Forward primer ACG CCT TTA ATC CCA GCA CTT; reverse primer-GCT GTC CTG GAA CTC ACT TTG TAG; probe AGA GGC AGG CGG ATT TCT GAG TTC GA

R-Repeats Assay: Forward primer ACA GAA GTG GAT GCT CAC AGT CA; reverse primer GCT CCT TGG GCA CTT TCT CTA GC; probe CAC AGG GCC CCC AAT GGA GGA

CHO49 Assay: Forward primer TGGAGAGATGGCTCGAGGTT; reverse primer TGGTTGCTGGGAATTGAACTC; probe- AGAGCAC CAACTGCTCTCCAGAGGTCC

CHO63 Assay: Forward primer CACACCTTAGAGGGCCAGTAC; reverse primer CAGCTGGTCTGGAACTCACATAA; probe CTGCCCGCTCTGCCTCTGG

CHO-G Repeat Assay: Forward primer TGGAGGGTGAGTCCTTCCAA; reverse primer CACAAGGGTAAGGGCCTAGA; probe CCTGCTCCAAATGATATGAC AGACTTTGAAGATCT.

Plasmid Control DNA TaqMan Assay (fLuc Assay)

Forward primer AAG CGA CCA ACG CCT TGA T; reverse primer AGT GTT CGT CTT CGT CCC AGT AAG; probe-TAT GTC TCC AGA ATG TAG CCA TCC ATC CTT GT

The primer-probe set for our plasmid control DNA TaqMan assay, referred to as the *fLuc assay*, was targeted to the luciferase reporter sequence of the plasmid pGL3-Basic vector (Invitrogen, www.invitrogen.com). Its 5'-3' sequence is also listed in the "Sequences" box.

DNA used to generate standard curves in TaqMan assays for NS0 DNA and CHO DNA was isolated from parental cell lines of NS0 and CHOK1 using Qiagen's blood and cell culture DNA Maxi Kits (www1.qiagen.com) according to the vendor's protocol. The plasmid was prepared by Hind III linearization followed by phenol/chloroform extraction and ethanol precipitation. The DNA standards and control plasmid pGL3-Basic DNA standards were quantitated by a Molecular Probes PicoGreen assay (<http://probes.invitrogen.com>), which is validated against A260. Mouse genomic DNA was purchased from Clontech (www.clontech.com).

All TaqMan PCR assays were performed using 7900HT sequence detection system (Applied Biosystems). Unless otherwise stated, all data points consisted of four replicates. Each TaqMan qPCR reaction well contained 10 μ L of input sample DNA template, 12.5 μ L 2 \times TaqMan PCR Master Mix (Applied Biosystems), 900 nM forward and reverse primer, 250 nM probe and 30 ng/ μ L baker's yeast tRNA (Sigma) in a total reaction volume of 25 μ L. The thermocycler conditions were 50 °C for two minutes, 95 °C for 10 minutes, then 40 cycles of 95 °C for 15 seconds followed by 60 °C for one minute. All data were analyzed using sequence detection system version 2.0.

DNA Extraction: All 96-well format DNA extractions were performed with Qiagen's QIAamp 96 DNA blood kit. Following the characterization of the TaqMan probe and primer sets and DNA standards, a statistical experimental design approach guided the optimization of the 96-well DNA extraction protocol with the vendor's protocol as a starting

point. The vendor's extraction protocol was optimized for additives [Poly(A), Poly(dA), BSA, Amersham Biosciences, carrier RNA (tRNA)], wash times and elution conditions (number of elutions, elution centrifugation speed, elution centrifugation time). Investigators at Genentech had previously published that adding BSA and Poly(A) as carriers before extraction dramatically improved recovery of MVM genomic DNA in viral clearance assessments of mammalian cell substrate-derived biotherapeutics (6).

RESULTS AND DISCUSSION

Primer-Probe Design and TaqMan qPCR: Highly repetitive and dispersed DNA families are characteristic of the genomes of mammalian species, which consist of long or short interspersed nuclear elements (LINEs or SINEs, respectively). We designed five primer-probe sets to target the SINEs of both NS0 (B1 and R repeats) (2, 3) and CHO (CHO49, CHO63 and CHO-G repeats) (4, 5) genomes, respectively.

Figure 1 demonstrates that the CHO49 assay detected a broad range of CHO DNA from 10 fg to 10⁷ fg (CHO-K1 genomic DNA as template). The log-linear regression plot shows a strong linear relationship ($R^2 > 0.99$) between the log of the DNA quantity and the threshold cycle (Ct) value, which is defined as the number of PCR cycle at which an increase in fluorescence above a baseline signal is first detected. The greater the number of target copies in the reaction, the lower the Ct value (7). Both CHO63 and CHO-G-repeat assays gave much lower sensitivity than did CHO49 assays (data not shown). Therefore, they were not further evaluated.

We analyzed 21 standard curves generated by the CHO49 assay to evaluate their consistency. The average slope of the standard curves is -3.38, which represents 96% efficiency of the PCR. The standard deviation of the slope is 0.1, and the RSD of the slope is 3%. The slopes of other CHO genomic DNA samples tested independently were all within two standard deviations of the average slope. We concluded that there is acceptable consistency in the standard curves.

Figure 2 demonstrates that both B1 and R-repeats assays detected a broad range of mouse DNA from 1 to 10⁶ fg (Clontech mouse genomic DNA as template). The log-linear regression plot shows a strong linear relationship ($R^2 > 0.99$) between the log of the DNA quantity and the Ct value. The B1 assay gave slightly lower Ct values for equivalent template concentrations, which indicates increased sensitivity.

We analyzed 10 standard curves generated by the B1 assay to evaluate for consistency. The average slope of the standard curves is -3.32 with standard deviation of 0.06 (2% RSD). We concluded that there is acceptable level of consistency in the standard curves.

Well-to-well variability within a single plate for both NS0 assays was assessed using 100 fg of Clontech mouse genomic DNA per well in the PCR reactions. Each of 92 wells received 100 fg of input template DNA, and four wells were used as no-template controls (NTCs). The B1 assay rendered a mean Ct value of 28.98 with a standard deviation

Figure 1: CHO49 assay

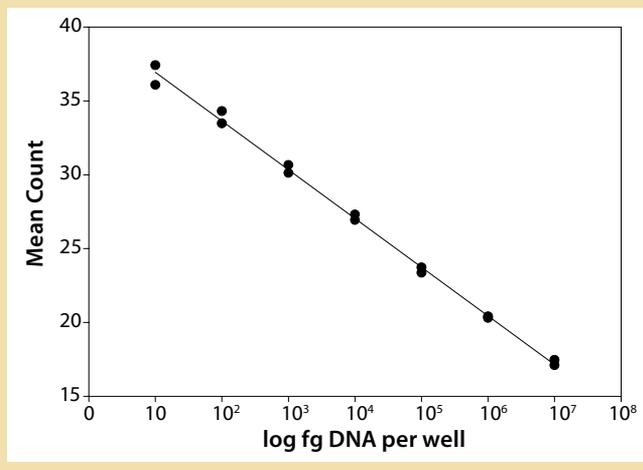


Figure 2: Comparison of R-repeat and B1 assays

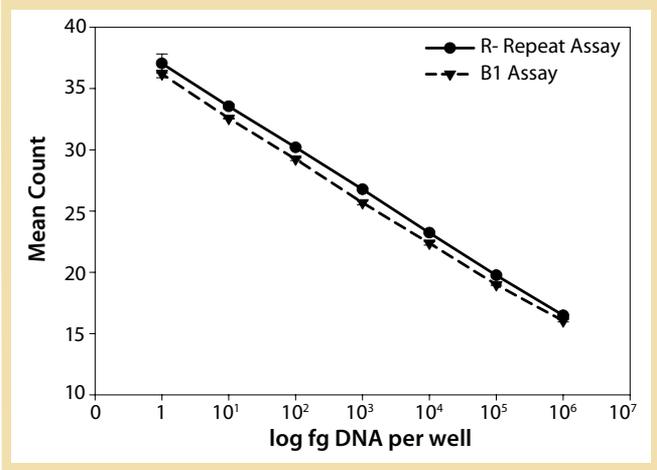


Figure 3A: Intact NS0 DNA as standards and KpnI, NdeI, SspI digested NS0 DNA as unknowns (B1 probe/primer set)

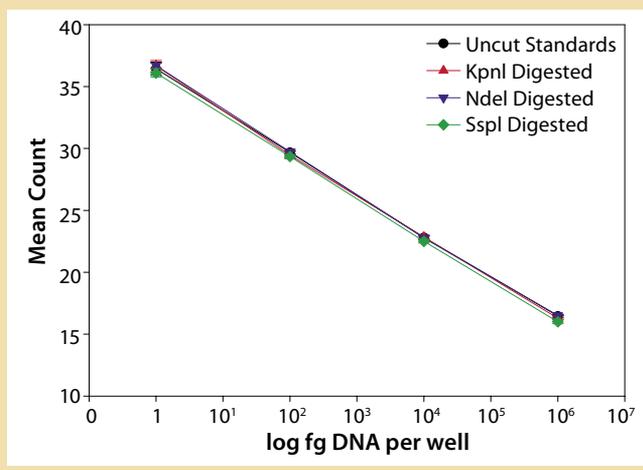
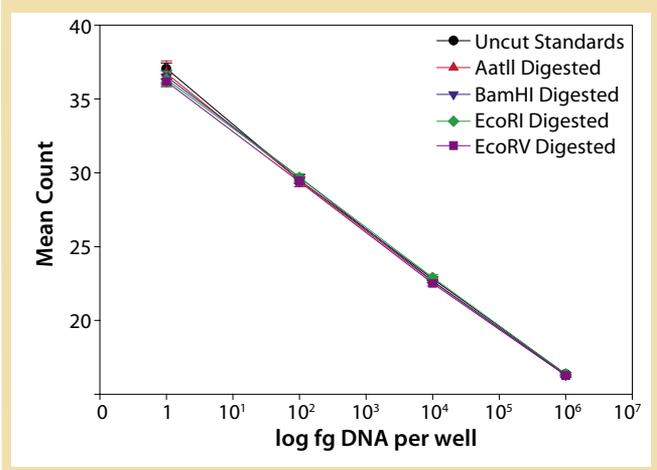


Figure 3B: Intact NS0 DNA as standards and AatII, BamHI, EcoRI, EcoRV digested NS0 DNA as unknowns (B1 probe/primer set)



of 0.18 and a %RSD of 0.64 for a 100-fg DNA input. And the R-repeats assay gave a mean Ct value of 30.09 with a standard deviation of 0.16 and a %RSD of 0.54 for a 100-fg DNA input.

As with the NS0 assays, we also assessed well-to-well variability within a single plate in the fLuc assay using 1×10^4 copies of Hind III linearized pGL3-Basic plasmid DNA per well in the PCR reactions. We found that the mean Ct value is 21.68 with a standard deviation of 0.30 and a %RSD of 1.37.

Effect of NS0 Genomic DNA Size:

Because we had evidence that in-process samples contain a substantial portion of degraded or sheared genomic DNA (data not shown), we surmised that characterization of DNA standards must include a “does template size matter?” component. To this end, an experiment was conducted in which NS0 DNA was subjected to

restriction enzyme digestion with the various cutters listed: AatII, BamHI, EcoRI, EcoRV, KpnI, NdeI, and SspI. Each of the digested DNA samples was diluted to a uniform concentration along with a sample of intact, undigested NS0 genomic DNA.

Four TaqMan assays were run, two for each of the B1, with the intact, undigested NS0 DNA serving as standards and the variously digested DNA as unknowns. As evidenced by Figure 3, none of the restriction enzyme digestions made a noticeable difference in TaqMan signal as compared with the signal generated by intact, undigested NS0 genomic DNA. Similar results were obtained from R-repeats assays (data not shown).

Optimization of Extraction

Conditions: Because TaqMan qPCR assays are conducted in 96-well format, we were interested in

exploring the feasibility of a 96-well DNA extraction method. Qiagen’s QIAamp 96 DNA blood kit provides a high-throughput 96-well DNA extraction format allowing the simultaneous processing of up to 192 samples. The method is based on the selective binding properties of a silica-gel membrane. DNA adsorption occurs during a brief centrifugation step. Following two washing steps, purified DNA is then eluted from the QIAamp 96 plate in low salt Buffer AE. Each well of the QIAamp 96 plate binds nucleic acids greater than 200 bases in length, and purified DNA from intact cells ranges in size up to 50 kb, with fragments of approximately 20–30 kb predominating. (As mentioned earlier, the DNA isolated from in-process samples had more smaller fragments.) DNA in this range undergoes complete denaturation during thermal

cycling and is amplified efficiently.

We devised a JMP-based experimental design to explore optimal extraction conditions with the response of DNA recovery rate. The following scheme and B1 assay were applied to evaluate extraction efficiency:

That scheme was based on no cross-reactivity between the B1 and fLuc assays and their respective DNA. There is no difference between B1 TaqMan signals generated by NS0 DNA alone and NS0 DNA mixed with pGL3-Basic plasmid DNA (data not shown).

Design of Experiments: We performed three DOE (design of experiments, JMP version 5.1, SAS Institute Inc., www.sas.com) to optimize the extraction process. The first experiment was full factorial in design looking at seven variables with two replicates or extraction wells per condition. The fishbone diagram (Figure 4) indicates the parameters addressed in the first DOE. Table 1 describes wash and elution steps for the six extraction plates in terms of centrifugation times and speeds. Plate numbers three (*) and four (*) represent the midpoints in this experiment. This required 64 wells per extraction plate, six extraction plates, and 12 TaqMan runs giving a total of 384 data points.

Each extraction well received 220 μL of extraction solutions, which consist of the following solutions: 20 μL of NS0 genomic DNA (5×10^3 fg) mixed with 5×10^3 copies of the linearized plasmid control DNA, pGL3-Basic. The DNA amounts chosen represented the midpoint of the standard curves used in both TaqMan assays:

- 50 μL of Poly(A) (1 \times , 12 μg), or H_2O
- 50 μL of Poly(dA) (1 \times , 8.89 μg), or H_2O
- 50 μL of carrier RNA (tRNA, 1 \times , 12 μg), or H_2O
- 50 μL of BSA (1 \times , 560 μg), or H_2O .

Following the extractions, eluates from each of the six extraction plates were assayed with both the B1 and fLuc TaqMan probe and primer sets.

Observations coming out of the first DOE indicated that increased wash time resulted in decreased

Figure 4: Cause-and-effect diagram for the extraction step; parameters in red are to be included in the first experimental design

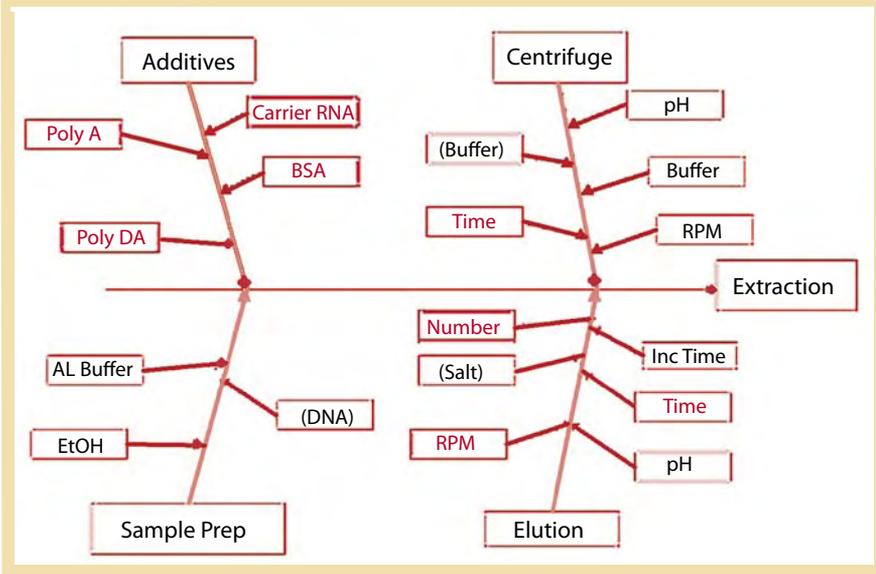
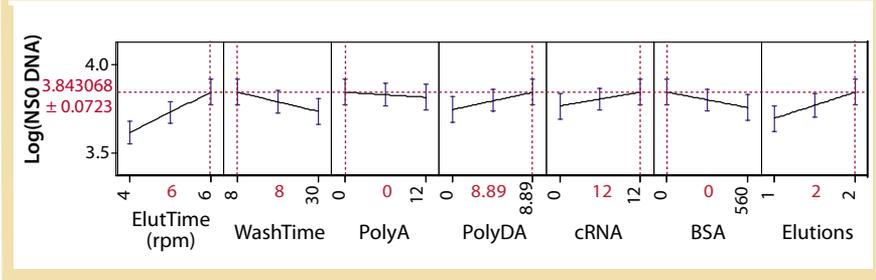


Figure 5: Prediction profiles of the first DOE; dotted vertical red lines indicate predicted best conditions



plasmid DNA recovery and that addition of Poly(dA) yielded better NS0 DNA recovery than did addition of Poly(A). Also, the number of elutions, elution time, and cRNA also improved recovery. Figure 5 shows the optimal conditions with a predicted best recovery of $64.4 \pm 6.3\%$. From this point onward we abandoned exploration of the Poly(A) and BSA additives to extraction mixtures.

The second DOE was performed to further optimize extraction conditions based on the trends observed in the first DOE. The following extraction parameters were varied (with ranges in parenthesis): elution time (8 and 10 minutes); Poly(dA) (1.2, 1.4 and 1.6 \times , related to 1 \times as indicated in the first DOE); and cRNA (1.2, 1.4 and 1.6 \times), while fixing DNA input and holding low value of wash time (Wash 1 at 6,000 rpm for one minute and Wash 2 at 6,000 rpm for eight minutes), with no addition of Poly(A) and BSA.

As expected, recovery levels were on average much higher than in the

first DOE. Poly(dA) and elution time exhibited significant main effects. No significant effect of cRNA was observed in this range. We decided to use the concentration Poly(dA) = 1.4 \times rather than 1.6 \times because the latter increased variability, and there really is not much extraction to gain by going to 1.6 \times (data not shown).

The Third DOE full standard curve range dilutions of both NS0 and the plasmid DNA were subjected to extraction. We further compared 1 \times to 1.4 \times of Poly(dA) with cRNA (1 \times) and applied six-minute elutions either two or three times at 6,000 rpm.

The results showed that every condition holds up for $[\text{DNA}] > 100$ fg/well, whereas at 10 fg/well a marked increase in variability was observed (Figure 6). The recoveries were less dependent on concentration for the 1 \times cRNA/Poly(dA) cases. The 1.4 \times Poly(dA) gave larger variability at 10 fg/well and stronger dependence on concentration. We selected two elutions for easier operation. Table 2 summarizes the final extraction conditions.

Table 1: Wash and elution conditions

Plate	First Wash (6K rpm)	Second Wash (6K rpm)	Elution
1	1 min	8 min	4K rpm/ 4 min
2	4 min	30 min	4K rpm/ 4 min
3*	2 min	15 min	6K rpm/ 4 min
4*	2 min	15 min	6K rpm/ 4 min
5	1 min	8 min	6K rpm/ 6 min
6	4 min	30 min	6K rpm/ 6 min

Table 2: The final extraction conditions

cRNA	Poly dA	First Wash	Second Wash	Elution
1.0x (12 µg/ 50 µL)	1.0x (8.89 mg/ 50 mL)	1 min	8 min	6,000 rpm twice

Assay Repeatability: A total of 56 actual process stream samples from two NS0 cell lines producing monoclonal antibodies and other biomolecules were analyzed under optimal extraction conditions and quantified by the B1 TaqMan assay. The samples were collected from protein A main stream and after various downstream processes at different dilutions. Each sample had four replicates (one per well). A total of 224 data points were generated. As shown in Figure 7A, there is a strong relationship between the standard deviation and the mean of each group of four wells. Further analysis of this relationship indicates that it is linear until ~10 fg/well. A linear relationship between mean and standard deviation implies a constant relative standard deviation (RSD,

also known as coefficient of variability or COV). Figure 7B clearly indicates how the COV remains constant ~20 ± 20%

The constant variability observed in the log- data allows a pooled estimated of the standard deviation. Because this estimate is obtained from the 224 data points, it is much more accurate than each of the individual ones. A pooled standard deviation of the log-transformed data can be used to obtain an RSD of the untransformed data using the following equation from the *US Pharmacopeial Forum* (8):

$$RSD = 100\% \times (10^s - 1)$$

where *s* is the pooled standard deviation of the log-transformed data. The pooled standard deviation in log-scale is 0.0888, which gives an RSD of 22.7%. With that value, the RSD of an average of four wells is estimated to be 10.8%.

Threshold and TaqMan Assay

Equivalency: Fourteen samples of NS0 cell lines and six samples of CHO cell lines from various steps in their purification processes were assayed by both Threshold (two to three replicates) and TaqMan (four replicates) methods to assess their equivalency. Figure 8 contains a linear regression of Threshold and TaqMan analyzed for the NS0 cell lines, both in log-scale, with 95% confidence intervals obtained from the replicated

Figure 6: The third DOE results

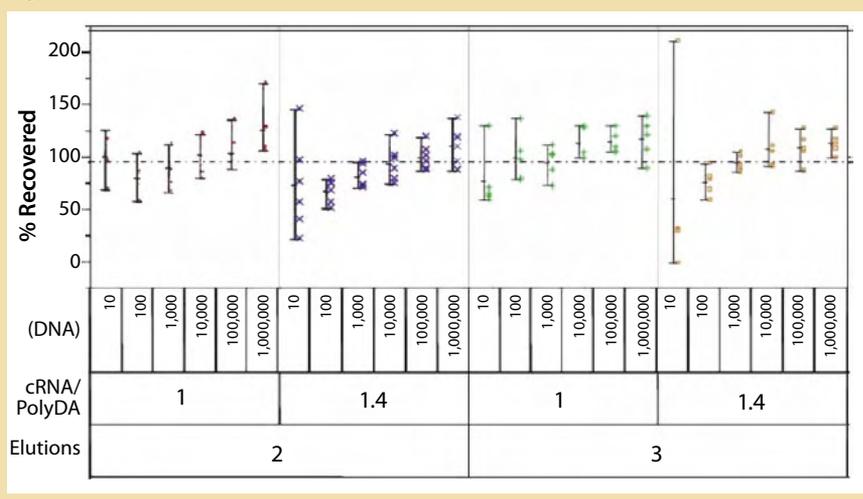


Figure 7A: Relationship between mean and standard deviation for the 56 process samples

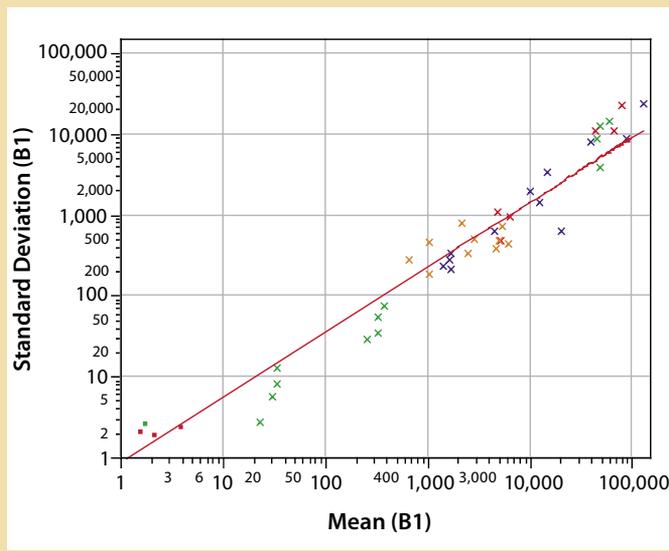


Figure 7B: RSD as a function of the mean DNA level for the 56 process samples

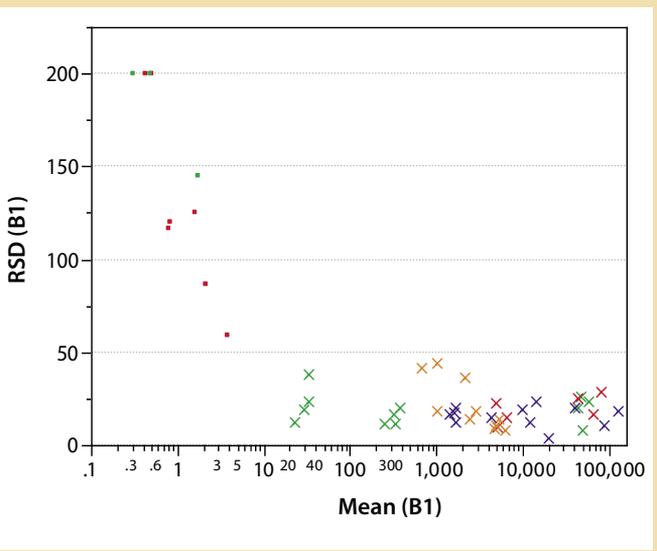
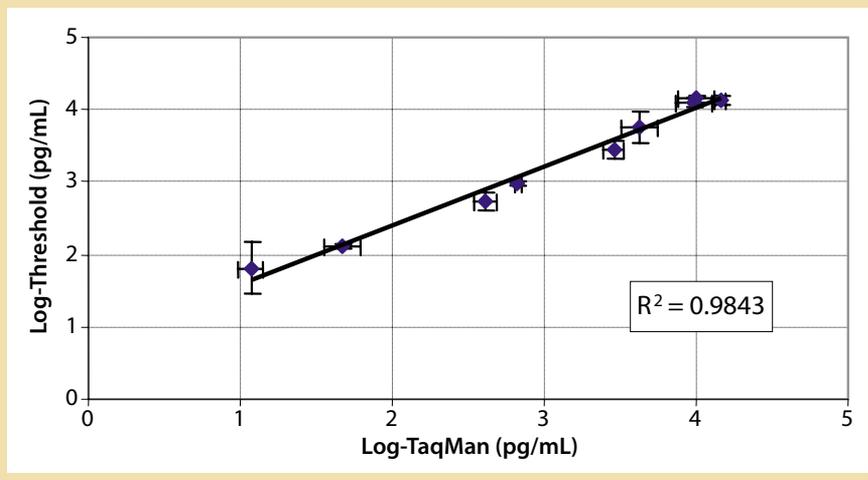


Figure 8: Log-TaqMan compared with log-Threshold correlation

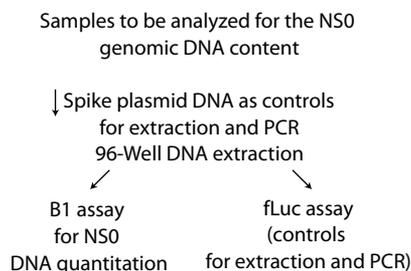


assays. The R^2 is 0.9843, which is higher than that from untransformed data (R^2 is 0.9416, plot is not shown). The analysis with log-transformed data is preferred because it expands the scale at lower DNA concentrations, at which most of the data are located. In addition, both assays showed no increased variability with DNA concentration. A log-transform is usually recommended in this case. A good correlation between the Threshold and the TaqMan CHO49 assay was also obtained (data not shown).

HIGH THROUGHPUT, LOW COST

This article summarizes the development of generic residual NS0 and CHO DNA assays for process stream and drug substance samples using the B1 and R-repeats TaqMan assays and the CHO49 assay. The method is the same for all matrices and requires an upstream DNA extraction procedure using modified Qiagen's QIAamp 96 DNA blood kit.

Under those optimal conditions, we found that recoveries from prepared NS0 samples achieved almost 100% over the 10^1 – 10^6 range of fg DNA. Repeatability of the assay was about 11% (RSD, when four wells per sample are used). Standard curves showed good consistency. Correlation of the B1 TaqMan assay with the Threshold assay was 98%, which demonstrated the equivalency of both assays. The advantages of the TaqMan assay are high throughput and relatively low cost. We are able to



analyze at least 20 samples per day by TaqMan assay, which is equivalent to at least six days of work by Threshold assay. According to Xu and Brorson (9), z of real-time quantitative PCR (TaqMan) technology, including host cell DNA quantification, have been accepted by regulatory agencies for biological safety and quantity evaluation.

ACKNOWLEDGMENTS

We would like to thank the following people for their help during the development: Anna Bedwell, Dayue Chen, Warren Emery, Wen Luo, Renee Lynch, Peter Lambooy, Tami Scott, Doris Weisman, Tracy Williams, Patrick Dunlap, Kamal Egodage, and Matthew Hilton.

REFERENCES

- 1 World Health Organization (WHO). WHO Expert Committee on Biological Standardization: Highlights of the 46th meeting, October 1996. *WHO Weekly Epidemiological Record* 72 (1997): 141–145.
- 2 Krayev AS, et al. The Nucleotide Sequence of the Ubiquitous Repetitive DNA Sequence B1 Complementary to the Most Abundant Class of Mouse Fold-back RNA. *Nuc. Aci. Res.* 8(6) 1980: 1201–1215.
- 3 Gebhard W, et al. A New Family of Interspersed Repetitive DNA Sequences in the Mouse Genome. *J. Mol. Biol.* 157(3) 1982:

453–471.

4 Haynes SR, et al. The Chinese Hamster Alu-Equivalent Sequence: A Conserved Highly Repetitious, Interspersed Deoxyribonucleic Acid Sequence in Mammals Has a Structure Suggestive of a Transposable Element. *Mol. Cell. Biol.* 1(7) 1981: 573–583.

5 Miles A, Meuth M. G-Repeats: A Novel Hamster SINE Family. *Nuc. Aci. Res.* 17(18) 1989: 7221–7228.

6 Zhan, et al. Detection of Minute Virus of Mice Using Real Time Quantitative PCR in Assessment of Virus Clearance During the Purification of Mammalian Cell Substrate Derived Biotherapeutics. *Biologicals* 30(4) 2002: 259–270.

7 Heid C, et al. Real Time Quantitative PCR. *Genome Methods* 6 (1996): 986–994.

8 *US Pharmacopeial Forum* (USP 1010). 30(1) 2004: 241.

9 Xu Y, Brorson K. An Overview of Quantitative PCR Assays for Biological Quantity and Safety Evaluation. *Process Validation for Manufacturing of Biologics and Biotechnology Products*. Brown F, Lubiniecki AS (Eds), vol. 113, *Dev. Bio.* Karger: Basel, Switzerland, 2003; 89–98. 🌐

Corresponding author **Luhong He** is a senior research scientist at Eli Lilly and Company, Indianapolis, IN 46285; 1-317-651-2334, helu@lilly.com. * **Ellen Monson** was a senior research scientist at Eli Lilly and Company and led the NS0 assay and initiated the CHO assay until her departure. **Guillermo Miro-Quesada** is a senior research scientist, and **John Calley** is a principal research scientist. **Daryl Venable** was an assistant senior bioanalytical chemist at Eli Lilly and Company.